



**HAL**  
open science

# Etude structurale à haute résolution spatio-temporelle des effets des modulateurs d'assemblage de capsidie du virus de l'hépatite B

Kalouna Kra

► **To cite this version:**

Kalouna Kra. Etude structurale à haute résolution spatio-temporelle des effets des modulateurs d'assemblage de capsidie du virus de l'hépatite B. Biophysique [physics.bio-ph]. Université Paris-Saclay, 2023. Français. NNT : 2023UPASQ066 . tel-04616401

**HAL Id: tel-04616401**

**<https://theses.hal.science/tel-04616401>**

Submitted on 18 Jun 2024

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

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

# Etude structurale à haute résolution spatio-temporelle des effets des modulateurs d'assemblage de capsid du virus de l'hépatite B

*Structural study at high spatio-temporal resolution of the effects of hepatitis B virus capsid assembly modulators*

## Thèse de doctorat de l'université Paris-Saclay

École doctorale n° 569, Innovation Thérapeutique : du Fondamental à l'Appliqué (ITFA)  
Spécialité de doctorat : Biologie Structurale  
Graduate School : Health and Drug Sciences. Référent : Faculté de pharmacie

Thèse préparée dans les unités de recherche, **Institute for Integrative Biology of the Cell (I2BC) (Université Paris-Saclay, CEA, CNRS) et Laboratoire de Physique des Solides (Université Paris-Saclay, CNRS)**, sous la direction de **Stéphane BRESSANELLI**, directeur de recherche, et la co-direction de **Guillaume TRESSET**, directeur de recherche

Thèse soutenue à Paris-Saclay, le 11 décembre 2023, par

**Kalouna KRA**

## Composition du Jury

Membres du jury avec voix délibérative

<b>Sylvie NESSLER</b> Professeure, Université Paris-Saclay	Présidente
<b>Cendrine MOSKALENKO</b> Professeure, Ecole normale supérieure de Lyon	Rapporteuse & Examinatrice
<b>David DURANTEL</b> Directeur de recherche, Université Lyon 1	Rapporteur & Examineur
<b>Pau BERNADO</b> Directeur de recherche, Université de Montpellier	Examineur

**Titre :** Etude structurale à haute résolution spatio-temporelle des effets des modulateurs d'assemblage de capsid du virus de l'hépatite B

**Mots clés :** HBV Core, Modulateurs d'assemblage, Cinétique, Diffusion des rayons X résolue en temps, Cryomicroscopie électronique

**Résumé :** Le virus de l'hépatite B (VHB) est l'un des plus sérieux pathogènes humains et les traitements actuellement disponibles permettent uniquement de contrôler l'infection mais pas d'en guérir. Les modulateurs d'assemblage sont des candidats au développement de nouveaux médicaments anti-VHB. C'est une nouvelle classe de molécules qui perturbent l'assemblage en capsides. Ces molécules sont également des inhibiteurs puissants de la réplication du VHB *in vitro*. Les mécanismes d'action des modulateurs ont été dégagés à partir d'études *in tubo* sur la protéine de la capsid, la protéine Core. Cependant, l'action des modulateurs sur les chemins d'assemblage et les intermédiaires hors équilibre qui conditionnent la formation des capsides demeurent inconnus. Dans cette thèse, une approche innovante combinant physique des objets biologiques et virologie structurale a été mise en œuvre pour étudier à haute résolution spatiale

et temporelle, l'influence de différents modulateurs sur l'assemblage de capsides, formées à partir de trois protéines recombinante de Core. Ce travail combine ainsi diffusion des rayons X aux petits angles résolue en temps sur sources synchrotron et cryomicroscopie électronique. Grâce à ces techniques, nous avons pu montrer que les modulateurs accélèrent la cinétique d'assemblage des capsides du virus de l'hépatite B et conduisent à une modification de la structure des objets formées en fonction du type de modulateurs et de sa concentration lors de l'assemblage. Nous avons également analysé les effets d'un nouveau type de modulateur qui conduirait à la formation d'un nombre plus faible de capsides. Enfin, nous montrons que l'ARN pré-génomique influence également la vitesse d'assemblage des capsides.

**Title :** Structural study at high spatio-temporal resolution of the effects of hepatitis B virus capsid assembly modulators

**Keywords :** HBV Core, Capsid assembly modulators, Kinetics, Time-resolved SAXS, Cryogenic electron microscopy

**Abstract :** Hepatitis B virus (HBV) is one of the most serious human pathogens, and currently available treatments can only control infection, not cure it. Assembly modulators are candidates for the development of new anti-HBV drugs. They are a new class of molecules that disrupt capsid assembly. These molecules are also powerful inhibitors of HBV replication *in vitro*. The mechanisms of action of these modulators have been elucidated by *in tubo* studies on the capsid protein Core. However, the action of modulators on the assembly pathways and non-equilibrium intermediates that condition capsid formation remains unknown. In this thesis, an innovative approach combining physics of biological objects and structural virology was used to study, at high spatial and temporal resolution, the influence of different modulators on the assembly of capsids formed from three recombinant Core proteins.

This work combines time-resolved small-angle X-ray scattering on synchrotron sources and cryo-electron microscopy. Thanks to these techniques, we were able to show that modulators accelerate the assembly kinetics of hepatitis B virus capsids and lead to a modification of the structure of the objects formed as a function of the type of modulator and its concentration during assembly. We have also analyzed the effects of a new type of modulator which would lead to the formation of a lower number of capsids. Finally, we show that pregenomic RNA also influences capsid assembly speed.

## Remerciements

Je remercie tout d'abord mes directeurs de thèse, Stéphane Bressanelli et Guillaume Tresset. Durant ces trois années de thèse, vous avez su me guider, me conseiller, me transmettre votre savoir, ce qui m'a permis d'en apprendre un peu plus chaque jour. Je vous remercie pour la confiance et l'autonomie accordées, deux paramètres qui ont été indispensables au bon déroulé de ce projet. Merci également d'avoir sacrifié vos soirées, week-end et nuits pour les expériences de diffusion au synchrotron. J'en suis extrêmement reconnaissante.

Ensuite, un grand merci à Laetitia Gargowitsch et Jéril Degrouard pour leur appui lors des expériences réalisées tout au long de cette thèse. Laetitia, ton aide a très été précieuse et indispensable pour la purification des protéines et de l'ARN. Je me souviendrais également de toutes nos discussions au bureau ! Quant à Jéril, je te remercie particulièrement pour ton aide lors des expériences de microscopie. Les débuts ont été laborieux et on a passé énormément de temps avant d'enfin réussir à observer ces capsides mais nous y sommes arrivés ! Je retiendrais également de ces journées de travail, ta bonne humeur, ton humour et ta patience infaillible !

Je remercie également Cendrine Moskalenko et David Durantel de m'avoir fait l'honneur d'être les rapporteurs de mes travaux de thèse. Merci également à la présidente du jury, Sylvie Nessler et à Pau Bernado d'avoir accepté d'être jury dans mon comité de suivi de thèse ainsi qu'à ma soutenance de thèse.

J'aimerais par ailleurs remercier tous les personnes avec qui j'ai collaboré pendant ma thèse. Merci à :

- Fabien Montel pour son accueil dans son laboratoire de recherche à Lyon et la mise à disposition de son matériel. A ce titre, je remercie aussi Léa Chazot-Franguiadakis qui a pris le

temps de me montrer ses expériences lors de mes déplacements. J'en garde d'excellents souvenirs ;

- Roya Zandi et Siyu Li pour leurs expériences de simulation d'assemblage de la capsid et pour leur participation à la rédaction de mon article scientifique en tant que premier auteure ;
- Frank Amblard et Raymond Schinazi qui m'ont fourni deux des quatre modulateurs d'assemblage que j'ai utilisé durant ma thèse ;
- Elizaveta Bazerbe, ma seule et unique stagiaire à ce jour, qui m'a beaucoup aidée lors de son stage, sur les expériences d'assemblage de la Cp183 ;
- Aurélien Thureau et Javier Pérez pour leur aide lors de mes nombreuses nuits à SOLEIL ;
- Lauren Matthews pour son aide indispensable lors des expériences de cinétique d'assemblage à l'ESRF.

Je souhaite remercier mes collègues thésards : Adrien Royet, Thomas Bugea et Rémi Ruedas. Merci pour tous nos échanges, les fous rires, les cinés, les soirées et les nuits au synchrotron ! Merci également à Thibault Tubiana pour ton humour et pour avoir sacrifié ton week-end pour m'accompagner à l'ESRF faire des cinétiques d'assemblage.

Je remercie aussi les « Master brothers » : Sana Aimeur, Salim Bouhazam et Charles Cirio. Merci pour votre amitié et soutien depuis toutes ces années ! Un grand merci également à Diana Saounera, Marie-Widning Dumé et Gwendoline Thieblin qui m'ont permis de penser à autre chose qu'à la thèse. Merci aussi pour votre amitié et encouragements depuis toutes ces années.

Merci également à tous les membres de l'équipe SOBIO et l'équipe IMAPP. Merci pour votre humour, pour les journées « tri des germes de blés » et les soirées au GDR Médyna. Merci à : Marion Badot, Yves Boulard, Sonia Fieulaine, Virginie Gervais, Marie-Hélène le Du,

Amélie Leforestier, Renata Naporano Bicev, Maïté Paternostre, Laura Pieri, Tamizhmalar Sundararajan, Fatima Taiki, Neha Tripathi, et Kahina Vertchik. Mention spéciale à Sonia Fieulaine, véritable guide sans qui je serais perdue au labo.

Je remercie l'Agence Nationale de Recherche sur le Sida et les hépatites virales pour le financement de mes recherches.

Je tiens aussi à remercier Hélène Débat, Marc Nadal et Julie Menetrey de m'avoir fait découvrir le monde de la recherche, ce qui a sans conteste, eu un impact sur mon orientation professionnelle.

Je remercie tout particulièrement Jean Yves Ouattara pour m'avoir accompagné tout au long de cette aventure. Merci pour cette assistance sans faille et de m'avoir permis d'aborder cette étape de ma vie sereinement.

Un grand merci aux membres de ma famille : Jean-Charles Kra, Jean-Marc Kra, Pétronille Ottavi Kra et Carelle Kra. Merci pour votre soutien et d'avoir été à l'écoute. Carelle, merci d'avoir participé à l'édition de ce document.

Enfin j'ai une pensée particulière pour Marcel Ottavi et pour Victor Antoine Ottavi. Au premier, merci beaucoup pour ton aide sans laquelle cette thèse aurait été très compliqué, au deuxième, tu avais vu juste 😊.

# Communications scientifiques

## Articles

- **Structures and kinetic assembly pathways of hepatitis B virus capsids in the presence of class E modulators.** Kalouna Kra, Laetitia Gargowitsch, Ana Arteni, Jéril Degrouard, Javier Pérez, Theyencheri Narayanan, Stéphane Bressanelli et Guillaume Tresset, *en préparation*

**Contribution** : Préparation des échantillons, réalisation des expériences de photométrie de masse, de SAXS, de TR-SAXS et de microscopie électronique, analyse des données, réalisation des figures, participation à l'écriture de l'article.

- **Energetics and kinetic assembly pathways of hepatitis B virus capsids in the presence of antivirals.** Kalouna Kra, Siyu Li, Laetitia Gargowitsch, Jéril Degrouard, Javier Pérez, Roya Zandi, Stéphane Bressanelli, et Guillaume Tresset, *ACS nano* 2023, <https://doi.org/10.1021/acsnano.3c03595>

**Contribution** : Préparation des échantillons, réalisation des expériences de SAXS et TR-SAXS et de cryomicroscopie électronique, analyse des données, réalisation des figures, participation à l'écriture de l'article.

- **Relationships between RNA topology and nucleocapsid structure in a model icosahedral virus.** Laurent Marichal, Laetitia Gargowitsch, Rafael Leite Rubim, Christina Sizun, Kalouna Kra, Stéphane Bressanelli, Yinan Dong, Sanaz Panahandeh, Roya Zandi, et Guillaume Tresset, *Biophysical Journal* 2021, <https://doi.org/10.1016/j.bpj.2021.08.021>

**Contribution** : Analyse des images de cryomicroscopie électronique.

## Présentations orales

- Réunion annuelle AC42 Réseau National Hépatites (Paris, France, 2023)
- Les Journées Francophones de Virologie (Paris, France, 2023)
- Réunion annuelle AC42 Réseau National Hépatites (France, 2022)
- 66th Biophysical Society Annual Meeting (San Francisco, USA, 2022)
- Réunion annuelle AC42 Réseau National Hépatites (Paris, France, 2021)

## Poster

- Gordon Research conference : Physical Virology (Lucca, Italy, 2023)
- The international Conference on Antiviral Research (Lyon, France, 2023)
- Les Journées Francophones de Virologie (France, 2021)

# Table des matières

Chapitre 1 : Introduction générale .....	1
Bibliographie .....	4
Chapitre 2 : Le virus de l'Hépatite B et les modulateurs d'assemblage.....	6
1. Généralités sur les virus .....	6
1.1 Description.....	6
1.2 Classification.....	7
1.3 Capsides virales et nombre de triangulation .....	8
2. Le virus de l'hépatite B .....	11
2.1 Découverte du virus de l'hépatite B .....	11
2.2 Génotypes et sérotypes .....	11
2.3 Prévalence.....	13
2.4 Transmission et prévention .....	14
2.5 Symptômes et traitements .....	15
3. Cycle viral du virus de l'hépatite B.....	17
4. Les différentes protéines du virus de l'hépatite B .....	19
5. La protéine Core.....	23
5.1 Structure de la protéine entière.....	23
5.2 Fonctions de la protéine Core .....	24
5.3 Domaine C-terminal .....	24
5.4 Autoassemblage de la capside .....	25
6. Les modulateurs d'assemblage.....	27
6.1 Mécanisme d'action .....	27
6.2 Les modulateurs d'assemblage de type A .....	29
6.3 Les modulateurs d'assemblage de type E .....	30
Bibliographie .....	31

Chapitre 3 : Principes des différentes méthodes utilisées.....	44
1. Expression et purification des protéines.....	44
1.1 Transformation et expression en système bactérien.....	44
1.2 Purification des protéines par chromatographie d'exclusion de taille.....	47
2. Méthodes de caractérisation optiques.....	48
2.1 Diffusion dynamique de la lumière.....	48
2.2 Photométrie de masse.....	53
3. Diffusion des rayons X aux petits angles et la diffusion des rayons X aux petits angles résolue dans le temps.....	54
3.1 Principe de la diffusion des rayons X aux petits angles.....	55
3.2 Instrumentation.....	58
3.3 Analyse des courbes SAXS.....	60
3.4 Diffusion des rayons X aux petits angles résolue dans le temps.....	67
4. Microscopie électronique et traitement d'images.....	70
4.1 Anatomie d'un microscope électronique.....	70
4.2 Contraste en microscopie électronique.....	73
4.3 Préparation de l'échantillon.....	76
4.4 Traitement de l'image : analyse de particules uniques.....	77
Bibliographie.....	79
Chapitre 4 : Effets des modulateurs d'assemblage sur la cinétique et la structure des capsides formées à partir de la protéine Cp149.....	82
1. Introduction.....	82
2. Assemblage de la protéine cp149 en absence de modulateurs.....	83
2.1 Expression et purification de la protéine dimérique cp149.....	83
2.2 En présence d'acétate d'ammonium, la protéine dimérique cp149 s'autoassemble en capside T=4.....	88
3. Article : <i>Energetics and Kinetics Assembly Pathway of Hepatitis B Virus Capsids in the Presence of Antivirals</i> .....	93

3.1	Introduction .....	93
3.2	Résumé de l'article .....	123
3.3	Complément de l'article : Caractérisation par DLS et microscopie électronique des capsides formées en présence des modulateurs de type A et de type E .....	123
4.	Assemblage de la protéine cp149 en présence de deux nouvelles molécules GLP-26 et RS-3427 .....	127
4.1	Introduction .....	127
4.2	<i>In vitro</i> , l'assemblage en présence des deux molécules conduit à la formation d'une grande majorité de capsides T=4 .....	129
4.3	GLP-26 accélère la cinétique d'assemblage et induit des changements structuraux.....	132
4.4	Contrairement à GLP-26, RS-3427 n'accélère pas la formation des capsides et entraîne de légères modifications structurales.....	137
4.5	Conclusion : Les modulateurs de type CAM-E ; une classe diverse et variée	145
5.	Conclusion du chapitre .....	146
6.	Matériels et méthodes .....	148
	Bibliographie .....	158
	Chapitre 5 : Effets des modulateurs d'assemblage sur la cinétique et la structure des capsides formées à partir des protéines Cp183 et des protéines Cp183-EEE.....	159
1.	Introduction.....	159
2.	Un pH basique et le chlorure de guanidinium sont nécessaires pour maintenir la protéine Cp183-EEE sous sa forme dimérique .....	160
3.	<i>In vitro</i> , les capsides formées à partir des protéines Cp183 et Cp183-EEE sont de type T=3 .....	164
4.	Les modulateurs d'assemblage perturbent la structure des capsides Cp183 et Cp183-EEE .....	169
5.	Conclusion du chapitre .....	174

6. Matériels et méthodes .....	176
Bibliographie .....	182
Chapitre 6 : Conclusions et perspectives .....	183
Bibliographie .....	188
Annexes .....	190

## Liste des figures

Figure 2.1 : Schéma simplifié d'un virus nu (à gauche sur l'image) et d'un virus enveloppé (à droite sur l'image).....	6
Figure 2.2 : Classification de Baltimore .....	7
Figure 2.3 : Description des axes de symétrie dans une capsidie icosaédrique .....	9
Figure 2.4 : Exemple de calcul du nombre de triangulation T et des paramètres h et k pour une capsidie de symétrie icosaédrique T=4 (A) et de symétrie icosaédrique T=3 (B) .....	9
Figure 2.5 : Variabilité des géotypes et sérotypes du virus de l'hépatite B.....	12
Figure 2.6 : Prévalence du virus de l'hépatite B.....	13
Figure 2.7 : Evolution des symptômes causés par le virus de l'hépatite B, vers le carcinome hépatocellulaire.....	15
Figure 2.8 : Cycle viral du virus de l'hépatite B .....	17
Figure 2.9 : Génome du virus de l'hépatite B .....	19
Figure 2.10 : Topologie des protéines d'enveloppe .....	20
Figure 2.11 : Structure de la protéine AgHBe .....	21
Figure 2.12 : Structure de la polymérase du virus de l'hépatite B .....	22
Figure 2.13 : Dimère de la protéine Core .....	23
Figure 2.14 : Domaine C-terminal (CTD) de la protéine Core .....	24
Figure 2.15 : Capsidie du virus de l'hépatite B.....	26
Figure 2.16 : Effet des modulateurs d'assemblage (CAMs) sur le cycle répliatif du VHB.....	28
Figure 2.17 : Structure chimique des modulateurs d'assemblage les plus courants	29
Figure 3.1 : Schéma d'un vecteur d'expression .....	45
Figure 3.2 : Principe de la chromatographie d'exclusion de taille.....	47
Figure 3.3 : Principe de la diffusion dynamique de la lumière .....	49
Figure 3.4 : Fonction d'autocorrélation .....	50
Figure 3.5 : Distribution en nombre, volume et intensité .....	52
Figure 3.6 : Principe de la photométrie de masse. ....	53
Figure 3.7 : Vecteur d'onde .....	55
Figure 3.8 : Motif de diffusion .....	57

Figure 3.9 : Schéma d'un synchrotron.....	59
Figure 3.10 : Soustraction du tampon .....	61
Figure 3.11 : Représentation de Guinier.....	63
Figure 3.12 : Fonction de distribution de distance.....	65
Figure 3.13 : Représentation de Kratky.....	66
Figure 3.14 : Stopped-flow .....	68
Figure 3.15 : Schéma d'un microscope .....	71
Figure 3.16 : La fonction de transfert de contraste .....	74
Figure 3.17 : Méthodes de préparation des grilles de microscopie électronique .....	76
Figure 4.1 : Gel SDS-PAGE 14% issu de l'étape de centrifugation sur un gradient de sucrose des capsides de Cp149 .....	84
Figure 4.2 : Gel SDS-PAGE 14% issu de la purification des capsides de Cp149 ....	84
Figure 4.3 : Chromatogrammes de purification par exclusion de taille de la Cp149.	85
Figure 4.4 : Gel SDS-PAGE 14% issu de la purification des dimères de Cp149.....	87
Figure 4.5 : Mesure par DLS des dimères de la Cp149 obtenus après la purification..	88
Figure 4.6 : Mesures par DLS des objets formés après l'assemblage de la Cp149 dans différents tampons.....	89
Figure 4.7 : Micrographies des capsides assemblées dans différents tampons et observées en coloration négative au microscope électronique .....	90
Figure 4.8 : Courbes SAXS des capsides assemblées dans différents tampons à pH 9 ou à pH 7,5.....	91
Figure 4.9 : Analyse par photométrie de masse de la distribution de masse de la protéine Cp149.....	92
Figure 4.10 : Mesures par DLS des objets formés après l'assemblage de la Cp149 en présence des modulateurs à différentes concentrations .....	124
Figure 4.11 : Micrographies des objets formés lors de l'assemblage de la Cp149 à partir des dimères de la Cp149 et en présence de modulateurs à différentes concentrations.....	125
Figure 4.12 : Assemblage de la capside Cp149 en présence de la CAM-A pendant 24 h .....	126
Figure 4.13 : Structures chimiques de GLP-26 (1) et de RS-3427 (2).....	127
Figure 4.14 : Caractérisation optique des objets formés en présence de GLP-26 et RS-3427 .....	129

Figure 4.15 : Courbes SAXS des capsides assemblées en présence des deux molécules .....	131
Figure 4.16 : Mesures TR-SAXS de l'assemblage de capsides Cp149 en présence de la molécule GLP-26.....	132
Figure 4.17 : Micrographies des capsides assemblées en présence ou non de GLP-26 et observées en microscopie électronique en coloration négative .....	134
Figure 4.18 : Données CryoEM obtenues pour l'assemblage des capsides en présence de GLP-26 .....	135
Figure 4.19 : Mesures TR-SAXS de l'assemblage de capsides Cp149 en présence de la molécule RS-3427 .....	137
Figure 4.20 : Micrographies des capsides assemblées en présence ou non de RS-3427 et observées en microscopie électronique en coloration négative .....	138
Figure 4.21 : Données CryoEM obtenues pour l'assemblage des capsides en présence de 90 $\mu$ M de RS-3427 .....	139
Figure 4.22 : Données cryoEM obtenues pour l'analyse des capsides en présence de RS-3427 et présentant des défauts.....	141
Figure 4.23 : Reconstruction de la structure atomique de la capsid Cp149 .....	143
Figure 4.24 : Mesures TR-SAXS de l'assemblage de capsides Cp149 en présence de 90 $\mu$ M de JNJ-632 .....	145
Figure 5.1 : Mesures par DLS du diamètre hydrodynamique dimère de Cp183-EEE dans différents tampons .....	161
Figure 5.2 : Mesures par DLS du diamètre hydrodynamique du dimère de Cp183-EEE à différentes concentrations .....	163
Figure 5.3 : Mesures TR-SAXS de l'assemblage de capsides de Cp183-EEE .....	164
Figure 5.4 : Mesures par DLS des objets formés après l'assemblage de la Cp183 et de la Cp183-EEE en présence ou non d'ARNpg.....	166
Figure 5.5 : Micrographies des objets formés lors de l'assemblage de la Cp183 et de la Cp183-EEE et observés en microscopie électronique par coloration négative ..	167
Figure 5.6 : Courbes SAXS des capsides formées en présence de l'ARNpg.....	168
Figure 5.7 : Micrographies des objets formés lors de l'assemblage de la Cp183 et observés en microscopie électronique par coloration négative .....	171
Figure 5.8 : Courbes SAXS des objets formés en présence des modulateurs.....	172

## Liste des tableaux

Tableau 3.1 : Récapitulatif des caractéristiques des lignes de lumières ID02 et SWING .....	59
Tableau 5.1 : Diamètre hydrodynamique selon l'intensité et le nombre de particules mesurées .....	169

## Liste des abréviations

**AAs** Acide aminés

**ACF** Fonction d'autocorrélation

**ADN** Acide désoxyribonucléique

**ADNccc** Acide désoxyribonucléique circulaire covalamment clos

**ADNrc** Acide désoxyribonucléique circulaire relâché

**AmAc** Acétate d'ammonium

**ARN** Acide ribonucléique

**ARNm** Acide ribonucléique messenger

**ARNpg** Acide ribonucléique pré-génomique

**CAMs** Capsid assembly modulators

**CAM-A** Capsid assembly modulators de type A

**CAM-E** Capsid assembly modulators de type E

**Cp149** Core protein 149

**Cp183** Core protein 183

**CTD** C-terminal domain

**CTF** Fonction de transfert de contraste

**DLS** Diffusion dynamique de la lumière

**DTT** Dithiothréitol

**FDA** Food drug administration

**GLP** Glyoxamoylpyrroloxamide

**GPA** Glyoxamide-pyrrolamide

**HAPs** Hétéroaryldihydropyrimidines

**HCC** Carcinome hépatocellulaire

**MCS** Site multiple de clonage

**NES** Séquence d'export nucléaire

**NLS** Séquence de localisation nucléaire

**NTCP** Sodium taurocholate cotransporting polypeptide

**NTD** N-terminal domain

**PEG-IFN** Interféron pégylé

**PPA** Phénylpropénamide

**PDI** Indice de polydispersité

**PSF** Point spread function

**RNaseH** Ribonucléase H

**SAXS** Diffusion des rayons X aux petits angles

**SBA** Sulfamoylbenzamide

**SEC** Size exclusion chromatography

**SVP** Sous-particule virale

**TP** Terminal protein

**TR-SAXS** Diffusion des rayons X aux petits angles résolue en temps

**VHB** Virus de l'hépatite B

# Chapitre 1 : Introduction générale

Le virus de l'hépatite B est un enjeu crucial de santé publique. Malgré l'existence d'un vaccin sûr et efficace (WHO 2020; Pattyn et al. 2021), plus de 292 millions de personnes dans le monde sont infectées de façon chronique par ce virus. Les traitements disponibles actuellement sont les monothérapies par pegInterféron-alpha ou par analogue de nucléosides (Taverniti et al. 2022). Cependant, ces traitements ne sont pas satisfaisants puisqu'ils permettent uniquement de contrôler l'infection, mais pas d'en guérir et doivent être pris à vie le plus souvent. Il est donc nécessaire de développer de nouveaux médicaments.

Le virus de l'hépatite B est un virus enveloppé avec une nucléocapside icosaédrique contenant le génome viral. La réplication du génome viral est extrêmement complexe et implique plusieurs compartiments cellulaires différents. Durant son cycle viral, un des ARN messagers du virus de l'hépatite B appelé ARN pré-génomique, va être encapsidé au sein de nucléocapsides. La capsid joue donc un rôle essentiel dans le cycle viral du virus de l'hépatite B et le constituant de cette capsid est la protéine Core. C'est une protéine dimérique de 183 (ou 185 selon les souches) acides aminés composée de deux domaines. Les 149 résidus N-terminaux forment un domaine pouvant s'autoassembler, *in vivo* et *in vitro* en capsides icosaédriques  $T = 4$  (120 dimères). Les 34 (ou 36) résidus restants forment un domaine intrinsèquement désordonné appelé le domaine C-terminal. Ce domaine interagit avec l'ARN pré-génomique et est nécessaire pour son encapsidation (Zlotnick et al. 2015; Venkatakrisnan et Zlotnick 2016; Nassal 1992).

Parmi les nouvelles molécules développées pour traiter l'hépatite B, les modulateurs d'assemblage sont des molécules dirigées spécifiquement contre la protéine Core (Taverniti et al. 2022). Ces molécules se lient à l'interface entre les dimères de Core et perturbent ainsi l'assemblage en capsid. Selon leur mode d'action, deux classes de modulateurs ont été définies :

- Les molécules de la classe A (pour *Aberrant*) qui favorisent la formation des très grosses structures non capsidiques et pléiomorphes.

- Les molécules de la classe E (pour *Empty*) qui promeuvent l'assemblage de capsides indistinguables des capsides normales en microscopie électronique à coloration négative, mais ne contenant pas d'ARN prégénomique.

De nombreuses études (Katen et al. 2013; Venkatakrishnan et al. 2016; Bourne et al. 2008) sur l'effet des modulateurs d'assemblage ont été réalisées *in vitro* afin de définir leurs mécanismes d'action, et leurs effets sur la capside du virus de l'hépatite B. Une très grande partie des données structurales, ont été obtenues à l'équilibre, dans des conditions où les capsides sont d'abord assemblées avant l'ajout des modulateurs. Cependant, les effets cinétiques et structuraux des modulateurs lors de l'assemblage de la capside n'ont pas encore été étudiés.

Ce projet de thèse avait donc pour but d'étudier par des méthodes biophysiques et structurales, les voies d'assemblage de la capside du virus de l'hépatite B et l'impact des modulateurs d'assemblage sur ces processus. Pour cela, trois capsides recombinantes de Core, la Cp149, la Cp183 et la Cp183-EEE ainsi que l'ARN (acide ribonucléique) prégénomique ont été produits et purifiés. Puis les capsides ont été désassemblées, et les dimères dissociés et purifiés. Ensuite, les expériences d'assemblage des capsides avec ou sans modulateurs d'assemblage et/ou sans ARN ont été réalisées. Enfin, les analyses structurales et de la cinétique d'assemblage de ces capsides ont été effectuées. Je voulais ainsi caractériser à haute résolution spatio-temporelle les chemins cinétiques d'assemblage et l'influence des modulateurs d'assemblage sur la cinétique mais également sur la structure des capsides formées. Cette étude repose principalement sur deux méthodes : la diffusion des rayons X aux petits angles, statique et résolue en temps (*Time-Resolved Small-Angle X-Ray Scattering*, TR-SAXS) et la cryomicroscopie électronique en transmission.

Cette thèse sera organisée comme suit. Dans le deuxième chapitre, je ferai une revue de la littérature concernant le virus de l'hépatite B, le processus d'autoassemblage de sa capside ainsi que sur les modulateurs d'assemblage. Le troisième chapitre sera consacré aux principes des méthodes biochimiques, biophysiques et structurales utilisées durant cette thèse. Ensuite, le chapitre 4 exposera les résultats de notre étude concernant les effets des modulateurs d'assemblage sur la capside du virus de l'hépatite B formée à partir de dimères de la protéine Cp149. Cette protéine contient uniquement le domaine d'autoassemblage de la protéine Core. S'en suivra le chapitre

5, dédié à notre étude de l'assemblage de la protéine Core entière en présence d'ARN pré-génomique et de modulateurs d'assemblage. Enfin, dans le chapitre 6, une discussion sur les résultats et la conclusion générale seront présentées.

## Bibliographie

Bourne, Christina, Sejin Lee, Bollu Venkataiah, Angela Lee, Brent Korba, M. G. Finn, et Adam Zlotnick. 2008. « Small-Molecule Effectors of Hepatitis B Virus Capsid Assembly Give Insight into Virus Life Cycle ». *Journal of Virology* 82 (20): 10262-70. <https://doi.org/10.1128/JVI.01360-08>.

Katen, Sarah P., Zhenning Tan, Srinivas Reddy Chirapu, M. G. Finn, et Adam Zlotnick. 2013. « Assembly-Directed Antivirals Differentially Bind Quasiequivalent Pockets to Modify Hepatitis B Virus Capsid Tertiary and Quaternary Structure ». *Structure (London, England: 1993)* 21 (8): 1406-16. <https://doi.org/10.1016/j.str.2013.06.013>.

Nassal, M. 1992. « The Arginine-Rich Domain of the Hepatitis B Virus Core Protein Is Required for Pregenome Encapsidation and Productive Viral Positive-Strand DNA Synthesis but Not for Virus Assembly ». *Journal of Virology* 66 (7): 4107-16. <https://doi.org/10.1128/JVI.66.7.4107-4116.1992>.

Pattyn, Jade, Greet Hendrickx, Alex Vorsters, et Pierre Van Damme. 2021. « Hepatitis B Vaccines ». *The Journal of Infectious Diseases* 224 (Suppl 4): S343-51. <https://doi.org/10.1093/infdis/jiaa668>.

Taverniti, Valerio, Gaëtan Ligat, Yannick Debing, Dieudonne Buh Kum, Thomas F. Baumert, et Eloi R. Verrier. 2022. « Capsid Assembly Modulators as Antiviral Agents against HBV: Molecular Mechanisms and Clinical Perspectives ». *Journal of Clinical Medicine* 11 (5): 1349. <https://doi.org/10.3390/jcm11051349>.

Venkatakrishnan, Balasubramanian, Sarah P. Katen, Samson Francis, Srinivas Chirapu, M. G. Finn, et Adam Zlotnick. 2016. « Hepatitis B Virus Capsids Have Diverse Structural Responses to Small-Molecule Ligands Bound to the Heteroaryldihydropyrimidine Pocket ». *Journal of Virology* 90 (8): 3994-4004. <https://doi.org/10.1128/JVI.03058-15>.

Venkatakrishnan, Balasubramanian, et Adam Zlotnick. 2016. « The Structural Biology of Hepatitis B Virus: Form and Function ». *Annual Review of Virology* 3 (1): 429-51. <https://doi.org/10.1146/annurev-virology-110615-042238>.

WHO. 2020. « Journée mondiale contre l'hépatite : accélérer l'élimination de l'hépatite B chez la mère et chez l'enfant ». 2020. <https://www.who.int/fr/news/item/27-07-2020->

world-hepatitis-day-fast-tracking-the-elimination-of-hepatitis-b-among-mothers-and-children.

Zlotnick, Adam, Balasubramanian Venkatakrishnan, Zhenning Tan, Eric Lewellyn, William Turner, et Samson Francis. 2015. « Core Protein: A Pleiotropic Keystone in the HBV Lifecycle ». *Antiviral Research* 121 (septembre): 82-93. <https://doi.org/10.1016/j.antiviral.2015.06.020>.

# Chapitre 2 : Le virus de l'Hépatite B et les modulateurs d'assemblage

## 1. Généralités sur les virus

### 1.1 Description

Les virus sont des complexes supramoléculaires contenant un matériel génétique. Ils agissent comme des parasites intracellulaires obligatoires, nécessitant des organismes vivants pour se répliquer et synthétiser les protéines virales.

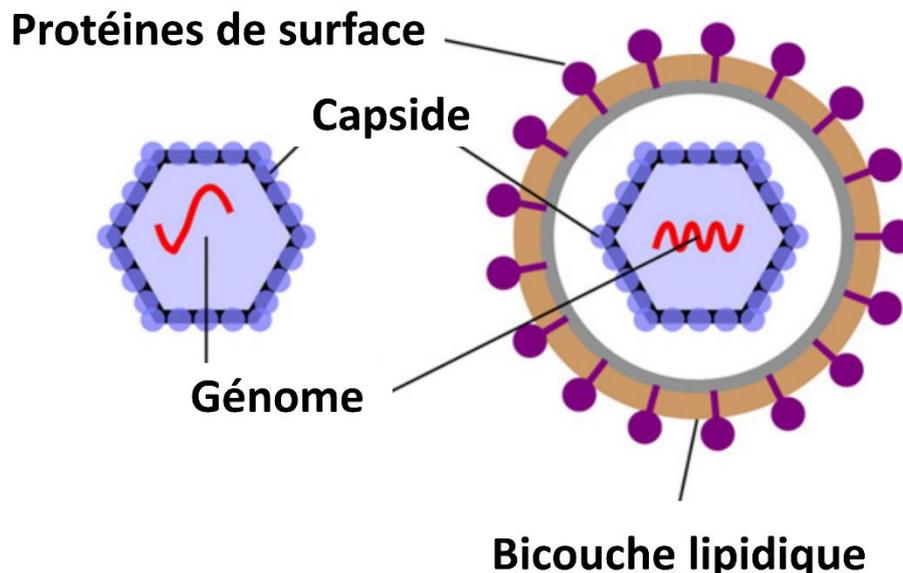


Figure 2.1 : Schéma simplifié d'un virus nu (à gauche sur l'image) et d'un virus enveloppé (à droite sur l'image). Adaptée de (Reddy et Sansom 2016).

Ils sont composés de plusieurs éléments (Figure 2.1) :

- Un matériel génétique sous la forme d'ADN (acide désoxyribonucléique) ou d'ARN (acide ribonucléique), pouvant être double ou simple brin. Ce matériel forme le génome viral.
- Une capsid constituée de protéines, servant à protéger le matériel génétique et les autres constituants du virus de l'environnement extérieur.
- Eventuellement, une enveloppe composée de lipides hérités des membranes des cellules hôtes. Cette enveloppe est associée aux protéines virales

permettant au virus de se lier spécifiquement aux cellules hôtes, facilitant ainsi leur entrée dans les cellules et leur infection.

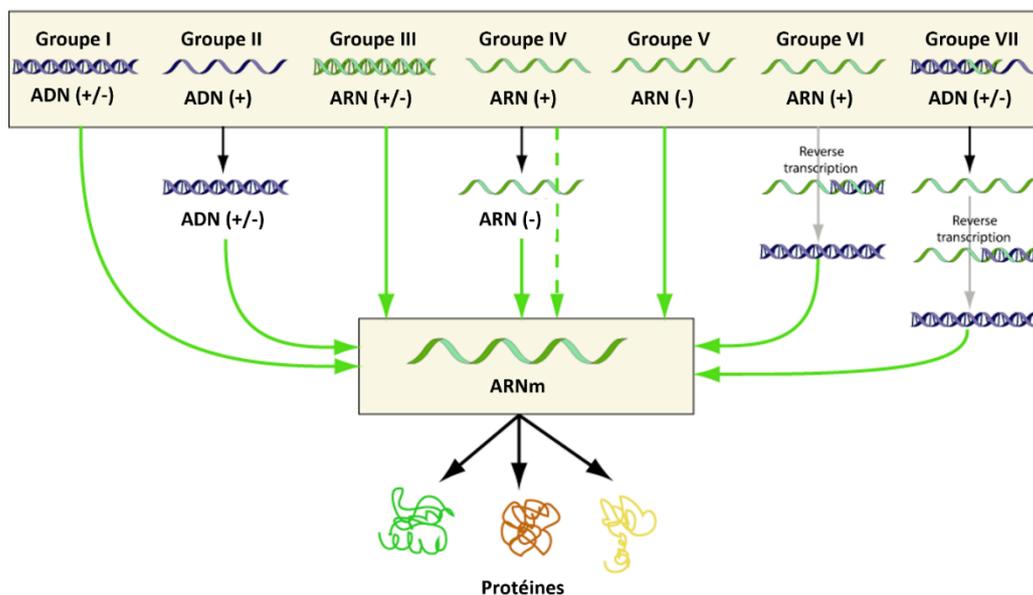
Il existe deux principales catégories de virus en fonction de leur composition :

- Les virus nus ou non-enveloppés sont composés uniquement des deux premiers éléments, c'est-à-dire du matériel génétique et de la capside. Ils ne possèdent pas d'enveloppe. Le virus de la poliomyélite est un virus nu.
- Les virus dits enveloppés, comme le virus de l'hépatite B (VHB) sont composés des trois éléments : le matériel génétique, la capside et l'enveloppe.

## 1.2 Classification

Les virus ne peuvent pas être classés parmi les archées, les bactéries et les eucaryotes. Ainsi, deux classifications ont donc été proposées.

La première classification, a été établie par le Comité International de Taxonomie des Virus (ICTV). Elle classe les virus selon différents niveaux tels que la classe, l'ordre, la famille, le genre, et l'espèce. En 2023, cette classification répertoriait 40 classes, 72 ordres, 264 familles et 11 273 espèces.



**Figure 2.2 : Classification de Baltimore.** Adaptée du site ViralZone ([Baltimore classification ~ ViralZone \(expasy.org\)](https://www.expasy.org/viralzone)).

La seconde méthode de classification est la classification de Baltimore (Figure 2.2), proposée en 1971 (Baltimore 1971). Elle repose sur deux critères : la nature de l'acide

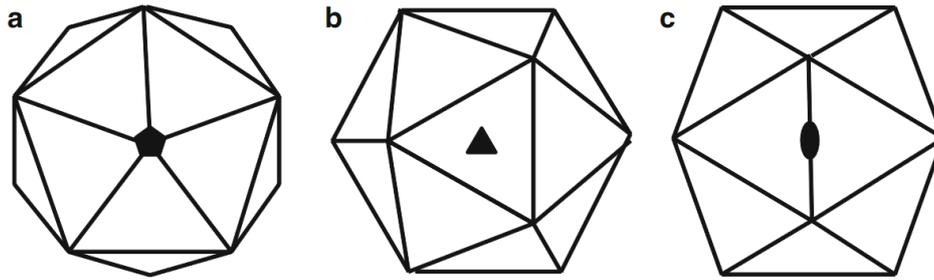
nucléique du génome viral et le mécanisme de synthèse des ARNs messagers (ARNm) permettant la fabrication des protéines virales. Cette classification comporte sept groupes :

- Les virus à ADN double brin forment le groupe 1. Chez ces virus, l'ARNm est directement transcrit à partir de l'ADN double brin et de l'ARN polymérase de la cellule hôte.
- Le groupe 2 contient les virus à ADN simple brin de même polarité que l'ARNm. Cet ADN est converti en ADN double brin permettant ensuite la synthèse de l'ARNm.
- Le groupe 3 comprend les virus à ARN double brin. Le virus contient une ARN polymérase ARN-dépendante qui permet la transcription de l'ARN double brin en ARNm.
- Le groupe 4 réunit les virus à ARN simple brin de polarité positive. Ici, l'ARN peut servir directement d'ARN messager.
- Le groupe 5 englobe les virus à ARN simple brin de polarité négative. L'ARNm est directement transcrit à partir de ce brin. Le virus contient une ARN polymérase ARN-dépendante qui synthétise un brin d'ARN positif.
- Le groupe 6 contient les rétrovirus à ARN simple brin. L'ARN est converti en ADN simple puis double brin avant d'être utilisé pour la synthèse de l'ARNm.

Enfin, les pararétrovirus à ADN double brin forment le groupe 7. Dans ces virus, l'ADN est directement transcrit en ARNm. Le processus de réplication virale implique un intermédiaire ARN, l'ARNpg (l'ARN pré-génomique) et nécessite donc la transcription inverse de l'ARN simple brin en ADN double brin.

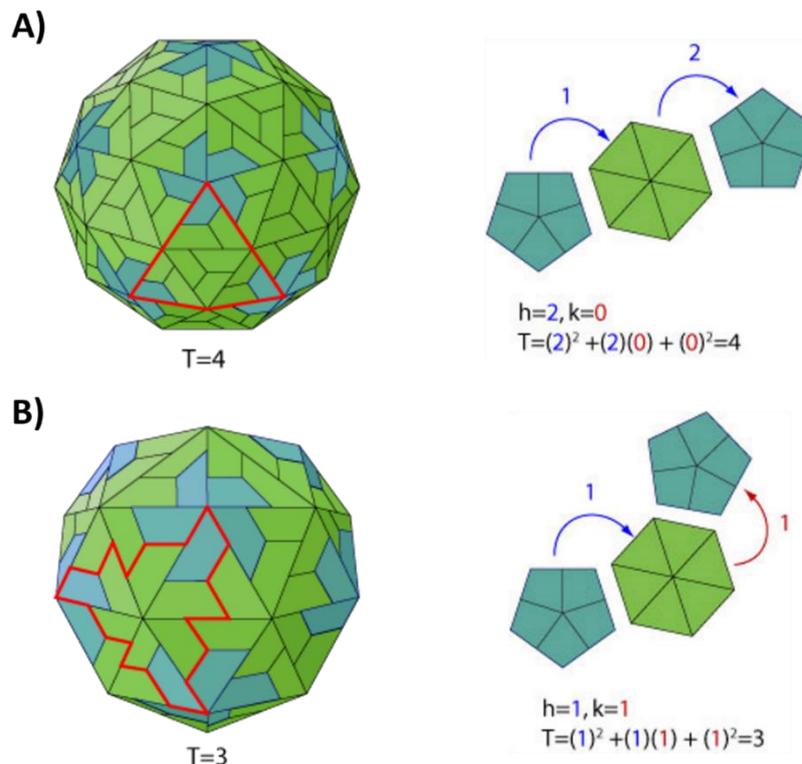
### **1.3 Capsides virales et nombre de triangulation**

Les capsides sont des objets protéiques protégeant le génome viral. Il existe deux grands types de capsides : d'une part, les capsides allongées, caractérisées par une géométrie hélicoïdale, où les protéines et le génome s'agencent sous la forme d'une hélice et d'autre part, les capsides icosaédriques qui représentent environ la moitié des virus sur la Terre (Bruinsma et Klug 2015).



**Figure 2.3 : Description des axes de symétrie dans une capsidie icosaédrique.** L'axe 5 est représenté en (a), l'axe 3 est représenté en (b) et l'axe 2 est représenté en (c). Adaptée de (Prasad et Schmid 2011).

Les capsides icosaédriques ont une géométrie fixe (Caspar et Klug 1962; Perlmutter et Hagan 2015). Elles se composent de 60 sous-unités identiques et s'organisent en 20 faces, 30 arêtes et 12 sommets. Elles possèdent 6 axes d'ordre 5 au niveau des sommets, 10 axes d'ordre 3 au centre de chaque face et 15 axes d'ordre 2 au niveau des arêtes (Figure 2.3), (Prasad et Schmid 2011).



**Figure 2.4 : Exemple de calcul du nombre de triangulation T et des paramètres h et k pour une capsidie de symétrie icosaédrique T=4 (A) et de symétrie icosaédrique T=3 (B).** Adaptée du site ViralZone ([T=4 icosahedral capsid protein ~ ViralZone \(expasy.org\)](https://viralzone.expasy.org) et [T=3 icosahedral capsid protein ~ ViralZone \(expasy.org\)](https://viralzone.expasy.org) ).

Cependant, il est important de noter que les capsides des virus icosaédriques ne sont pas limitées à 60 copies d'une même protéine. Par exemple, la capside du CCMV (Cowpea Chlorotic Mottle Virus) est composée de 180 copies de la même protéine et la capside du VHB est composée de 240 copies de la même protéine. Cette variation dans le nombre de copies de la protéine de capsides a conduit Caspar et Klug à proposer la théorie de la quasi-équivalence (Caspar et Klug 1962).

Selon cette théorie, les sous-unités protéiques interagissent de façon quasi-équivalente afin de s'assembler en capsides icosaédriques ayant plus de 60 sous-unités. Ainsi, une capside icosaédrique comporte  $60T$  sous-unités, composée de 12 pentagones et d'un nombre variable d'hexagones. Le nombre de triangulation  $T$  (Figure 2.4), correspond au nombre de conformations distinctes d'une même sous-unité. La valeur de  $T$  est calculée à l'aide de la formule suivante :  $T = h^2 + hk + k^2$  où  $h$  correspond au nombre de pas nécessaires pour aller du sommet d'un pentamère (assemblage de 5 sous-unités) à un autre pentamère dans un sens donné, en passant par un hexamère, et  $k$  est l'équivalent de  $h$  mais dans un autre sens (Hulo et al. 2011). Chaque pas correspond au chemin entre le sommet d'un pentamère et celui d'un hexamère (assemblage de 6 sous-unités).

Ainsi, le nombre d'hexamères entre deux pentamères vaut  $h + k - 1$ , et le nombre total d'hexamères dans la capside est égal à  $10(T-1)$ . Pour des valeurs de  $T$  supérieures à 1, les sous-unités de la capside icosaédrique s'organisent donc en agencements de pentamères et d'hexamères ayant des contacts presque identiques.

Un virus icosaédrique de type  $T = 4$  est donc composé de 240 copies d'une même protéine dans 4 conformations légèrement différentes appelées A, formant les pentamères, B, C et D, formant les hexamères.

## **2. Le virus de l'hépatite B**

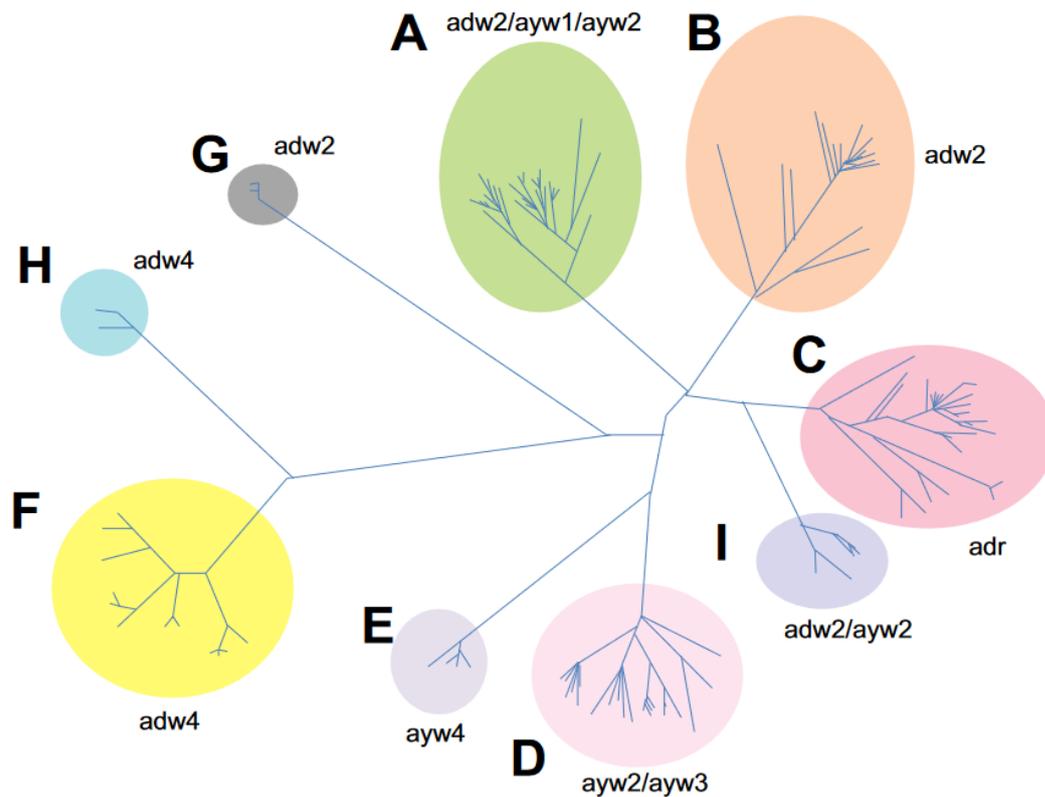
### **2.1 Découverte du virus de l'hépatite B**

L'histoire de l'infection par le VHB chez l'Homme remonte probablement à au moins 4500 ans avant Jésus-Christ (Mühlemann et al. 2018). Cependant, ce n'est qu'en 1885 que la première épidémie d'hépatite B a été observée (Lurman A 1885). En 1965, une équipe dirigée par le professeur Blumberg a découvert un premier antigène chez un patient australien (Blumberg et al. 1967). À l'origine, cet antigène a été associé à la leucémie, mais en 1968, les premiers liens entre l'hépatite B et cet antigène ont été établis par Prince et son équipe (Prince 1968). Deux ans plus tard, en 1970, la morphologie du virus de l'hépatite B a pu être observée par microscopie électronique grâce aux travaux de l'équipe du professeur Dane (Dane, Cameron, et Briggs 1970). C'est d'ailleurs pour cette raison que les particules du VHB sont nommées « particules de Dane ». Enfin, en 1979, la séquence complète du virus de l'hépatite B a été publiée (Galibert et al. 1979) ; s'ensuit l'expérimentation des premiers vaccins l'année suivante (1980).

### **2.2 Génotypes et sérotypes**

L'hépatite B est causée par le virus du même nom. Selon la classification de Baltimore, le VHB appartient au groupe 7, le groupe des virus à ADN double brin, qui se répliquent grâce à un intermédiaire ARN. Ce virus appartient à la famille des Hepadnaviridae qui se divise en plusieurs genres selon la catégorie d'animaux infectés par ce virus. Parmi ces genres, on retrouve les Orthohepadnavirus qui infectent spécifiquement les mammifères, y compris les humains et les chimpanzés. Ce genre est composé de virus enveloppés avec des capsides icosaédriques (Rajoriya et al. 2017).

L'hépatite B peut être classée selon dix génotypes allant de A à J (Figure 2.5), chacun composé de sous-génotypes (Ho, Jeevan-Raj, et Netter 2020). Les différences présentes dans ces séquences nucléotidiques atteignent au moins 8 % entre les génotypes. Il existe 40 sous génotypes, avec des différences de 4 à 8 % dans les séquences nucléotidiques (Tong et Revill 2016; Rajoriya et al. 2017).



**Figure 2.5 : Variabilité des génotypes et sérotypes du virus de l'hépatite B.** Adaptée de (Rajoriya et al. 2017).

Parmi les 10 génotypes, 5 contribuent à environ 97 % des infections chroniques de l'hépatite B. Les génotypes A, B, C, D, et E représentent respectivement 17 %, 14 %, 26 %, 22 % et 18 % de ces infections. La distribution des génotypes du VHB varie en fonction de la population géographique et ethnique. Le génotype A est fréquent en Amérique du Nord, Europe et Afrique. Les génotypes B et C sont fréquents en Asie de l'Est. Le génotype D quant à lui, prédomine en Europe du Sud, au Moyen-Orient et en Inde. Enfin le génotype E est principalement présent en Afrique subsaharienne (Locarnini et al. 2013; Tong et Revill 2016; Rajoriya et al. 2017). Ces variations influencent non seulement la gravité des maladies et leurs symptômes, mais également l'efficacité des traitements.

En outre, le VHB peut être classé en 4 principaux groupes de sérotypes (Figure 2.5) : *adw*, *adr*, *ayw* et *ayr*. Ces sérotypes permettent de distinguer les variants de la protéine S (cette protéine sera décrite dans la partie 4 de ce chapitre) tel que précisé ci-après :

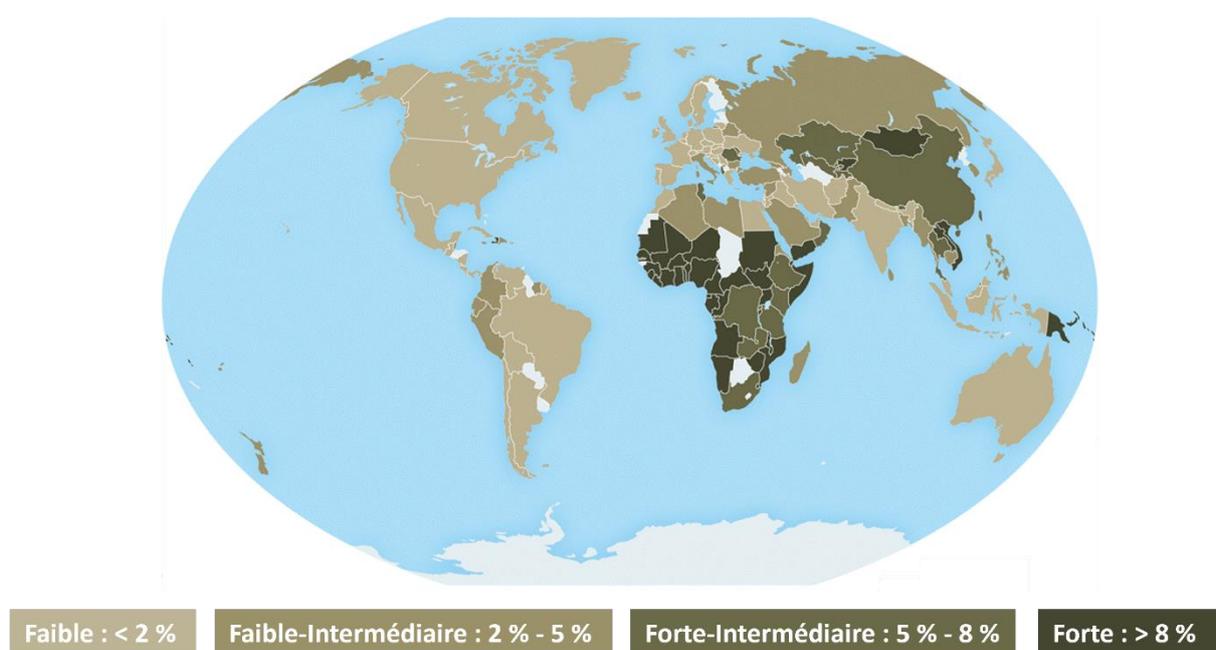
- Le sous-type *d* correspond au variant K122,
- Le sous-type *y* correspond au variant R122,

- Le sous-type *w* correspond au variant K160 et,
- Le sous-type *r* correspond au variant R160.

Bien qu'il existe une corrélation entre les sérotypes et les génotypes du VHB, les variations de sérotypes sont faibles et ne reflètent pas la distribution du VHB dans la classification géographique et épidémiologique, ainsi que dans les génotypes. Par conséquent, les génotypes plutôt que les sérotypes sont de plus en plus utilisés dans les études épidémiologiques du VHB (Locarnini et Bowden 2012; Rajoriya et al. 2017).

## 2.3 Prévalence

L'hépatite B est une maladie infectieuse largement répandue à travers le monde. Selon les données de l'OMS, on estime qu'en 2019, près de 296 millions de personnes sont touchées par l'hépatite B (WHO 2020). Parmi elles, seulement 10 % avaient été diagnostiquées et 22 % des malades diagnostiqués sont sous traitements.



**Figure 2.6 : Prévalence du virus de l'hépatite B.** Adaptée du site du centre pour le contrôle et la prévention des maladies ([Global Viral Hepatitis: Millions of People are Affected | CDC](#)).

La séroprévalence évalue le pourcentage de personnes dans une population donnée, ayant été exposées à un microorganisme et qui développent des anticorps spécifiques à des niveaux significatifs. La séroprévalence du VHB, évaluée à partir de la présence de l'Antigène HBs, est en moyenne de 3,6 % (Chang et Nguyen 2017). Selon ce critère, l'endémicité de ce virus est classée comme suit :

- Faible si elle est inférieure à 2 %,
- Faible-intermédiaire entre 2 et 5 %,
- Forte-intermédiaire entre 5 et 8 % et,
- Elevée au-delà de 8 %.

La distribution de l'hépatite B peut être divisée en trois zones (Figure 2.6). Dans les régions d'Europe de l'Ouest et du Nord, en Australie et en Amérique du Nord, la prévalence est faible. En Europe de l'Est, Russie, au Proche-Orient, en Amérique du Sud, ainsi que dans les pays méditerranéens, la prévalence se situe entre 2 et 8 %. Enfin, en Asie du Sud-Est, en Chine et en Afrique subsaharienne, la prévalence dépasse les 8 % (Schweitzer et al. 2015). Dans certains pays d'Afrique, celle-ci peut même atteindre 20 % de la population. Près de la moitié de la population humaine vit dans des zones à forte endémie d'hépatite B et 1/3 de la population mondiale a été ou est en contact avec le virus de l'hépatite B. On estime que pas moins de 2 milliards de personnes ont été infectées par l'hépatite B à un moment donné.

Chaque année, ce virus est responsable du décès d'environ 1 million de personnes, principalement en Afrique (Locarnini et al. 2015). Cela fait de l'infection par l'hépatite B, l'une des principales causes de décès chez l'être humain.

## 2.4 Transmission et prévention

Le VHB peut être transmis de différentes manières, dont la transmission périnatale et la transmission dite horizontale. La transmission périnatale de la mère à l'enfant est fréquente dans les zones de forte endémie et se produit pendant l'accouchement ou durant les premiers mois après la naissance (Dény et Zoulim 2010). Dans ce cas de figure, les risques de transmission sont plus élevés lors d'une infection aiguë par le VHB au cours du 3<sup>e</sup> trimestre de grossesse. A noter qu'à ce jour, aucune transmission placentaire n'a été démontrée. Concernant la transmission horizontale, elle se produit principalement par exposition à du sang contaminé ou par des rapports sexuels non protégés impliquant l'échange de fluides biologiques contaminés.

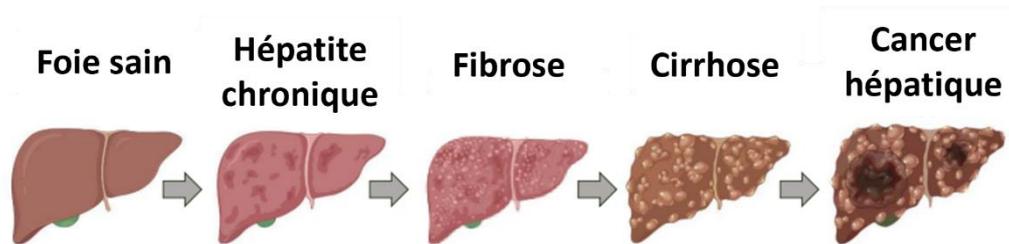
Le virus peut rester infectieux à l'extérieur du corps humain jusqu'à 7 jours, ce qui peut entraîner une infection chez une personne non-vaccinée, entrant en contact avec des surfaces contaminées (Bond et al. 1981).

A ce jour, la vaccination demeure le moyen le plus efficace pour prévenir l'hépatite B. D'après les recommandations de l'OMS, une série de 3 à 4 doses de vaccin permet de couvrir 94 % des nouveau-nés avec 95 % d'efficacité. En 2005, un modèle mathématique a montré que la vaccination avait permis de sauver environ 1 million de vies en 2000 et d'empêcher l'infection d'environ 65 millions de personnes (Goldstein et al. 2005).

## 2.5 Symptômes et traitements

Il existe 2 formes d'hépatite B : l'hépatite aiguë et l'hépatite chronique.

L'hépatite B aiguë est asymptomatique et guérie spontanément dans 90 à 95 % des cas en 6 mois (Hollinger et Lau 2006). Néanmoins, elle peut provoquer de la fatigue, de la nausée, des douleurs abdominales, mais également des jaunisses chez certains patients. De plus, chez 5 à 10 % des malades, cette forme peut évoluer vers une infection chronique. Cette condition est plus souvent observée chez des enfants infectés très jeunes, avant l'âge de 5 ans ou chez des patients immunodéprimés.



**Figure 2.7 : Evolution des symptômes causés par le virus de l'hépatite B, vers le carcinome hépatocellulaire.** Adaptée de (Crouchet, Schuster, et Baumert 2021).

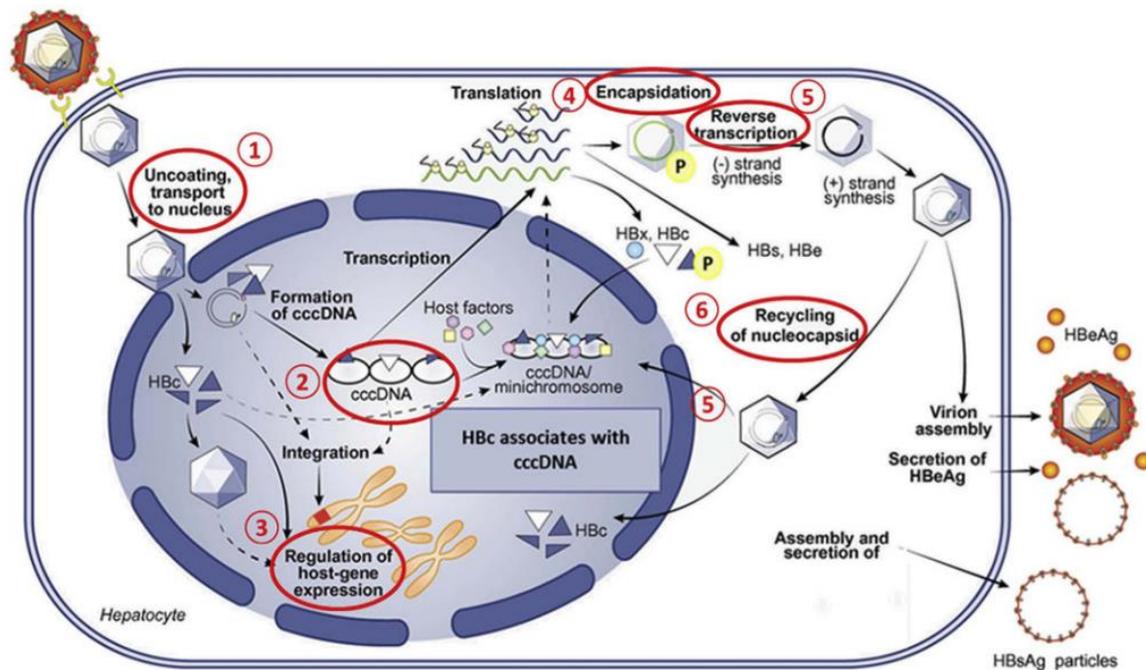
Les hépatites chroniques peuvent être plus ou moins actives. Cette forme d'hépatite se caractérise par la persistance de l'antigène de surface HBs 6 mois après l'infection (Lavanchy 2004). Dans 70 à 90 % des cas, les patients atteints d'hépatite B chronique sont asymptomatiques. Cependant, dans certains cas, l'infection par l'hépatite B peut évoluer vers une forme active, entraînant des complications allant de la fibrose à la cirrhose (C. Wang et al. 2017). La cirrhose est le principal déclencheur de complications hépatiques liées au VHB. Elle peut conduire à l'insuffisance hépatocellulaire et peut favoriser l'apparition d'un carcinome hépatocellulaire (HCC), potentiellement mortel (Figure 2.7).

Le diagnostic de l'hépatite B repose sur des tests moléculaires et sérologiques, notamment des analyses immunologiques permettant de détecter divers marqueurs tels que les antigènes viraux HBs, HBe et HBc ainsi que leurs anticorps respectifs. La période d'incubation de ce virus varie de 30 à 180 jours, ainsi ce virus est détectable sur une durée allant de 30 à 60 jours après l'infection (WHO 2020).

Actuellement, il n'y a pas de traitement pour guérir d'une infection causée par le virus de l'hépatite B. Les médicaments utilisés sont les analogues de nucléosides et les interférons alpha pégylés (PEG-IFN). Les analogues de nucléosides inhibent la polymérase du virus et donc la réplication virale. Ils réduisent la charge virale, retardent la progression des cirrhoses, et réduisent ainsi les risques de cancer hépatique. Toutefois, ils n'éliminent pas l'ADNccc (ADN circulaire covalamment clos) viral, responsable de la persistance de l'infection. Enfin, la nécessité d'un traitement à long terme, conduit à l'apparition de résistances. Les PEG-IFN qui activent le système immunitaire inné et affectent la réplication virale du VHB sont bien tolérés par une petite population de malades et peuvent avoir des effets secondaires graves (Mak et al. 2017).

### 3. Cycle viral du virus de l'hépatite B

Malgré la complexité du cycle viral de l'hépatite B (Figure 2.8), une schématisation en plusieurs étapes peut être établie. Cette schématisation, montre l'importance de la protéine Core au sein du cycle viral.



**Figure 2.8 : Cycle viral du virus de l'hépatite B.** 1) Core régule le transport et la libération nucléaire du génome viral. 2) Core s'associe à l'ADNccc. 3) Core module l'expression des gènes de l'hôte. 4) Core est nécessaire à l'encapsidation de l'ARN polymérase. 5) Core est nécessaire à la transcription inverse, et 6) la persistance de l'ADNccc implique le recyclage des nucléocapsides dans le noyau. Adaptée de (Diab et al. 2018).

#### 1. Entrée dans la cellule :

Tout d'abord, la grande protéine de surface du VHB, appelée protéine L, se lie aux récepteurs NTCP (sodium taurocholate cotransporting polypeptide) présents sur les hépatocytes (Urban et al. 2014). Cette interaction va permettre l'entrée des particules virales dans le cytoplasme des hépatocytes. Le mécanisme précis d'entrée, n'est pas complètement clair, mais il pourrait impliquer la fusion de l'enveloppe virale avec la membrane cellulaire ou l'endocytose de la particule virale (Huang et al. 2012; Macovei et al. 2010).

## **2. Transport de la nucléocapside vers le noyau :**

Lors de la pénétration dans le cytoplasme, la nucléocapside qui renferme l'ADN viral, est libérée et transportée vers le noyau grâce aux microtubules (Rabe, Glebe, et Kann 2006).

## **3. Importation dans le noyau :**

La nucléocapside contenant le génome viral interagit avec les importines via les séquences de localisation nucléaire (NLS) et est transportée à travers le pore nucléaire où elle sera désassemblée (Schmitz et al. 2010; Gallucci et Kann 2017). Cette étape permet la libération du génome viral et de la protéine de capsid dans le noyau (Rabe et al. 2009).

## **4. Formation du mini-chromosome et transcription :**

Après l'entrée du génome dans le noyau, l'ADNrc (ADN relâché, circulaire) de la nucléocapside est converti en un ADN circulaire covalentement clos appelé ADNccc (Michael Nassal 2008). Cet ADN est surenroulé et associé à des histones ainsi qu'à certaines protéines du VHB dont la protéine de capsid pour former le minichromosome. Le minichromosome est transcrit par l'ARN polymérase II cellulaire, ADN-dépendante (Michael Nassal 2008). Cela conduit à la synthèse des ARNm viraux et de l'ARN pré-génomique (ARNpg).

## **5. Traduction :**

Ces ARNs sont ensuite traduits pour produire les différentes protéines du virus. Les ARNm viraux codent pour les protéines S, M et L, la protéine E et la protéine X. L'ARN pré-génomique joue un double rôle en tant qu'ARN messager pour la synthèse de la protéine Core (la protéine de la capsid) et de la polymérase virale, ainsi qu'en tant que matrice pour la transcription inverse par la polymérase virale.

## **6. Encapsidation et devenir de la nucléocapside :**

Puis l'ARNpg et la polymérase sont encapsidés (Michael Nassal 2015), et l'ARNpg est converti en ADNrc (Schreiner et Nassal 2017). Pour finir, cette nucléocapside mature soit acquiert une enveloppe et est sécrétée, soit est recyclée vers le noyau pour maintenir le pool d'ADNccc.

La formation de capsides vides est également observée dans le cycle viral du VHB (Ning et al. 2011)

#### 4. Les différentes protéines du virus de l'hépatite B

Le génome du VHB est un ADN relâché circulaire, partiellement double brin de 3,2 kb (Figure 2.9).

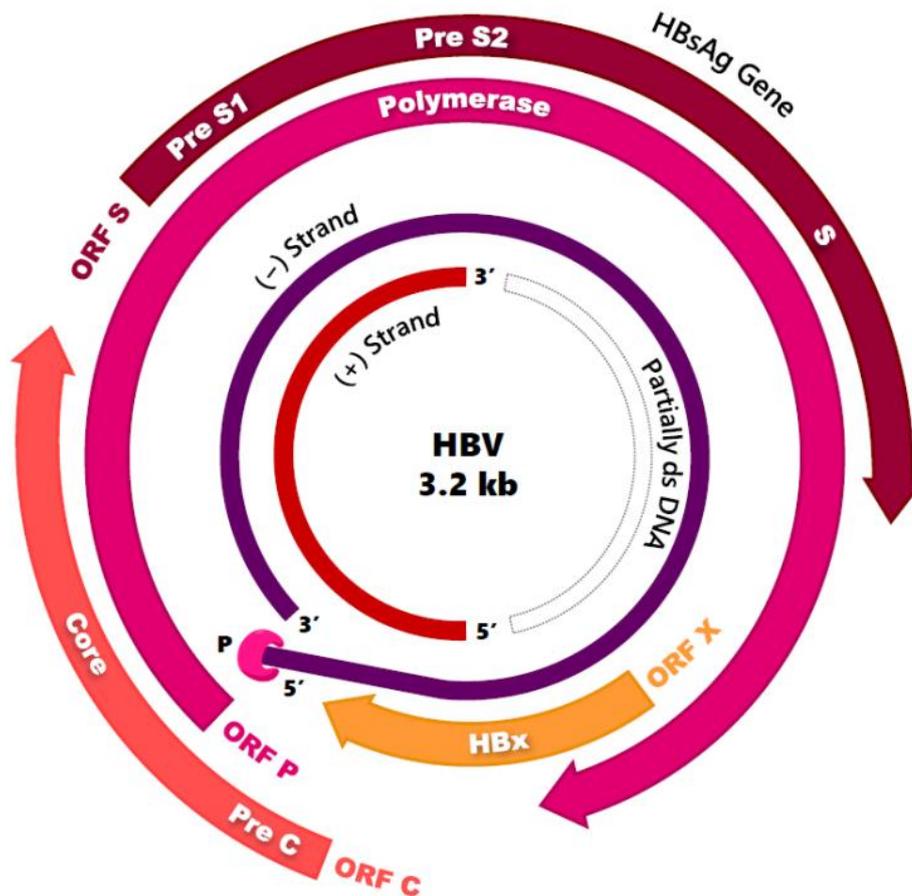


Figure 2.9 : Génome du virus de l'hépatite B. Adaptée de (Prifti et al. 2021).

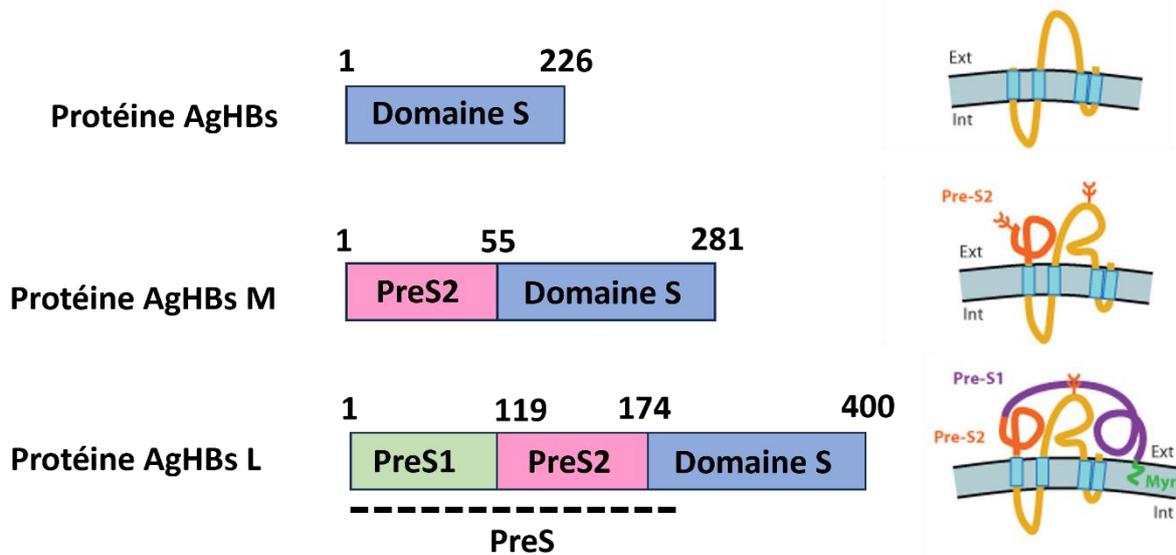
Après réparation dans le noyau en ADNccc, le brin positif contient 4 phases ouvertes de lecture. Ces 4 phases ouvertes de lecture, codent pour 7 protéines virales différentes : la polymérase, la protéine Core (AgHBc), la protéine E (AgHBe), la protéine X (HBx), et les protéines d'enveloppe, L M et S (AgHBs L, AgHBs M, AgHBs).

##### La protéine Core :

En raison de ses fonctions, la protéine Core est l'une des protéines les plus importantes du VHB. Elle sera présentée dans la partie suivante.

## Les protéines d'enveloppe :

Le virus de l'hépatite B possède trois protéines d'enveloppes différentes : AgHBs, AgHBs M et AgHBs L. Ces protéines sont retrouvées à la surface des virions infectieux et peuvent également former spontanément des structures sphériques ou filamenteuses vides appelés sous-particules virales (SVPs) (Heermann et al. 1984). Les SVPs sont non infectieuses, mais elles pourraient favoriser la propagation des virions et ainsi participer au maintien de l'infection.



**Figure 2.10 : Topologie des protéines d'enveloppe.** Les schémas de gauche représentent les différents domaines des protéines AgHBs, AgHBs M et AgHBs L. Les schémas de droite représentent les positions des différents domaines de ces protéines par rapport à la membrane. Chez la protéine AgHBs M, les domaines PréS2 et S peuvent être glycosylés. Chez la protéine AgHBs L, le domaine PréS peut traverser la membrane (non représenté ici), le domaine S est glycosylé et le domaine PréS1 subit une myristoylation. Les protéines orientées vers l'intérieur sont impliquées dans la morphogenèse du virus, et les protéines orientées vers l'extérieur sont impliquées dans la reconnaissance des récepteurs. Adaptée du site ViralZone ([Surface proteins ~ ViralZone \(expasy.org\)](http://Surface%20proteins%20~%20ViralZone%20(expasy.org))).

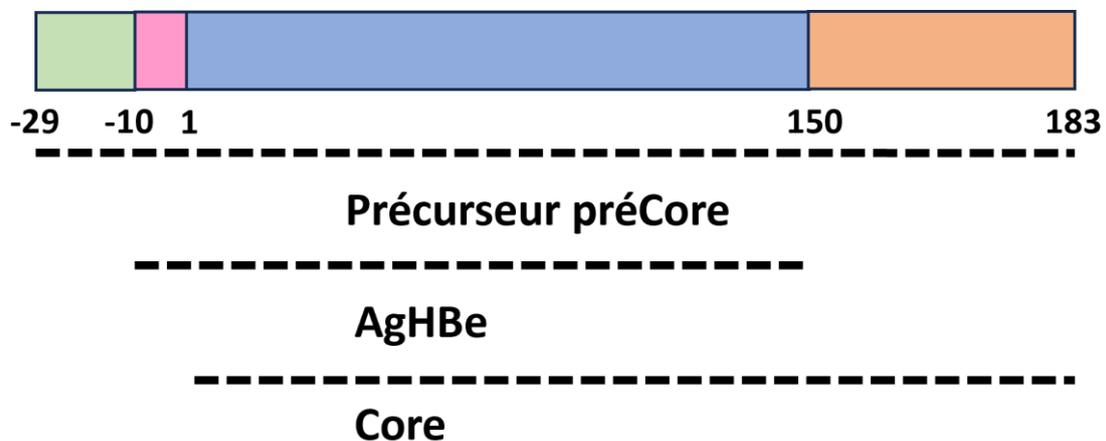
Ces protéines ont une région C-terminale commune appelée la région S qui est en fait la protéine S. La protéine AgHBs (Figure 2.10) est composée de 226 acides aminés (AAs) (Eble, Lingappa, et Ganem 1986). Elle représente environ 2/3 de la composition totale de l'enveloppe des SVPs. De plus, AgHBs est un marqueur de l'infection virale (Brunetto et al. 2010) et est à la base des vaccins anti-VHB recombinants actuels.

La protéine AgHBs M (Figure 2.10) contient le domaine preS2 dans la région N-terminale, en plus de la région S. Le domaine preS2 est composé de 55 AAs (Ganem et Varmus 1987). Le rôle de cette protéine reste à clarifier étant donné qu'elle n'est pas nécessaire pour l'infection ni pour la morphogénèse virale (Bruss et Ganem 1991).

Enfin, la protéine AgHBs L (Figure 2.10) contient le domaine S et le domaine preS. Ce dernier comprend les domaines preS2 et preS1. Le domaine preS1 varie en longueur de 108 à 119 AAs selon les géotypes. La protéine AgHBs L est essentielle pour la formation des sous-particules virales filamenteuses, pour l'acquisition de l'enveloppe des capsides du VHB et pour la liaison avec le récepteur hépatocytaire NTCP lors de l'entrée du virus (via le domaine preS1) (Glebe et al. 2005; Le Duff, Blanchet, et Sureau 2009; Ni et al. 2014; Yan et al. 2012; Wong et al. 2017).

### La protéine E :

AgHBe est une protéine de 17 kDa produite par traduction de l'ARN préCore de 3,5 kb. Elle est codée par le même cadre de lecture que la protéine Core et possède 90 % d'identité de séquence en acides aminés avec elle mais a des rôles complètement différents.



**Figure 2.11 : Structure de la protéine AgHBe.** La protéine AgHBe est obtenue par maturation du précurseur de la préCore. Adaptée de (Duriez et al. 2014).

C'est une protéine hydrophobe, contenant par rapport à Core 10 résidus supplémentaires en N-terminus, précédés d'un peptide signal qui lui permet de passer dans les voies de sécrétion. Après maturation dans ces voies, la forme sécrétée AgHBe contient les résidus -10 à ~150 dans la numérotation de Core (Figure 2.11). AgHBe n'est pas impliquée dans la réplication virale, mais participe à la tolérance de

l'hôte envers le virus en affectant le système immunitaire (Jegaskanda et al. 2014; Lang et al. 2011). De plus, la sécrétion d'AgHBe étant corrélée avec de hauts niveaux d'expression de la réplication des virions du VHB (J. Chen et al. 2017), cet antigène est également un biomarqueur pour suivre la réplication virale du VHB.

### La polymérase :

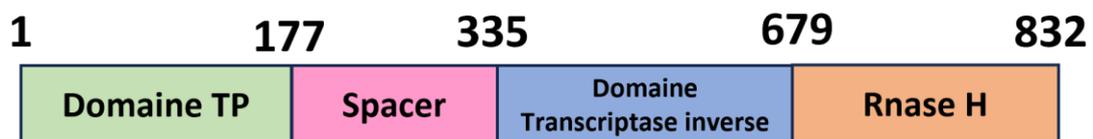


Figure 2.12 : Structure de la polymérase du virus de l'hépatite B. Adaptée de (Clark et Hu 2015).

La polymérase (Figure 2.12) est l'une des protéines les plus importantes du VHB. Elle est la plus grande protéine du VHB, a une masse moléculaire de 90 kDa, et est composée de 832 acides aminés. Cette protéine est divisée en 4 domaines : le domaine TP (Terminal Protein), le spacer, le domaine de la transcriptase inverse et le domaine ribonucléase H (RNase H). Le domaine TP situé à l'extrémité N-terminale permet la liaison à l'ARNpg et sert d'amorce pour initier la synthèse du brin négatif d'ADN (Bartenschlager et Schaller 1988; Jones et Hu 2013). Le spacer, dont la fonction n'est pas clairement définie, est un domaine flexible. Il permettrait la connexion entre les différents domaines et leur bon fonctionnement. Du côté C-terminal, deux domaines sont présents : Le domaine de transcriptase inverse, qui permet la synthèse de l'ADN (Radziwill, Tucker, et Schaller 1990), et le domaine RNase H, responsable de la digestion de l'ARNpg durant la transcription inverse (Wei et Peterson 1996; Ko et al. 2014).

La polymérase est la cible des antiviraux à action directe actuels, en particulier des analogues de nucléos(t)ides (Zoulim et Locarnini 2013).

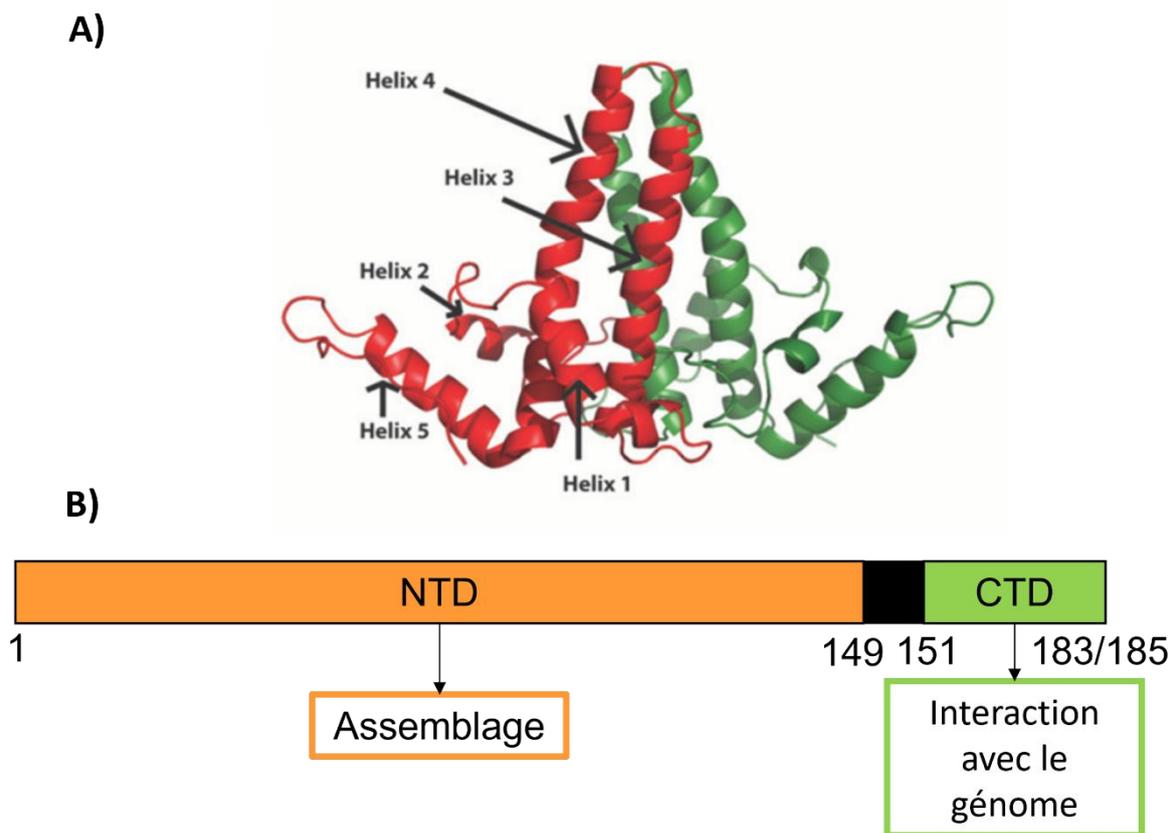
### La protéine X :

La protéine X est une protéine de 154 acides aminés, non structurée. Bien que HBx ne soit pas impliquée dans la structure des virions du VHB, elle joue un rôle dans la réplication virale via le contrôle de la transcription de l'ADNccc avec lequel elle interagit (Lucifora et al. 2011). De plus, elle jouerait un rôle important dans la régulation indirecte de nombreux gènes hôtes. Par ailleurs, elle a un potentiel oncogénique et jouerait également un rôle dans le développement des HCC liés au VHB.

## 5. La protéine Core

### 5.1 Structure de la protéine entière

La protéine Core est une protéine dimérique (Figure 2.13.A) (Zlotnick et al. 2015) de 183 (ou 185 ou 195, selon les souches) acides aminés (Locarnini et al. 2013) composée de deux domaines. Les 149 résidus N-terminaux (Cp149) forment un domaine structuré appelé NTD (N Terminal Domain) et composé de 5 hélices alpha (Birnbaum et Nassal 1990; Wynne, Crowther, et Leslie 1999). Le NTD (Figure 2.13.B) est responsable de la dimérisation et de l'autoassemblage *in vivo*, mais également *in vitro* des capsides du VHB.



**Figure 2.13 : Dimère de la protéine Core.** A) Structure cristallographique du dimère de Core dans le contexte de la capside, adaptée de (Venkatakrishnan et Zlotnick 2016). B) Domaines de la protéine Core.

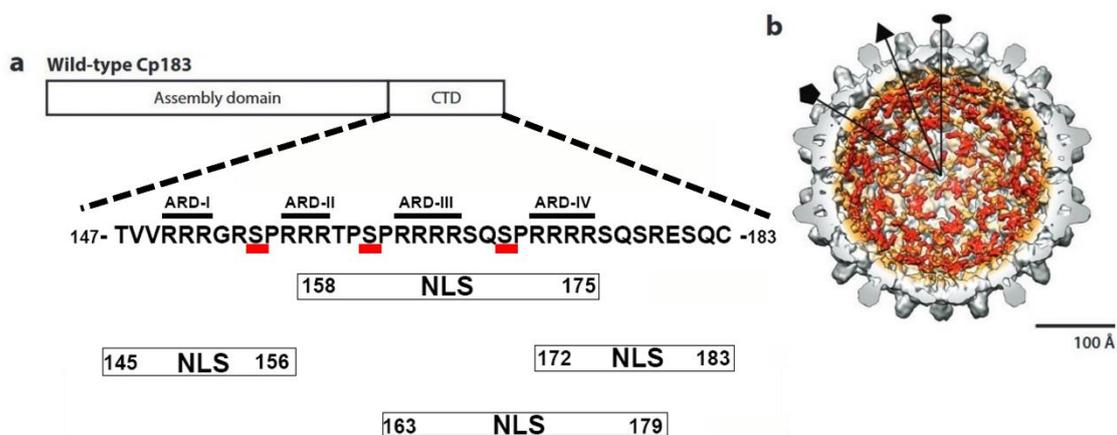
Les acides aminés restants, forment un domaine intrinsèquement désordonné, riche en arginines et sérines (Figure 2.13.B) appelé CTD (C terminal Domain) (J. C.-Y. Wang, Dhason, et Zlotnick 2012).

## 5.2 Fonctions de la protéine Core

La protéine Core est le constituant de la capsid du VHB. La protéine Core va aussi avoir plusieurs autres rôles, probablement sous sa forme dimérique. Dans les hépatocytes infectés par le VHB, la protéine Core s'accumule dans le noyau, elle aurait donc des fonctions nucléaires lors de l'infection par le VHB. En effet, la Core interagit avec l'ADNccc lors de la formation du minichromosome viral (Bock et al. 2001). Sa présence a également été corrélée à un état de transcription active de l'ADNccc, puisqu'elle pourrait l'activer en interagissant avec cet ADN au niveau des îlots CpG (Y.-H. Guo et al. 2011). Core aurait ainsi un rôle important dans la régulation de la transcription virale. De plus, elle serait également impliquée dans la régulation de nombreux gènes cellulaires en se liant à leurs promoteurs et inhibant ainsi leur expression (Y. Guo et al. 2012). Enfin, l'interaction de l'ADNccc avec certaines enzymes telles que la cytidine déaminase, qui peut conduire à sa dégradation est médiée par la protéine Core. Elle participerait donc à la destruction de l'ADNccc (Lucifora et al. 2014).

## 5.3 Domaine C-terminal

Le CTD est un domaine riche en arginines et sérines pouvant être phosphorylées. Ce domaine comprend également les séquences de localisation nucléaire (NLS-nuclear localization signal) et d'export nucléaire (NES-nuclear export signal).



**Figure 2.14 : Domaine C-terminal (CTD) de la protéine Core.** (a) Schéma du domaine d'assemblage et du CTD, y compris la séquence du CTD. S155, S162 et S170 sont considérées comme critiques pour l'encapsidation de l'ARN pré-génomique. (b) Vue en coupe d'une reconstruction par cryomicroscopie

électronique d'une capsid e Cp183 vide. La densité en couleur correspond au CTD d'après l'ajustement d'une structure cristalline d'une capsid e Cp149. Adaptée de (Venkatakris hnan et Zlotnick 2016).

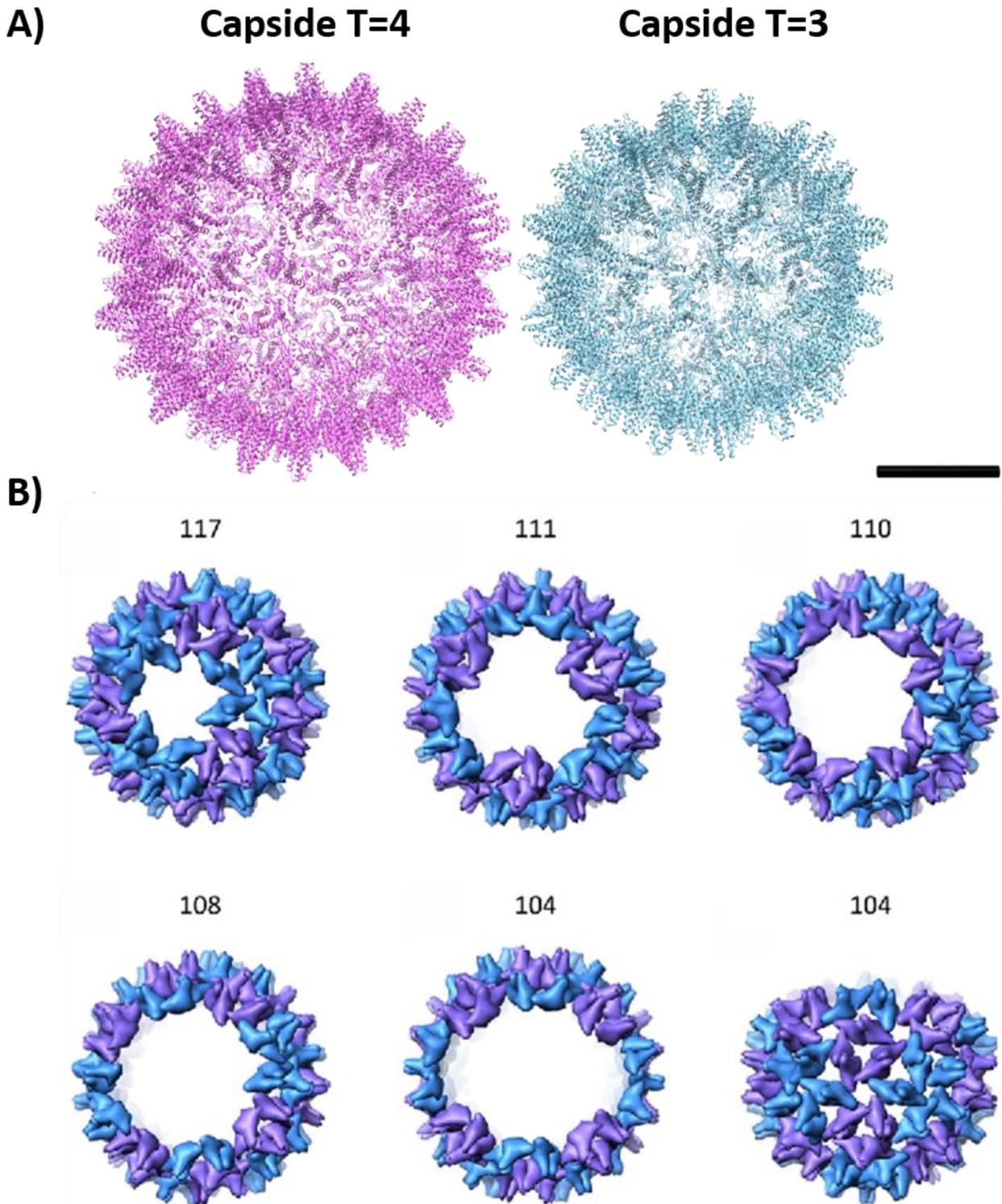
Sous sa forme phosphorylée, le CTD interagit avec l'ARNpg dans le cytoplasme et est nécessaire pour son encapsidation (Gallina et al. 1989; M. Nassal 1992). Des études ont montré que la phosphorylation des sérines 155, 162 et 170 (Figure 2.14) sont importantes pour l'encapsidation de l'ARNpg (Kann et Gerlich 1994). De plus, le CTD a une activité chaperonne permettant le bon repliement de l'ARNpg et sa transcription inverse en ADNrc.

Les déphosphorylations successives du CTD qui accompagnent la synthèse de l'ADNrc à partir de l'ARNpg, entraînent l'exposition de certains CTD et donc des NLS à l'extérieur de la nucléocapsid e mature. Cette exposition des NLS permet l'interaction du CTD avec les importines et conduit à l'acheminement de la nucléocapsid e mature au noyau. Alternativement, la maturation associée à la synthèse d'ADNrc conduit à l'interaction avec l'AgHbS L pour l'acquisition de l'enveloppe et la sécrétion de la cellule (M. Nassal 1992; Michael Nassal 2008; Schmitz et al. 2010; Rabe et al. 2009; C. Chen et al. 2016; Selzer et al. 2015).

Les phosphatases responsables de la déphosphorylation du CTD, ainsi que les kinases permettant sa phosphorylation n'ont pas encore été identifiées (Chabrolles et al. 2018). Cependant, des travaux suggèrent que les protéines kinases A et C, mais également les protéines kinases dépendantes des cyclines 1 et 2 jouent un rôle dans la phosphorylation du CTD.

## 5.4 Autoassemblage de la capsid e

La protéine Core peut s'autoassembler *in vivo* et *in vitro* en capsides icosaédriques de triangulation  $T = 4$  mais également en capsides icosaédriques de triangulation  $T = 3$  (Figure 2.15.A). La majorité des capsides formées (environ 95 %) sont des capsides de type  $T = 4$  (Venkatakris hnan et al. 2016; M. Nassal 1992). Ces capsides sont composées des 120 dimères de la protéine Core, présente sous 4 conformations différentes notées A, B, C et D. Les capsides  $T = 3$  sont composés de 90 dimères de la protéine Core, présente sous 3 conformations notées A, B et C.



**Figure 2.15 : Capside du virus de l'hépatite B.** A) Structure des capsides du VHB de triangulation T=4 (code PDB : 1QGT) et de triangulation T=3 (code PDB : 6BVN). La barre d'échelle est égale à 10 nm. B) Modèles possibles des intermédiaires tardifs observés durant l'assemblage de la capside. Adaptée de (Pierson et al. 2014).

L'assemblage de la Core suit un processus de nucléation-croissance (Perlmutter et Hagan 2015; Endres et Zlotnick 2002). La nucléation consiste en la formation d'un noyau de dimères. Puis, durant l'élongation, les sous-unités dimériques libres sont ajoutées au fur et à mesure jusqu'à la formation de la capsid complète. Quelques espèces intermédiaires au cours de la phase d'élongation peuvent être observées (Oliver et al. 2020). De plus, plusieurs intermédiaires tardifs (Figure 2.15.B) comprenant 104 à 117 dimères de Core ont pu être observés par spectrométrie de masse et par cryomicroscopie électronique (Pierson et al. 2014; Lutomski et al. 2018).

Plusieurs facteurs tels que la force ionique (Ceres et Zlotnick 2002), la stabilité de l'interface entre les sous-unités (Patterson et al. 2020) ou la surface de contact entre les sous-unités (Garmann et al. 2014) influencent l'assemblage de la capsid du VHB. Toutefois, ce processus ne nécessite pas la présence de protéines chaperonnes.

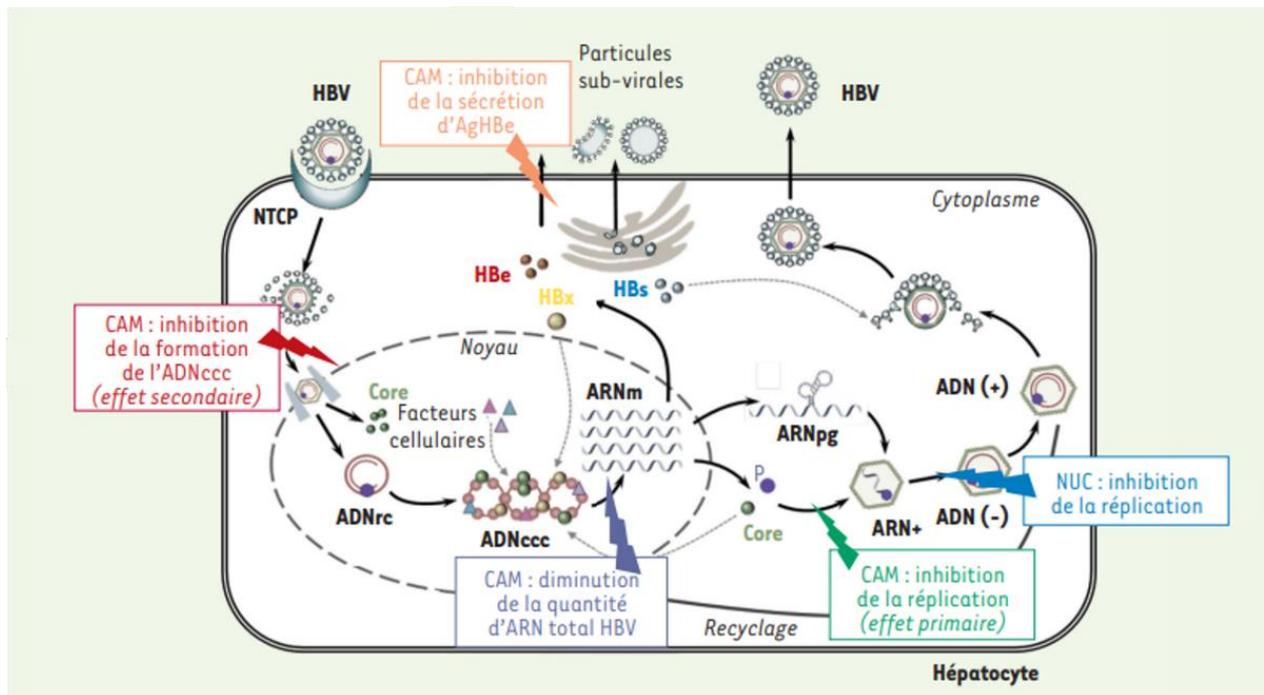
Le TR-SAXS (Time-resolved small-angle X-ray scattering, qui signifie diffusion de rayons X aux petits angles résolue en temps) permet de suivre l'assemblage de capsides virales (Kler et al. 2012; Tresset et al. 2013; Chevreuil et al. 2018; Asor et al. 2020; Law-Hine et al. 2015; 2016; Chevreuil et al. 2020; Burke et al. 2018). Les données obtenues lors de ces différentes études ont ainsi révélé que l'assemblage de la protéine Cp149 (uniquement composée du domaine NTD) dans des conditions douces, à pH = 7,5 et avec une concentration en sel monovalent égale à 150 mM suit bien un processus à deux étapes impliquant uniquement les dimères de Core et la capsid. En présence d'une plus forte concentration de sel, la présence de capsides icosaédriques de triangulation  $T = 3$  et l'accumulation brève d'intermédiaires contenant 7 à 35 sous-unités dimériques peuvent être observés (Asor et al. 2020). Dans une autre étude, lors de l'assemblage de la Cp149 à pH 9, une étape supplémentaire a été observée. Durant cette étape, les capsides malformées se réarrangent pour former des capsides correctes.

## **6. Les modulateurs d'assemblage**

### **6.1 Mécanisme d'action**

Les modulateurs d'assemblage (CAMs-capsids assembly modulators) sont des petites molécules antivirales qui ciblent spécifiquement la protéine Core du VHB (Taverniti et al. 2022). En ciblant une protéine importante dans le cycle viral du VHB, ces molécules

sont considérées comme des candidates prometteuses pour le traitement de l'hépatite B. Plusieurs CAMs sont en cours d'essais cliniques, soit en monothérapie, soit en combinaison avec d'autres antiviraux.

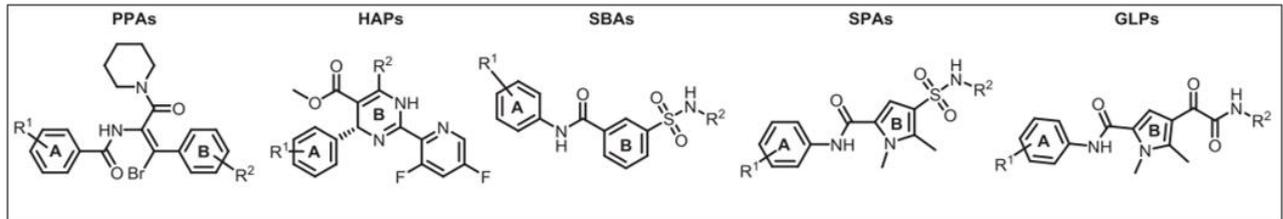


**Figure 2.16 : Effet des modulateurs d'assemblage (CAMs) sur le cycle réplcatif du VHB.** Les CAMs peuvent interférer avec plusieurs étapes du cycle viral. Adaptée de (Chabrolles et al. 2018).

Leur effet primaire consiste à perturber l'assemblage de la capsid virale. Les CAMs accélèrent le processus d'assemblage de la capsid de manière dose-dépendante, en se liant à une poche hydrophobe qui facilite les contacts entre les dimères de la protéine Core dans la capsid finale (Stray et al. 2005; Katen et al. 2010; Klumpp et al. 2015; Berke et al. 2017; Kra et al. 2023). Cette perturbation empêche l'encapsidation de l'ARNpg, et conduit à l'inhibition de la réplication virale, empêchant ainsi la formation de l'ADNrc. Ces molécules peuvent également inhiber la réplication virale des mutants VHB résistants aux analogues de nucléosides (Diab et al. 2018).

En culture cellulaire, les CAMs ont également plusieurs autres effets. Ces molécules inhibent la formation de l'ADNccc, probablement en perturbant le trafic cellulaire des nucléocapsides, contribuant ainsi à lutter contre la persistance du VHB dans les cellules (Berke et al. 2017; Zlotnick et al. 2015; Zoulim 2005). Elles induisent une diminution de la quantité totale d'ARNm viraux et de la sécrétion de la protéine AgHBe. Chez les cellules HepaRG infectées par le VHB et dans les cellules hépatocytes

primaires, un traitement à long terme ou à des concentrations élevées de CAMs entraîne une réduction de la transcription de l'ARN viral, de la sécrétion de AgHBe (en induisant son accumulation dans la cellule), de la sécrétion de AgHBs ainsi que de l'ADNccc total (Lahlali et al. 2018; Berke et al. 2017; F. Guo et al. 2017). Enfin, en interagissant avec la protéine Core, les CAMs pourraient perturber l'expression de gènes cellulaires et également bloquer la voie de recyclage des nucléocapsides vers le noyau (Figure 2.16).



**Figure 2.17 : Structure chimique des modulateurs d'assemblage les plus courants.** Les HAPs appartiennent au groupe des CAM-A et les autres molécules appartiennent au groupe de CAM-E. Adaptée de (Nijampatnam et Liotta 2019).

L'effet primaire des CAMs est observé à très faibles doses, cependant les autres effets nécessitent de fortes doses et ne sont observables qu'en présence de certaines molécules. Les CAMs (Figure 2.17), qui inhibent donc des étapes précoces, mais également des étapes tardives du cycle viral du VHB, peuvent être classées en deux catégories en fonction de leur mode d'action primaire (Zoulim et al. 2022) : le type CAM-A (A pour « aberrant ») et le type CAM-E (E pour « empty », signifiant vide).

## 6.2 Les modulateurs d'assemblage de type A

Les modulateurs de type A induisent la formation de très grosses structures non capsidiques, pléiomorphes et de capsides mal-assemblées qui s'agrègent dans la cellule (Klump et al. 2015; Corcuera et al. 2018; Nijampatnam et Liotta 2019).

Cette catégorie de modulateurs regroupe les hétéroaryldihydropyrimidines (HAP) tels que BAY 41-4109, HAP-1, HAP-12, HAP-18, ainsi qu'un autre groupe de composés chimiques, les ciclopirox (Kang et al. 2019; Mohebbi et al. 2023).

Les HAPs se lient préférentiellement aux sous-unités hexamériques B et C de la protéine Core dans la capsid du VHB (Bourne, Finn, et Zlotnick 2006;

Venkatakrisnan et al. 2016) et peuvent en plus déstabiliser les nucléocapsides déjà assemblées (Schlicksup et al. 2018).

Les ciclopirox sont des composés approuvés par la FDA (Food and drug administration) et initialement utilisés pour le traitement des maladies de la peau dues à certains champignons. Cependant Kang et son équipe (Kang et al. 2019), ont démontré que ces composés peuvent également empêcher la formation de capsides et conduire à la formation de structures plus grandes et ouvertes.

### **6.3 Les modulateurs d'assemblage de type E**

Contrairement aux modulateurs de type A, les modulateurs de type E sont censés accélérer la formation de capsides du VHB morphologiquement intactes. Néanmoins, de récentes études ont montré que certains modulateurs de cette classe conduiraient à la formation d'objets allongés, n'ayant plus la morphologie de capsides (Lecoq et al. 2023; Kra et al. 2023). De plus, ces molécules inhibent la formation de nucléocapsides contenant de l'ARNpg (Campagna et al. 2013; Cole 2016).

Cette classe est composée de plusieurs types de molécules différentes, telles que les sulfamoylbenzamides (SBA) comme JNJ-632, les dérivés phénylpropénamides (PPA), (Berke et al. 2017; Zhou et al. 2017) et les glyoxamide-pyrrolamides (GPA).

Plusieurs études in vitro ont été menées sur les effets des modulateurs et il est important de noter que pour l'instant, aucun composé de CAMs n'a été en mesure de déclencher une diminution de l'HBsAg chez les patients, et l'arrêt du traitement est associé à un rebond viral (Taverniti et al. 2022).

## Bibliographie

Asor, Roi, Christopher John Schlicksup, Zhongchao Zhao, Adam Zlotnick, et Uri Raviv. 2020. « Rapidly Forming Early Intermediate Structures Dictate the Pathway of Capsid Assembly ». *Journal of the American Chemical Society* 142 (17): 7868-82. <https://doi.org/10.1021/jacs.0c01092>.

Baltimore, D. 1971. « Expression of Animal Virus Genomes ». *Bacteriological Reviews* 35 (3): 235-41. <https://doi.org/10.1128/br.35.3.235-241.1971>.

Bartenschlager, R, et H Schaller. 1988. « The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. » *The EMBO Journal* 7 (13): 4185-92.

Berke, Jan Martin, Pascale Dehertogh, Karen Vergauwen, Ellen Van Damme, Wendy Mostmans, Koen Vandyck, et Frederik Pauwels. 2017. « Capsid Assembly Modulators Have a Dual Mechanism of Action in Primary Human Hepatocytes Infected with Hepatitis B Virus ». *Antimicrobial Agents and Chemotherapy* 61 (8). <https://doi.org/10.1128/AAC.00560-17>.

Birnbaum, F., et M. Nassal. 1990. « Hepatitis B Virus Nucleocapsid Assembly: Primary Structure Requirements in the Core Protein ». *Journal of Virology* 64 (7): 3319-30. <https://doi.org/10.1128/JVI.64.7.3319-3330.1990>.

Blumberg, B. S., B. J. Gerstley, D. A. Hungerford, W. T. London, et A. I. Sutnick. 1967. « A Serum Antigen (Australia Antigen) in Down's Syndrome, Leukemia, and Hepatitis ». *Annals of Internal Medicine* 66 (5): 924-31. <https://doi.org/10.7326/0003-4819-66-5-924>.

Bock, C. T., S. Schwinn, S. Locarnini, J. Fyfe, M. P. Manns, C. Trautwein, et H. Zentgraf. 2001. « Structural Organization of the Hepatitis B Virus Minichromosome ». *Journal of Molecular Biology* 307 (1): 183-96. <https://doi.org/10.1006/jmbi.2000.4481>.

Bond, WalterW., MartinS. Favero, NormanJ. Petersen, CliftonR. Gravelle, JamesW. Ebert, et JamesE. Maynard. 1981. « SURVIVAL OF HEPATITIS B VIRUS AFTER DRYING AND STORAGE FOR ONE WEEK ». *The Lancet*, Originally published as Volume 1, Issue 8219, 317 (8219): 550-51. [https://doi.org/10.1016/S0140-6736\(81\)92877-4](https://doi.org/10.1016/S0140-6736(81)92877-4).

Bourne, Christina R., M. G. Finn, et Adam Zlotnick. 2006. « Global Structural Changes in Hepatitis B Virus Capsids Induced by the Assembly Effector HAP1 ». *Journal of Virology* 80 (22): 11055-61. <https://doi.org/10.1128/JVI.00933-06>.

Bruinsma, Robijn F., et William S. Klug. 2015. « Physics of Viral Shells ». *Annual Review of Condensed Matter Physics* 6 (1): 245-68. <https://doi.org/10.1146/annurev-conmatphys-031214-014325>.

Brunetto, Maurizia Rossana, Filippo Oliveri, Piero Colombatto, Francesco Moriconi, Pietro Ciccorossi, Barbara Coco, Veronica Romagnoli, et al. 2010. « Hepatitis B Surface Antigen Serum Levels Help to Distinguish Active from Inactive Hepatitis B Virus Genotype D Carriers ». *Gastroenterology* 139 (2): 483-90. <https://doi.org/10.1053/j.gastro.2010.04.052>.

Bruss, V., et D. Ganem. 1991. « The Role of Envelope Proteins in Hepatitis B Virus Assembly ». *Proceedings of the National Academy of Sciences of the United States of America* 88 (3): 1059-63. <https://doi.org/10.1073/pnas.88.3.1059>.

Burke, Andrew R., Maelenn Chevreuil, A. Paris, Vanessa de La Grange, C. Goldmann, Javier Perez, D. Constantin, et G. Tresset. 2018. « Nanoparticle-Templated Self-Assembly of Viral Capsids Probed by Time-Resolved Absorbance Spectroscopy and X-Ray Scattering ». <https://doi.org/10.1103/PHYSREVAPPLIED.10.054065>.

Campagna, Matthew R., Fei Liu, Richeng Mao, Courtney Mills, Dawei Cai, Fang Guo, Xuesen Zhao, et al. 2013. « Sulfamoylbenzamide Derivatives Inhibit the Assembly of Hepatitis B Virus Nucleocapsids ». *Journal of Virology* 87 (12): 6931-42. <https://doi.org/10.1128/JVI.00582-13>.

Caspar, D. L., et A. Klug. 1962. « Physical Principles in the Construction of Regular Viruses ». *Cold Spring Harbor Symposia on Quantitative Biology* 27: 1-24. <https://doi.org/10.1101/sqb.1962.027.001.005>.

Ceres, Pablo, et Adam Zlotnick. 2002. « Weak Protein-Protein Interactions Are Sufficient to Drive Assembly of Hepatitis B Virus Capsids ». *Biochemistry* 41 (39): 11525-31. <https://doi.org/10.1021/bi0261645>.

Chabrolles, H el ene, Thomas Lahlali, H elo ise Auclair, et Anna Salvetti. 2018. « Les multiples r oles de la prot eine Core du virus de l'h epatite B: Nouvelles pistes de recherche et enjeux th erapeutiques ». *m edecine/sciences* 34 (8-9): 693-700. <https://doi.org/10.1051/medsci/20183408016>.

Chang, Matthew S., et Mindie H. Nguyen. 2017. « Epidemiology of Hepatitis B and the Role of Vaccination ». *Best Practice & Research. Clinical Gastroenterology* 31 (3): 239-47. <https://doi.org/10.1016/j.bpg.2017.05.008>.

Chen, Chao, Joseph Che-Yen Wang, Elizabeth E. Pierson, David Z. Keifer, Mildred Delaleau, Lara Gallucci, Christian Cazenave, Michael Kann, Martin F. Jarrold, et Adam Zlotnick. 2016. « Importin  $\beta$  Can Bind Hepatitis B Virus Core Protein and Empty Core-Like Particles and

Induce Structural Changes ». *PLOS Pathogens* 12 (8): e1005802. <https://doi.org/10.1371/journal.ppat.1005802>.

Chen, Jingzhi, Maelenn Chevreuil, S. Combet, Y. Lansac, et G. Tresset. 2017. « Investigating the thermal dissociation of viral capsid by lattice model ». <https://doi.org/10.1088/1361-648X/aa8d88>.

Chevreuil, Maelenn, Didier Law-Hine, Jingzhi Chen, Stéphane Bressanelli, Sophie Combet, Doru Constantin, Jéril Degrouard, Johannes Möller, Mehdi Zeghal, et Guillaume Tresset. 2018. « Nonequilibrium Self-Assembly Dynamics of Icosahedral Viral Capsids Packaging Genome or Polyelectrolyte ». *Nature Communications* 9 (1): 3071. <https://doi.org/10.1038/s41467-018-05426-8>.

Chevreuil, Maelenn, L. Lecoq, Shishan Wang, Laetitia Gargowitsch, N. Nhiri, E. Jacquet, T. Zinn, S. Fieulaine, S. Bressanelli, et G. Tresset. 2020. « Nonsymmetrical Dynamics of the HBV Capsid Assembly and Disassembly Evidenced by Their Transient Species. » <https://doi.org/10.1021/acs.jpcc.0c05024>.

Clark, Daniel N, et Jianming Hu. 2015. « Unveiling the roles of HBV polymerase for new antiviral strategies ». *Future virology* 10 (3): 283-95. <https://doi.org/10.2217/fvl.14.113>.

Cole, Andrew G. 2016. « Modulators of HBV Capsid Assembly as an Approach to Treating Hepatitis B Virus Infection ». *Current Opinion in Pharmacology* 30 (octobre): 131-37. <https://doi.org/10.1016/j.coph.2016.08.004>.

Corcuera, Angelica, Katharina Stolle, Stefan Hillmer, Stefan Seitz, Ji-Young Lee, Ralf Bartenschlager, Alexander Birkmann, et Andreas Urban. 2018. « Novel Non-Heteroarylpyrimidine (HAP) Capsid Assembly Modifiers Have a Different Mode of Action from HAPs in Vitro ». *Antiviral Research* 158 (octobre): 135-42. <https://doi.org/10.1016/j.antiviral.2018.07.011>.

Crouchet, Emilie, Catherine Schuster, et Thomas F. Baumert. 2021. « Liver cell circuits and therapeutic discovery for advanced liver disease and cancer ». *Comptes Rendus. Biologies* 344 (3): 233-48. <https://doi.org/10.5802/crbio.64>.

Dane, D. S., C. H. Cameron, et M. Briggs. 1970. « Virus-like Particles in Serum of Patients with Australia-Antigen-Associated Hepatitis ». *Lancet (London, England)* 1 (7649): 695-98. [https://doi.org/10.1016/s0140-6736\(70\)90926-8](https://doi.org/10.1016/s0140-6736(70)90926-8).

Dény, P., et F. Zoulim. 2010. « Hepatitis B virus: From diagnosis to treatment ». *Pathologie Biologie, Hépatite B*, 58 (4): 245-53. <https://doi.org/10.1016/j.patbio.2010.05.002>.

Diab, Ahmed, Adrien Foca, Fabien Zoulim, David Durantel, et Ourania Andrisani. 2018. « The diverse functions of the hepatitis B core/capsid protein (HBc) in the viral life cycle: Implications for the development of HBc-targeting antivirals ». *Antiviral research* 149 (janvier): 211-20. <https://doi.org/10.1016/j.antiviral.2017.11.015>.

Duriez, Marion, Anne Thouard, Stéphane Bressanelli, Jean-Michel Rossignol, et Delphine Sitterlin. 2014. « Conserved Aromatic Residues of the Hepatitis B Virus Precore Propeptide Are Involved in a Switch between Distinct Dimeric Conformations and Essential in the Formation of Heterocapsids ». *Virology* 462-463 (août): 273-82. <https://doi.org/10.1016/j.virol.2014.06.013>.

Eble, B. E., V. R. Lingappa, et D. Ganem. 1986. « Hepatitis B Surface Antigen: An Unusual Secreted Protein Initially Synthesized as a Transmembrane Polypeptide ». *Molecular and Cellular Biology* 6 (5): 1454-63. <https://doi.org/10.1128/mcb.6.5.1454-1463.1986>.

Endres, Dan, et Adam Zlotnick. 2002. « Model-Based Analysis of Assembly Kinetics for Virus Capsids or Other Spherical Polymers ». *Biophysical Journal* 83 (2): 1217-30. [https://doi.org/10.1016/S0006-3495\(02\)75245-4](https://doi.org/10.1016/S0006-3495(02)75245-4).

Galibert, F., E. Mandart, F. Fitoussi, P. Tiollais, et P. Charnay. 1979. « Nucleotide Sequence of the Hepatitis B Virus Genome (Subtype Ayw) Cloned in E. Coli ». *Nature* 281 (5733): 646-50. <https://doi.org/10.1038/281646a0>.

Gallina, A, F Bonelli, L Zentilin, G Rindi, M Muttini, et G Milanese. 1989. « A recombinant hepatitis B core antigen polypeptide with the protamine-like domain deleted self-assembles into capsid particles but fails to bind nucleic acids. » *Journal of Virology* 63 (11): 4645-52.

Gallucci, Lara, et Michael Kann. 2017. « Nuclear Import of Hepatitis B Virus Capsids and Genome ». *Viruses* 9 (1): 21. <https://doi.org/10.3390/v9010021>.

Ganem, Don, et Harold E. Varmus. 1987. « The Molecular Biology of the Hepatitis B Viruses ». *Annual Review of Biochemistry* 56 (1): 651-93. <https://doi.org/10.1146/annurev.bi.56.070187.003251>.

Garmann, Rees F., Mauricio Comas-Garcia, Ajaykumar Gopal, Charles M. Knobler, et William M. Gelbart. 2014. « The Assembly Pathway of an Icosahedral Single-Stranded RNA Virus Depends on the Strength of Inter-Subunit Attractions ». *Journal of Molecular Biology, Molecular Interplay in Viral Assembly*, 426 (5): 1050-60. <https://doi.org/10.1016/j.jmb.2013.10.017>.

Glebe, Dieter, Stephan Urban, Eva V. Knoop, Nilgün Cag, Peter Krass, Stefanie Grün, Aiste Bulavaite, Kestutis Sasnauskas, et Wolfram H. Gerlich. 2005. « Mapping of the Hepatitis B

Virus Attachment Site by Use of Infection-Inhibiting PreS1 Lipopeptides and Tupaia Hepatocytes ». *Gastroenterology* 129 (1): 234-45. <https://doi.org/10.1053/j.gastro.2005.03.090>.

Goldstein, Susan T, Fangjun Zhou, Stephen C Hadler, Beth P Bell, Eric E Mast, et Harold S Margolis. 2005. « A mathematical model to estimate global hepatitis B disease burden and vaccination impact ». *International Journal of Epidemiology* 34 (6): 1329-39. <https://doi.org/10.1093/ije/dyi206>.

Guo, Fang, Qiong Zhao, Muhammad Sheraz, Junjun Cheng, Yonghe Qi, Qing Su, Andrea Cuconati, et al. 2017. « HBV Core Protein Allosteric Modulators Differentially Alter CccDNA Biosynthesis from de Novo Infection and Intracellular Amplification Pathways ». Édité par Aleem Siddiqui. *PLOS Pathogens* 13 (9): e1006658. <https://doi.org/10.1371/journal.ppat.1006658>.

Guo, Yanhai, Wei Kang, Xiaoying Lei, Yongnian Li, An Xiang, Yonglan Liu, Jinrong Zhao, Ju Zhang, et Zhen Yan. 2012. « Hepatitis B viral core protein disrupts human host gene expression by binding to promoter regions ». *BMC Genomics* 13 (1): 563. <https://doi.org/10.1186/1471-2164-13-563>.

Guo, Yan-Hai, Yong-Nian Li, Jin-Rong Zhao, Ju Zhang, et Zhen Yan. 2011. « HBc Binds to the CpG Islands of HBV CccDNA and Promotes an Epigenetic Permissive State ». *Epigenetics* 6 (6): 720-26. <https://doi.org/10.4161/epi.6.6.15815>.

Heermann, K. H., U. Goldmann, W. Schwartz, T. Seyffarth, H. Baumgarten, et W. H. Gerlich. 1984. « Large Surface Proteins of Hepatitis B Virus Containing the Pre-s Sequence ». *Journal of Virology* 52 (2): 396-402. <https://doi.org/10.1128/JVI.52.2.396-402.1984>.

Ho, Joan Kha-Tu, Beena Jeevan-Raj, et Hans-Jürgen Netter. 2020. « Hepatitis B Virus (HBV) Subviral Particles as Protective Vaccines and Vaccine Platforms ». *Viruses* 12 (2): 126. <https://doi.org/10.3390/v12020126>.

Hollinger, F. Blaine, et Daryl T.-Y. Lau. 2006. « Hepatitis B: The Pathway to Recovery through Treatment ». *Gastroenterology Clinics of North America* 35 (4): 895-931. <https://doi.org/10.1016/j.gtc.2006.10.002>.

Huang, Jian, Qing Deng, Qun Wang, Kun-Yu Li, Ji-Hong Dai, Niu Li, Zhi-Dong Zhu, et al. 2012. « Exome Sequencing of Hepatitis B Virus-Associated Hepatocellular Carcinoma ». *Nature Genetics* 44 (10): 1117-21. <https://doi.org/10.1038/ng.2391>.

Hulo, Chantal, Edouard de Castro, Patrick Masson, Lydie Bougueleret, Amos Bairoch, Ioannis Xenarios, et Philippe Le Mercier. 2011. « ViralZone: a knowledge resource to understand virus

diversity ». *Nucleic Acids Research* 39 (Database issue): D576-82. <https://doi.org/10.1093/nar/gkq901>.

Jegaskanda, S., S. H. Ahn, N. Skinner, A. J. Thompson, T. Ngyuen, J. Holmes, R. De Rose, et al. 2014. « Downregulation of Interleukin-18-Mediated Cell Signaling and Interferon Gamma Expression by the Hepatitis B Virus e Antigen ». *Journal of Virology* 88 (18): 10412-20. <https://doi.org/10.1128/JVI.00111-14>.

Jones, Scott A., et Jianming Hu. 2013. « Hepatitis B Virus Reverse Transcriptase: Diverse Functions as Classical and Emerging Targets for Antiviral Intervention ». *Emerging Microbes & Infections* 2 (9): e56. <https://doi.org/10.1038/emi.2013.56>.

Kang, Jung-Ah, Songwon Kim, Minji Park, Hyun-Jin Park, Jeong-Hyun Kim, Sanghyeok Park, Jeong-Ryul Hwang, et al. 2019. « Ciclopirox Inhibits Hepatitis B Virus Secretion by Blocking Capsid Assembly ». *Nature Communications* 10 (1): 2184. <https://doi.org/10.1038/s41467-019-10200-5>.

Kann, M, et W H Gerlich. 1994. « Effect of core protein phosphorylation by protein kinase C on encapsidation of RNA within core particles of hepatitis B virus. » *Journal of Virology* 68 (12): 7993-8000.

Katen, S., S. Chirapu, M. G. Finn, et A. Zlotnick. 2010. « Trapping of hepatitis B virus capsid assembly intermediates by phenylpropenamide assembly accelerators. » <https://doi.org/10.1021/cb100275b>.

Kler, Stanislav, Roi Asor, Chenglei Li, Avi Ginsburg, Daniel Harries, Ariella Oppenheim, Adam Zlotnick, et Uri Raviv. 2012. « RNA Encapsidation by SV40-Derived Nanoparticles Follows a Rapid Two-State Mechanism ». *Journal of the American Chemical Society* 134 (21): 8823-30. <https://doi.org/10.1021/ja2110703>.

Klumpp, Klaus, Angela M. Lam, Christine Lukacs, Robert Vogel, Suping Ren, Christine Espiritu, Ruth Baydo, et al. 2015. « High-Resolution Crystal Structure of a Hepatitis B Virus Replication Inhibitor Bound to the Viral Core Protein ». *Proceedings of the National Academy of Sciences* 112 (49): 15196-201. <https://doi.org/10.1073/pnas.1513803112>.

Ko, Chunkyu, Sooyoung Lee, Marc P. Windisch, et Wang-Shick Ryu. 2014. « DDX3 DEAD-Box RNA Helicase Is a Host Factor That Restricts Hepatitis B Virus Replication at the Transcriptional Level ». *Journal of Virology* 88 (23): 13689-98. <https://doi.org/10.1128/JVI.02035-14>.

Kra, Kalouna, Siyu Li, Laetitia Gargowitsch, Jéril Degrouard, Javier Pérez, Roya Zandi, Stéphane Bressanelli, et Guillaume Tresset. 2023. « Energetics and Kinetic Assembly

Pathways of Hepatitis B Virus Capsids in the Presence of Antivirals ». *ACS Nano* 17 (13): 12723-33. <https://doi.org/10.1021/acsnano.3c03595>.

Lahlali, Thomas, Jan Martin Berke, Karen Vergauwen, Adrien Foca, Koen Vandyck, Frederik Pauwels, Fabien Zoulim, et David Durantel. 2018. « Novel Potent Capsid Assembly Modulators Regulate Multiple Steps of the Hepatitis B Virus Life Cycle ». *Antimicrobial Agents and Chemotherapy* 62 (10): e00835-18. <https://doi.org/10.1128/AAC.00835-18>.

Lang, Tali, Camden Lo, Narelle Skinner, Stephen Locarnini, Kumar Visvanathan, et Ashley Mansell. 2011. « The Hepatitis B e Antigen (HBeAg) Targets and Suppresses Activation of the Toll-like Receptor Signaling Pathway ». *Journal of Hepatology* 55 (4): 762-69. <https://doi.org/10.1016/j.jhep.2010.12.042>.

Lavanchy, D. 2004. « Hepatitis B Virus Epidemiology, Disease Burden, Treatment, and Current and Emerging Prevention and Control Measures ». *Journal of Viral Hepatitis* 11 (2): 97-107. <https://doi.org/10.1046/j.1365-2893.2003.00487.x>.

Law-Hine, Didier, Anil K. Sahoo, Virginie Bailleux, Mehdi Zeghal, Sylvain Prevost, Prabal K. Maiti, Stéphane Bressanelli, Doru Constantin, et Guillaume Tresset. 2015. « Reconstruction of the Disassembly Pathway of an Icosahedral Viral Capsid and Shape Determination of Two Successive Intermediates ». *The Journal of Physical Chemistry Letters* 6 (17): 3471-76. <https://doi.org/10.1021/acs.jpcclett.5b01478>.

Law-Hine, Didier, Mehdi Zeghal, Stéphane Bressanelli, Doru Constantin, et Guillaume Tresset. 2016. « Identification of a Major Intermediate along the Self-Assembly Pathway of an Icosahedral Viral Capsid by Using an Analytical Model of a Spherical Patch ». *Soft Matter* 12 (32): 6728-36. <https://doi.org/10.1039/C6SM01060A>.

Le Duff, Yann, Matthieu Blanchet, et Camille Sureau. 2009. « The Pre-S1 and Antigenic Loop Infectivity Determinants of the Hepatitis B Virus Envelope Proteins Are Functionally Independent ». *Journal of Virology* 83 (23): 12443-51. <https://doi.org/10.1128/JVI.01594-09>.

Lecoq, Lauriane, Louis Brigandat, Rebecca Huber, Marie-Laure Fogeron, Shishan Wang, Marie Dujardin, Mathilde Briday, et al. 2023. « Molecular Elucidation of Drug-Induced Abnormal Assemblies of the Hepatitis B Virus Capsid Protein by Solid-State NMR ». *Nature Communications* 14 (1): 471. <https://doi.org/10.1038/s41467-023-36219-3>.

Locarnini, Stephen, et Scott Bowden. 2012. « Hepatitis B Surface Antigen Quantification: Not What It Seems on the Surface ». *Hepatology* 56 (2): 411. <https://doi.org/10.1002/hep.25732>.

- Locarnini, Stephen, Angelos Hatzakis, Ding-Shinn Chen, et Anna Lok. 2015. « Strategies to Control Hepatitis B: Public Policy, Epidemiology, Vaccine and Drugs ». *Journal of Hepatology* 62 (1 Suppl): S76-86. <https://doi.org/10.1016/j.jhep.2015.01.018>.
- Locarnini, Stephen, Margaret Littlejohn, Muhammad Nazri Aziz, et Lilly Yuen. 2013. « Possible Origins and Evolution of the Hepatitis B Virus (HBV) ». *Seminars in Cancer Biology* 23 (6 Pt B): 561-75. <https://doi.org/10.1016/j.semcancer.2013.08.006>.
- Lucifora, Julie, Silke Arzberger, David Durantel, Laura Belloni, Michel Strubin, Massimo Levrero, Fabien Zoulim, Olivier Hantz, et Ulrike Protzer. 2011. « Hepatitis B Virus X Protein Is Essential to Initiate and Maintain Virus Replication after Infection ». *Journal of Hepatology* 55 (5): 996-1003. <https://doi.org/10.1016/j.jhep.2011.02.015>.
- Lucifora, Julie, Yuchen Xia, Florian Reisinger, Ke Zhang, Daniela Stadler, Xiaoming Cheng, Martin F. Sprinzl, et al. 2014. « Specific and Nonhepatotoxic Degradation of Nuclear Hepatitis B Virus CccDNA ». *Science (New York, N.Y.)* 343 (6176): 1221-28. <https://doi.org/10.1126/science.1243462>.
- Lurman A. 1885. « Eine icterus epidemic ». *Berl Klin Wochenschr*, n° 22: 20-23.
- Lutomski, Corinne A., N. Lykтей, E. Pierson, Zhongchao Zhao, A. Zlotnick, et M. Jarrold. 2018. « Multiple Pathways in Capsid Assembly. » <https://doi.org/10.1021/jacs.8b01804>.
- Macovei, Alina, Cristina Radulescu, Catalin Lazar, Stefana Petrescu, David Durantel, Raymond A. Dwek, Nicole Zitzmann, et Norica Branza-Nichita. 2010. « Hepatitis B Virus Requires Intact Caveolin-1 Function for Productive Infection in HepaRG Cells ». *Journal of Virology* 84 (1): 243-53. <https://doi.org/10.1128/JVI.01207-09>.
- Mak, Lung-Yi, Danny Ka-Ho Wong, Wai-Kay Seto, Ching-Lung Lai, et Man Fung Yuen. 2017. « Hepatitis B core protein as a therapeutic target ». *Expert Opinion on Therapeutic Targets* 21 (12): 1153-59. <https://doi.org/10.1080/14728222.2017.1397134>.
- Mohebbi, Alireza, Toubia Ghorbanzadeh, Shabnam Naderifar, Fattaneh Khalaj, Fatemeh Sana Askari, et Ali Salehnia Sammak. 2023. « A Fragment-Based Drug Discovery Developed on Cyclopirox for Inhibition of Hepatitis B Virus Core Protein: An in Silico Study ». *PloS One* 18 (5): e0285941. <https://doi.org/10.1371/journal.pone.0285941>.
- Mühlemann, Barbara, Terry C. Jones, Peter de Barros Damgaard, Morten E. Allentoft, Irina Shevnina, Andrey Logvin, Emma Usmanova, et al. 2018. « Ancient Hepatitis B Viruses from the Bronze Age to the Medieval Period ». *Nature* 557 (7705): 418-23. <https://doi.org/10.1038/s41586-018-0097-z>.

Nassal, M. 1992. « The Arginine-Rich Domain of the Hepatitis B Virus Core Protein Is Required for Pregenome Encapsidation and Productive Viral Positive-Strand DNA Synthesis but Not for Virus Assembly ». *Journal of Virology* 66 (7): 4107-16. <https://doi.org/10.1128/JVI.66.7.4107-4116.1992>.

Nassal, Michael. 2008. « Hepatitis B Viruses: Reverse Transcription a Different Way ». *Virus Research* 134 (1-2): 235-49. <https://doi.org/10.1016/j.virusres.2007.12.024>.

———. 2015. « HBV CccDNA: Viral Persistence Reservoir and Key Obstacle for a Cure of Chronic Hepatitis B ». *Gut* 64 (12): 1972-84. <https://doi.org/10.1136/gutjnl-2015-309809>.

Ni, Yi, Florian A. Lempp, Stefan Mehrle, Shirin Nkongolo, Christina Kaufman, Maria Fälth, Jan Stindt, et al. 2014. « Hepatitis B and D Viruses Exploit Sodium Taurocholate Co-Transporting Polypeptide for Species-Specific Entry into Hepatocytes ». *Gastroenterology* 146 (4): 1070-83. <https://doi.org/10.1053/j.gastro.2013.12.024>.

Nijampatnam, Bhavitavya, et Dennis C Liotta. 2019. « Recent Advances in the Development of HBV Capsid Assembly Modulators ». *Current Opinion in Chemical Biology* 50 (juin): 73-79. <https://doi.org/10.1016/j.cbpa.2019.02.009>.

Ning, Xiaojun, David Nguyen, Laura Mentzer, Christina Adams, Hyunwook Lee, Robert Ashley, Susan Hafenstein, et Jianming Hu. 2011. « Secretion of Genome-Free Hepatitis B Virus – Single Strand Blocking Model for Virion Morphogenesis of Para-Retrovirus ». Édité par Jing-hsiung James Ou. *PLoS Pathogens* 7 (9): e1002255. <https://doi.org/10.1371/journal.ppat.1002255>.

Oliver, Ryan C., Wojciech Potrzebowski, Seyed Morteza Najibi, Martin Nors Pedersen, Lise Arleth, Najet Mahmoudi, et Ingemar André. 2020. « Assembly of Capsids from Hepatitis B Virus Core Protein Progresses through Highly Populated Intermediates in the Presence and Absence of RNA ». *ACS Nano* 14 (8): 10226-38. <https://doi.org/10.1021/acsnano.0c03569>.

Patterson, Angela, Zhongchao Zhao, Elizabeth Waymire, Adam Zlotnick, et Brian Bothner. 2020. « Dynamics of Hepatitis B Virus Capsid Protein Dimer Regulate Assembly through an Allosteric Network ». *ACS Chemical Biology* 15 (8): 2273-80. <https://doi.org/10.1021/acscchembio.0c00481>.

Perlmutter, Jason D., et Michael F. Hagan. 2015. « Mechanisms of Virus Assembly ». *Annual Review of Physical Chemistry* 66 (avril): 217-39. <https://doi.org/10.1146/annurev-physchem-040214-121637>.

- Pierson, E., David Z. Keifer, Lisa Selzer, L. Lee, Nathan C. Contino, J. Wang, A. Zlotnick, et M. Jarrold. 2014. « Detection of Late Intermediates in Virus Capsid Assembly by Charge Detection Mass Spectrometry ». <https://doi.org/10.1021/ja411460w>.
- Prasad, B. V. Venkataram, et Michael F. Schmid. 2011. « Principles of Virus Structural Organization ». *Viral Molecular Machines* 726 (novembre): 17-47. [https://doi.org/10.1007/978-1-4614-0980-9\\_3](https://doi.org/10.1007/978-1-4614-0980-9_3).
- Prifti, Georgia-Myrto, Dimitrios Moianos, Erofilii Giannakopoulou, Vasiliki Pardali, John Tavis, et Grigoris Zoidis. 2021. « Recent Advances in Hepatitis B Treatment ». *Pharmaceuticals* 14 (5): 417. <https://doi.org/10.3390/ph14050417>.
- Prince, A. M. 1968. « An Antigen Detected in the Blood during the Incubation Period of Serum Hepatitis ». *Proceedings of the National Academy of Sciences of the United States of America* 60 (3): 814-21. <https://doi.org/10.1073/pnas.60.3.814>.
- Rabe, Birgit, Mildred Delaleau, Andreas Bischof, Michael Foss, Irina Sominskaya, Paul Pumpens, Christian Cazenave, Michel Castroviejo, et Michael Kann. 2009. « Nuclear Entry of Hepatitis B Virus Capsids Involves Disintegration to Protein Dimers Followed by Nuclear Reassociation to Capsids ». Édité par Michael Farzan. *PLoS Pathogens* 5 (8): e1000563. <https://doi.org/10.1371/journal.ppat.1000563>.
- Rabe, Birgit, Dieter Glebe, et Michael Kann. 2006. « Lipid-Mediated Introduction of Hepatitis B Virus Capsids into Nonsusceptible Cells Allows Highly Efficient Replication and Facilitates the Study of Early Infection Events ». *Journal of Virology* 80 (11): 5465-73. <https://doi.org/10.1128/JVI.02303-05>.
- Radziwill, G., W. Tucker, et H. Schaller. 1990. « Mutational Analysis of the Hepatitis B Virus P Gene Product: Domain Structure and RNase H Activity ». *Journal of Virology* 64 (2): 613-20. <https://doi.org/10.1128/JVI.64.2.613-620.1990>.
- Rajoriya, Neil, Christophe Combet, Fabien Zoulim, et Harry L. A. Janssen. 2017. « How Viral Genetic Variants and Genotypes Influence Disease and Treatment Outcome of Chronic Hepatitis B. Time for an Individualised Approach? » *Journal of Hepatology* 67 (6): 1281-97. <https://doi.org/10.1016/j.jhep.2017.07.011>.
- Reddy, Tyler, et Mark S. P. Sansom. 2016. « Computational virology: From the inside out ». *Biochimica et Biophysica Acta (BBA) - Biomembranes*, New approaches for bridging computation and experiment on membrane proteins, 1858 (7, Part B): 1610-18. <https://doi.org/10.1016/j.bbamem.2016.02.007>.

- Schlicksup, Christopher John, Joseph Che-Yen Wang, Samson Francis, Balasubramanian Venkatakrishnan, William W Turner, Michael VanNieuwenhze, et Adam Zlotnick. 2018. « Hepatitis B virus core protein allosteric modulators can distort and disrupt intact capsids ». Édité par Karla Kirkegaard. *eLife* 7 (janvier): e31473. <https://doi.org/10.7554/eLife.31473>.
- Schmitz, André, Alexandra Schwarz, Michael Foss, Lixin Zhou, Birgit Rabe, Julia Hoellenriegel, Miriam Stoeber, Nelly Panté, et Michael Kann. 2010. « Nucleoporin 153 Arrests the Nuclear Import of Hepatitis B Virus Capsids in the Nuclear Basket ». *PLOS Pathogens* 6 (1): e1000741. <https://doi.org/10.1371/journal.ppat.1000741>.
- Schreiner, et Nassal. 2017. « A Role for the Host DNA Damage Response in Hepatitis B Virus CccDNA Formation-and Beyond? » *Viruses* 9 (5). <https://doi.org/10.3390/v9050125>.
- Schweitzer, Aparna, Johannes Horn, Rafael T. Mikolajczyk, Gérard Krause, et Jördis J. Ott. 2015. « Estimations of Worldwide Prevalence of Chronic Hepatitis B Virus Infection: A Systematic Review of Data Published between 1965 and 2013 ». *Lancet (London, England)* 386 (10003): 1546-55. [https://doi.org/10.1016/S0140-6736\(15\)61412-X](https://doi.org/10.1016/S0140-6736(15)61412-X).
- Selzer, Lisa, Ravi Kant, Joseph C.-Y. Wang, Brian Bothner, et Adam Zlotnick. 2015. « Hepatitis B Virus Core Protein Phosphorylation Sites Affect Capsid Stability and Transient Exposure of the C-Terminal Domain ». *Journal of Biological Chemistry* 290 (47): 28584-93. <https://doi.org/10.1074/jbc.M115.678441>.
- Stray, S. J., C. R. Bourne, S. Punna, W. G. Lewis, M. G. Finn, et A. Zlotnick. 2005. « A Heteroaryldihydropyrimidine Activates and Can Misdirect Hepatitis B Virus Capsid Assembly ». *Proceedings of the National Academy of Sciences* 102 (23): 8138-43. <https://doi.org/10.1073/pnas.0409732102>.
- Taverniti, Valerio, Gaëtan Ligat, Yannick Debing, Dieudonne Buh Kum, Thomas F. Baumert, et Eloi R. Verrier. 2022. « Capsid Assembly Modulators as Antiviral Agents against HBV: Molecular Mechanisms and Clinical Perspectives ». *Journal of Clinical Medicine* 11 (5): 1349. <https://doi.org/10.3390/jcm11051349>.
- Tong, Shuping, et Peter Revill. 2016. « Overview of Hepatitis B Viral Replication and Genetic Variability ». *Journal of Hepatology* 64 (1 Suppl): S4-16. <https://doi.org/10.1016/j.jhep.2016.01.027>.
- Tresset, Guillaume, Clémence Le Coeur, Jean-François Bryche, Mouna Tatou, Mehdi Zeghal, Annie Charpilienne, Didier Poncet, Doru Constantin, et Stéphane Bressanelli. 2013. « Norovirus Capsid Proteins Self-Assemble through Biphasic Kinetics via Long-Lived Stave-

like Intermediates ». *Journal of the American Chemical Society* 135 (41): 15373-81. <https://doi.org/10.1021/ja403550f>.

Urban, Stephan, Ralf Bartenschlager, Ralf Kubitz, et Fabien Zoulim. 2014. « Strategies to Inhibit Entry of HBV and HDV into Hepatocytes ». *Gastroenterology* 147 (1): 48-64. <https://doi.org/10.1053/j.gastro.2014.04.030>.

Venkatakrishnan, Balasubramanian, Sarah P. Katen, Samson Francis, Srinivas Chirapu, M. G. Finn, et Adam Zlotnick. 2016. « Hepatitis B Virus Capsids Have Diverse Structural Responses to Small-Molecule Ligands Bound to the Heteroaryldihydropyrimidine Pocket ». Édité par W. I. Sundquist. *Journal of Virology* 90 (8): 3994-4004. <https://doi.org/10.1128/JVI.03058-15>.

Venkatakrishnan, Balasubramanian, et Adam Zlotnick. 2016. « The Structural Biology of Hepatitis B Virus: Form and Function ». *Annual Review of Virology* 3 (1): 429-51. <https://doi.org/10.1146/annurev-virology-110615-042238>.

Wang, Cheng, Dong Ji, Jing Chen, Qing Shao, Bing Li, Jialiang Liu, Vanessa Wu, et al. 2017. « Hepatitis Due to Reactivation of Hepatitis B Virus in Endemic Areas Among Patients With Hepatitis C Treated With Direct-Acting Antiviral Agents ». *Clinical Gastroenterology and Hepatology: The Official Clinical Practice Journal of the American Gastroenterological Association* 15 (1): 132-36. <https://doi.org/10.1016/j.cgh.2016.06.023>.

Wang, Joseph C.-Y., Mary S. Dhason, et Adam Zlotnick. 2012. « Structural Organization of Pregenomic RNA and the Carboxy-Terminal Domain of the Capsid Protein of Hepatitis B Virus ». *PLoS Pathogens* 8 (9): e1002919. <https://doi.org/10.1371/journal.ppat.1002919>.

Wei, X., et D. L. Peterson. 1996. « Expression, Purification, and Characterization of an Active RNase H Domain of the Hepatitis B Viral Polymerase ». *The Journal of Biological Chemistry* 271 (51): 32617-22. <https://doi.org/10.1074/jbc.271.51.32617>.

WHO. 2020. « Hepatitis B ». WHO. 2020. <https://www.who.int/news-room/fact-sheets/detail/hepatitis-b>.

Wong, Danny Ka-Ho, Wai-Kay Seto, Ka-Shing Cheung, Chun-Kong Chong, Fung-Yu Huang, James Fung, Ching-Lung Lai, et Man-Fung Yuen. 2017. « Hepatitis B Virus Core-Related Antigen as a Surrogate Marker for Covalently Closed Circular DNA ». *Liver International: Official Journal of the International Association for the Study of the Liver* 37 (7): 995-1001. <https://doi.org/10.1111/liv.13346>.

Wynne, S. A., R. A. Crowther, et A. G. Leslie. 1999. « The Crystal Structure of the Human Hepatitis B Virus Capsid ». *Molecular Cell* 3 (6): 771-80. [https://doi.org/10.1016/s1097-2765\(01\)80009-5](https://doi.org/10.1016/s1097-2765(01)80009-5).

Yan, Huan, Guocai Zhong, Guangwei Xu, Wenhui He, Zhiyi Jing, Zhenchao Gao, Yi Huang, et al. 2012. « Sodium Taurocholate Cotransporting Polypeptide Is a Functional Receptor for Human Hepatitis B and D Virus ». *ELife* 1 (novembre): e00049. <https://doi.org/10.7554/eLife.00049>.

Zhou, Zheng, Taishan Hu, Xue Zhou, Steffen Wildum, Fernando Garcia-Alcalde, Zhiheng Xu, Daitze Wu, et al. 2017. « Heteroaryldihydropyrimidine (HAP) and Sulfamoylbenzamide (SBA) Inhibit Hepatitis B Virus Replication by Different Molecular Mechanisms ». *Scientific Reports* 7 (1): 42374. <https://doi.org/10.1038/srep42374>.

Zlotnick, Adam, Balasubramanian Venkatakrishnan, Zhenning Tan, Eric Lewellyn, William Turner, et Samson Francis. 2015. « Core Protein: A Pleiotropic Keystone in the HBV Lifecycle ». *Antiviral Research* 121 (septembre): 82-93. <https://doi.org/10.1016/j.antiviral.2015.06.020>.

Zoulim, Fabien. 2005. « New Insight on Hepatitis B Virus Persistence from the Study of Intrahepatic Viral CccDNA ». *Journal of Hepatology* 42 (3): 302-8. <https://doi.org/10.1016/j.jhep.2004.12.015>.

Zoulim, Fabien, et Stephen Locarnini. 2013. « Optimal Management of Chronic Hepatitis B Patients with Treatment Failure and Antiviral Drug Resistance ». *Liver International: Official Journal of the International Association for the Study of the Liver* 33 Suppl 1 (février): 116-24. <https://doi.org/10.1111/liv.12069>.

Zoulim, Fabien, Adam Zlotnick, Stephanie Buchholz, Eric Donaldson, John Fry, Anuj Gaggar, Jianming Hu, et al. 2022. « Nomenclature of HBV Core Protein-Targeting Antivirals ». *Nature Reviews Gastroenterology & Hepatology* 19 (12): 748-50. <https://doi.org/10.1038/s41575-022-00700-z>.

# Chapitre 3 : Principes des différentes méthodes utilisées

Afin d'étudier les effets des modulateurs d'assemblage sur la capsid du VHB, il est nécessaire dans un premier temps de produire et de purifier la protéine Core. Pour cela, la méthode d'expression de la Core chez *Escherichia coli* a été choisie. Le principe de cette méthode sera détaillé dans ce chapitre. La protéine est ensuite purifiée grâce à plusieurs techniques de purification telles que la centrifugation différentielle, la précipitation au sulfate d'ammonium et la purification par chromatographie d'exclusion de taille (SEC «*size exclusion chromatography*»). Le principe de la chromatographie d'exclusion de taille sera exposé en détail dans ce chapitre.

Une fois que la protéine Core est purifiée, les tests d'assemblage en présence des modulateurs d'assemblage peuvent être réalisés. Plusieurs méthodes permettant de caractériser les effets des modulateurs d'assemblage sur l'assemblage de la protéine Core ont été utilisées :

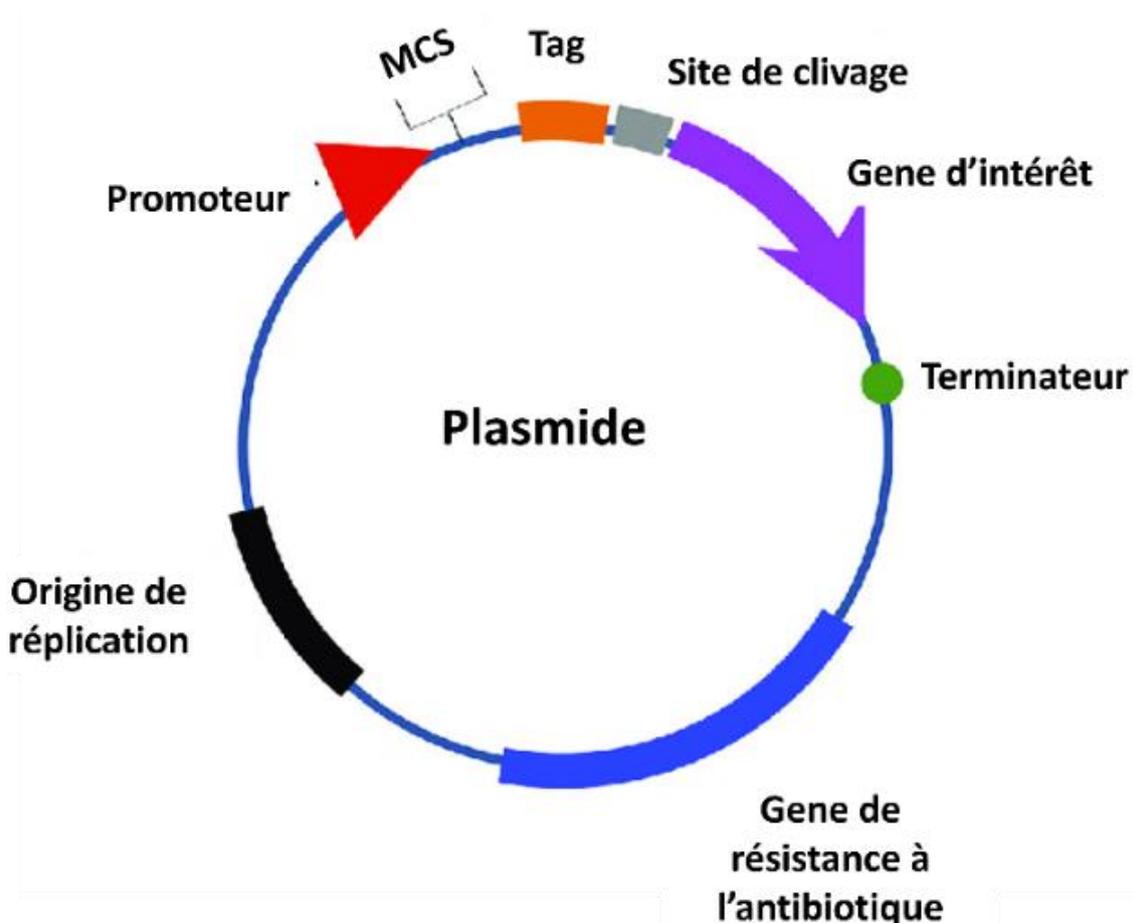
- La diffusion dynamique de la lumière,
- La photométrie de masse,
- La diffusion des rayons X aux petits angles résolue en temps ou non,
- La (cryo)microscopie électronique en transmission.

Les principes de ces techniques seront également présentés dans ce chapitre.

## 1. Expression et purification des protéines

### 1.1 Transformation et expression en système bactérien

Un vecteur d'expression, nommé ici vecteur pET et contenant le code génétique de cette protéine a été obtenu par clonage (Figure 3.1). Ce vecteur est constitué d'une séquence d'ADN circulaire contenant divers éléments permettant la synthèse de la protéine Core.



**Figure 3.1 : Schéma d'un vecteur d'expression.** Afin de faciliter la purification de la protéine d'intérêt, un tag peut être inclus dans le cadre de lecture du gène d'intérêt. Dans le cas du vecteur d'expression de la protéine Core, il n'y a ni tag, ni site de clivage. Adapté de (Mesa-Pereira et al. 2018).

### **Le promoteur**

Il s'agit d'une région spécifique reconnue par l'ARN polymérase, qui contrôle l'expression du gène d'intérêt. Dans notre cas, il s'agit d'un promoteur inducible, le promoteur T7. Ce promoteur contient un opéron LAC permettant l'induction de la transcription, uniquement en présence de lactose ou d'IPTG (isopropyl Beta-D-1-thiogalactopyranoside). Ainsi, il est possible de contrôler l'expression du gène d'intérêt et de l'activer seulement lorsque la quantité de bactéries nécessaire est présente dans le milieu de culture.

### **Le site de clonage multiple (MCS)**

Situé en aval du promoteur, il consiste en un segment d'ADN contenant de nombreux sites de restriction, conduisant ainsi l'insertion du gène à exprimer.

### **L'origine de réplication**

L'initiation de la réplication du plasmide est permise par cette séquence d'ADN. Dans notre cas, nous utilisons l'origine de réplication nommée pMB1 qui génère environ 15 à 20 copies du vecteur pET par bactérie.

### **Les marqueurs de sélection**

Ce sont des gènes de résistance aux antibiotiques utilisés pour la sélection clonale. Le vecteur peut ainsi contenir :

- Le gène exprimant la B-lactamase, conférant à la bactérie une résistance à l'ampicilline,
- Le gène exprimant la protéine chloramphénicol acétyl transférase, concédant une résistance au chloramphénicol ou,
- Le gène exprimant la protéine néomycine phosphotransférase, octroyant une résistance à la kanamycine.

Ces marqueurs sont essentiels pour la sélection des clones ayant le plasmide d'intérêt.

### **La séquence de Shine-Dalgarno**

Il s'agit d'une séquence située environ 10 nucléotides avant le codon ATG (le codon start), qui permet la liaison du ribosome pour initier la traduction.

### **Le terminateur de transcription**

Chez les procaryotes, la terminaison de la transcription nécessite une séquence de terminaison spécifique. Selon les gènes, deux types de terminateurs peuvent être distingués : La terminaison nécessitant l'intervention du facteur protéique Rho, appelée terminaison dépendant de rho, ou la terminaison intrinsèque, qui ne nécessite pas de facteur protéique.

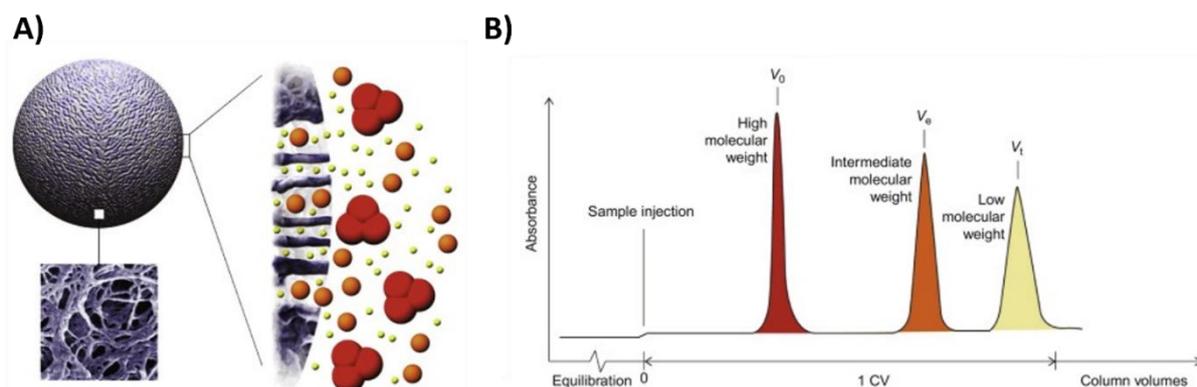
Une fois le vecteur obtenu, celui-ci est inséré dans les bactéries d'intérêt grâce à la technique de transformation bactérienne ou transgénèse. La transgénèse est un processus par lequel un plasmide d'intérêt est inséré dans une bactérie compétente afin qu'il soit exprimé. Les bactéries compétentes sont capables d'incorporer cet ADN soit par électroporation, autrement dit par l'application d'un choc électrique sur les membranes bactériennes afin de les perméabiliser, ou par choc thermique à 42°C.

Après la transformation, les bactéries sont cultivées dans un milieu de culture minimal contenant les antibiotiques permettant la sélection des bactéries possédant le plasmide d'intérêt. Une fois que la population bactérienne a atteint une densité suffisante dans le milieu, l'induction par ajout d'IPTG est réalisée afin de déclencher la synthèse de la protéine d'intérêt. Après l'induction, la température du milieu de culture est réduite pour ralentir la vitesse de synthèse de la protéine d'intérêt, et ainsi favoriser un repliement correct. Enfin, à la fin de la culture, les bactéries sont collectées en vue d'extraire les protéines.

L'utilisation de bactéries pour l'expression d'une protéine recombinante présente plusieurs avantages. Elle permet une croissance rapide dans un milieu de culture minimal, elle facilite le clonage du gène d'intérêt, et elle permet d'obtenir des niveaux d'expression très élevés.

## 1.2 Purification des protéines par chromatographie d'exclusion de taille

Après l'extraction des protéines, la protéine Core va être purifiée, notamment par chromatographie d'exclusion de taille.



**Figure 3.2 : Principe de la chromatographie d'exclusion de taille.** (A) Billes poreuses. (B) Exemple d'un profil d'élution. Les molécules de haut poids moléculaire sont éluées en premier, avant les molécules de faible poids moléculaire. Adaptée de (Hall 2018).

La chromatographie d'exclusion de taille (SEC, pour size exclusion chromatography) est une technique qui consiste en la séparation des molécules selon leur volume

hydrodynamique (Figure 3.2). Cette méthode repose sur l'utilisation de colonnes remplies d'un gel poreux et inerte pour la filtration de protéines, ainsi que d'un tampon d'éluion qui facilite le passage des molécules à travers la colonne. Les molécules entrent dans les pores du gel en fonction de leur taille. Celles de plus grande taille, pénétrant de ce fait moins facilement dans les pores, vont être éluées plus tôt que les molécules de plus petite taille.

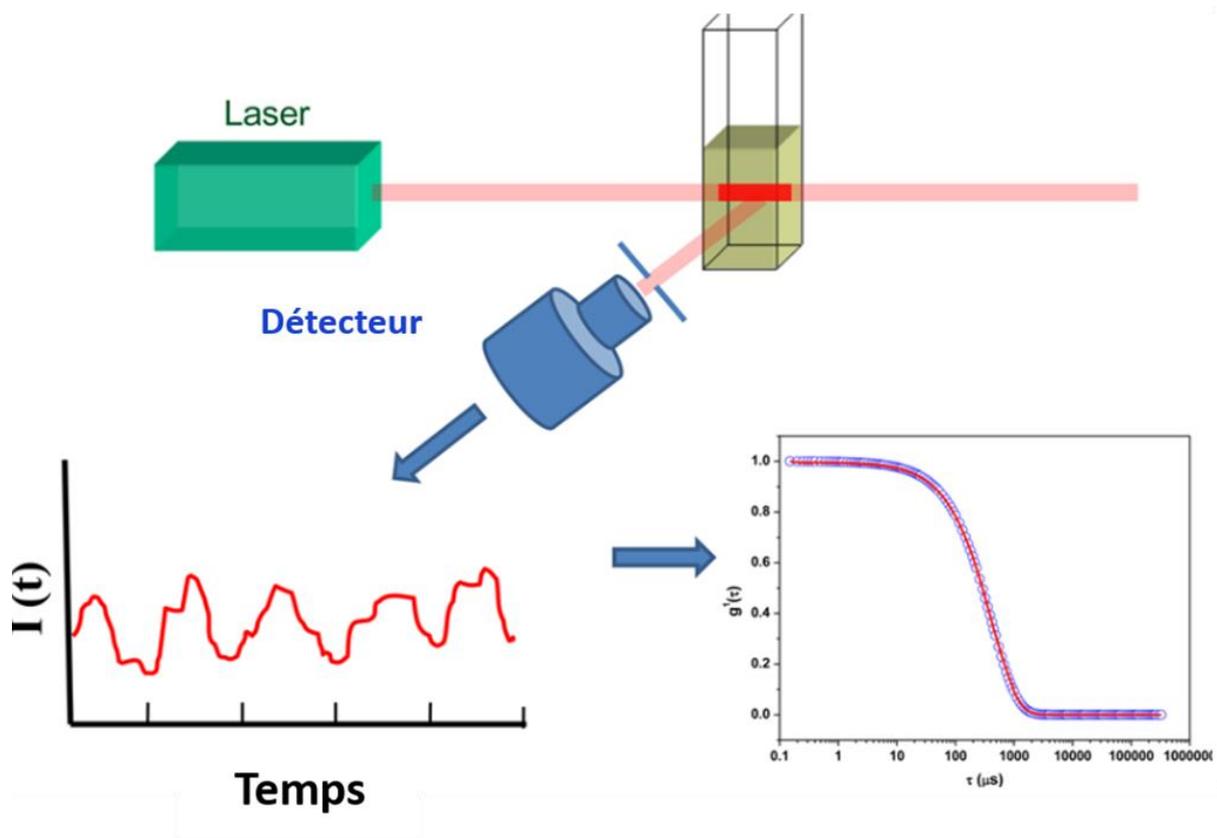
La SEC concourt à la séparation et la quantification des mélanges de protéine. Elle est couramment employée comme méthode de purification des protéines en biochimie, mais également pour le contrôle qualité de la production de protéines recombinantes. Lors d'étude d'interaction moléculaire, cette technique permet également d'analyser la formation de complexes.

## **2. Méthodes de caractérisation optiques**

### **2.1 Diffusion dynamique de la lumière**

La diffusion dynamique de la lumière (DLS, signifiant Dynamic Light Scattering) est une technique permettant d'étudier le comportement de diffusion de particules en solution et ainsi de déterminer leur taille. Elle est basée sur la théorie du mouvement Brownien des particules (Einstein 1905). Selon cette théorie, plus une particule est grande, plus elle se déplace lentement dans un liquide de viscosité et de température données. Les fluctuations de la lumière diffusée par les particules dépendent de la vitesse de diffusion de ces particules, qui est directement liée à leur distribution de taille.

Lors d'une mesure de DLS, un laser (dans notre cas, d'une longueur d'onde égale à 633 nm) éclaire les particules en solution, et la lumière diffusée est mesurée à un angle fixe (Figure 3.3).



**Figure 3.3 : Principe de la diffusion dynamique de la lumière.** Un faisceau laser éclaire l'échantillon et la lumière diffusée est captée par un détecteur. A partir de la fonction d'autocorrélation, le rayon hydrodynamique des particules est déterminé. Adaptée de (Hassan, Rana, et Verma 2015).

L'intensité de diffusion des particules s'écrit:

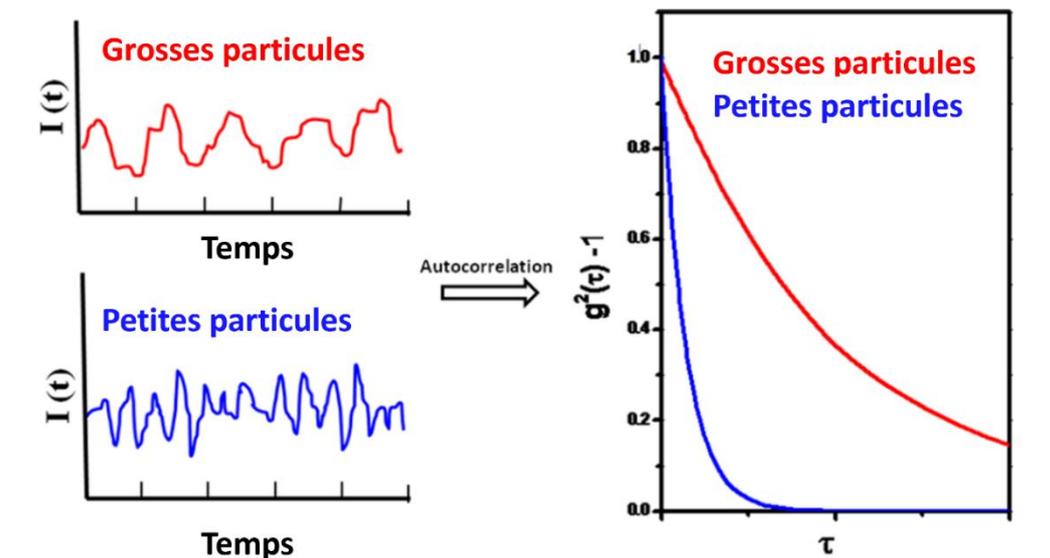
$$I = I_0 \frac{16\pi^4}{\lambda^4} \frac{1 + \cos^2\theta}{2s^2} \left( \frac{n^2 - 1}{n^2 + 2} \right)^2 r^6 \quad (1.1)$$

où  $\lambda$  correspond à la longueur d'onde du faisceau laser,  $\theta$  à l'angle de diffusion,  $s$  à la distance entre la particule diffusant et le détecteur,  $n$  au rapport de l'indice de réfraction de la particule sur l'indice de réfraction du milieu et  $r$  au rayon de la particule.

L'intensité mesurée est directement proportionnelle à la sixième puissance du rayon de la particule. Par conséquent, plus une particule est grande, plus elle contribue au signal de diffusion. De ce fait, la DLS présente une sensibilité accrue pour détecter les oligomères et les agrégats. Cette intensité de diffusion est autocorrélée en fonction d'un intervalle de temps défini nommé  $\tau$ . La fonction d'autocorrélation (ACF) sert à déterminer la vitesse à laquelle l'intensité fluctue et s'écrit :

$$g(\tau) = \frac{\langle I(t)I(t + \tau) \rangle}{\langle I(t)^2 \rangle} \quad (1.2)$$

où  $t$  correspond à un temps donné au moment de la mesure et  $\tau$  au temps de corrélation.



**Figure 3.4 : Fonction d'autocorrélation.** Schéma représentant l'intensité de la lumière diffusée et la fonction de corrélation associées pour des solutions contenant des particules de tailles différentes. Adaptée de (Hassan, Rana, et Verma 2015).

Lorsque de grandes particules sont mesurées, leur mouvement est lent, ce qui se traduit par une fluctuation lente de leur intensité. L'ACF décroît donc lentement au cours du temps. A l'inverse, lorsque des particules de petite taille sont mesurées, leur rapide mobilité entraîne une fluctuation rapide de leur intensité. L'ACF décroît donc très rapidement au cours du temps (Figure 3.4).

L'ACF peut être approximée par la fonction suivante :

$$g(\tau) = b_{\infty} + b_0 e^{-2\Gamma\tau} \quad (1.3)$$

où  $b_{\infty}$  correspond à la valeur de l'ACF à un temps infini,  $b_0$  à la valeur maximale de l'ACF à un temps proche de zéro, et  $\Gamma$  correspond au taux de décroissance.

Le taux de décroissance est proportionnel au coefficient de diffusion et peut s'écrire :

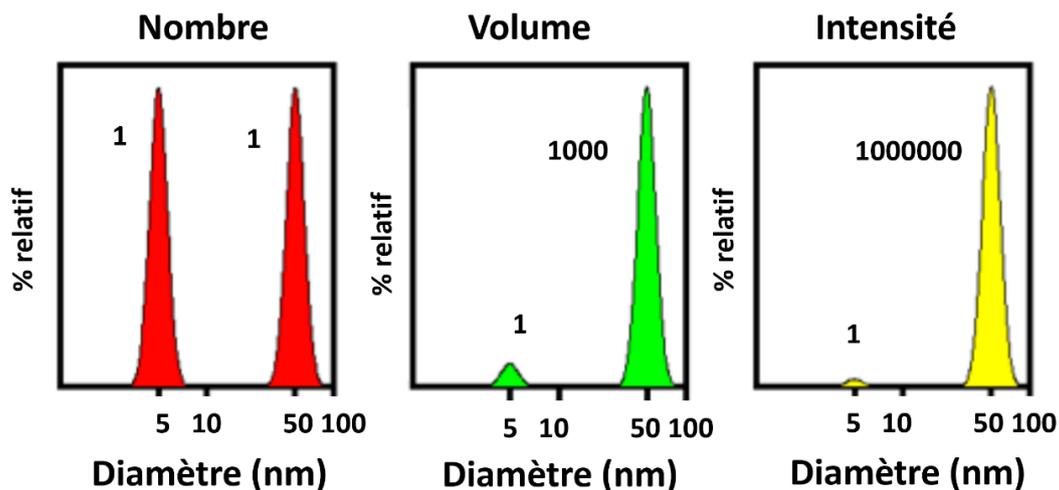
$$\Gamma = Dq^2 = D \left( \frac{4\pi n_0 \sin \frac{\theta}{2}}{\lambda} \right)^2 \quad (1.4)$$

où  $q$  correspond au nombre d'onde,  $n_0$  correspond à l'indice de réfraction du solvant,  $\theta$  à l'angle de diffusion et  $\lambda$  à la longueur d'onde du laser. De plus, le coefficient de diffusion  $D$  des particules, donné par la relation de Stokes-Einstein est inversement proportionnel au diamètre hydrodynamique des particules et dépend également de la viscosité du liquide et de la température. Ce coefficient s'écrit :

$$D = \frac{k_B T}{6\pi\eta r_h} \quad (1.5)$$

où  $k_B$  correspond à la constante de Boltzmann,  $T$  à la température,  $\eta$  à la viscosité du milieu et  $r_h$  au rayon hydrodynamique des particules.

Dans le cas de particules polydisperses, plusieurs algorithmes ont été développés, tels que l'analyse des cumulants (Frisken 2001), qui fournit l'indice de polydispersité (PDI) traduisant le degré d'homogénéité des rayons hydrodynamiques et donc des particules en solution. L'algorithme CONTIN (Provencher 1982) estime quant à lui la distribution des tailles en procédant à une transformée de Laplace régularisée de la fonction d'autocorrélation. Ces analyses sont couramment disponibles dans les logiciels de traitement des mesures de DLS (Falke et Betzel 2019; Stetefeld, McKenna, et Patel 2016).



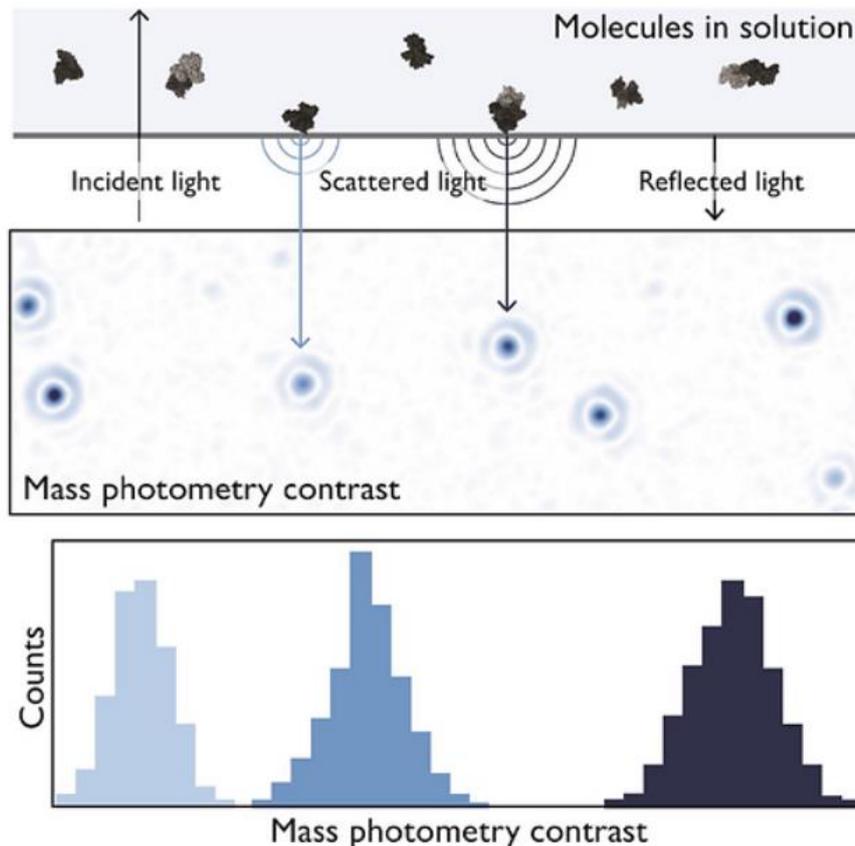
**Figure 3.5 : Distribution en nombre, volume et intensité.** Représentation de la distribution d'un échantillon contenant un mélange de particules de 5 nm et de 50 nm de diamètre. Adaptée du site chem.uci.edu ([Dynamic light scattering in 30 minutes.pdf \(uci.edu\)](http://chem.uci.edu/Dynamic%20light%20scattering%20in%2030%20minutes.pdf)).

Bien que la distribution de taille donnée par la DLS soit liée à l'intensité (distribution en intensité), elle peut aussi être convertie (Figure 3.5) en distribution selon le volume des particules (distribution en volume), puis en distribution selon le nombre de particules (distribution en nombres). Ainsi, lors de la mesure de deux particules de 5 nm et 50 nm, le pic correspondant à la particule de 50 nm sera 1 million de fois plus important que celui de la particule de 5 nm dans la distribution en intensité (car l'intensité est proportionnelle à la sixième puissance du rayon). Par ailleurs, il sera mille fois plus grand dans la distribution en volume, mais équivalent dans la distribution en nombre, car il y a autant de particule de 5 nm et de 50 nm.

Les appareils de DLS sont capables de mesurer des particules dont la taille varie entre 0,3 nm et 10 µm et dont la masse moléculaire est supérieure à 1 KDa. La mesure nécessite très peu de matériel, de l'ordre du microlitre, et la connaissance de la viscosité du tampon utilisé.

## 2.2 Photométrie de masse

La photométrie de masse est une nouvelle technique de molécule unique permettant de détecter des molécules et de quantifier leur masse moléculaire (Figure 3.6). Cette technique est basée sur les principes de microscopie à diffusion interférométrique.



**Figure 3.6 : Principe de la photométrie de masse.** La lumière diffusée par une molécule qui se pose sur une surface de mesure, interfère avec la lumière réfléchiée par cette surface. Le signal d'interférence est proportionnel à la masse de la molécule. Adaptée du site Refeyn ([Mass Photometry Technology | Refeyn](#)).

Lors d'une expérience de photométrie de masse, les molécules en solution entrent en contact avec une lame de verre exposée à un faisceau de lumière. Chacune des molécules va produire un contraste interférométrique mesurable et directement proportionnel à sa masse moléculaire. Plus la masse de la molécule est grande, plus le signal de diffusion est intense. La détection optique est donc basée sur la diffusion de la lumière et ne requiert pas de marquage ou de modification des protéines. La photométrie de masse mesure l'interférence entre la lumière diffusée par la molécule et la lumière réfléchiée par la surface de verre. Le signal généré par chaque particule

est quantifié et sa valeur de contraste, appelée le contraste interférométrique, est calculée. Les valeurs de contraste interférométriques sont corrélées à la masse moléculaire de l'objet mesuré (Wu et Piszczek 2021a; Young et al. 2018; Wu et Piszczek 2021b).

Cette technique nécessite environ 10  $\mu\text{L}$  de protéines à une concentration de l'ordre du nanomolaire. Plusieurs milliers de molécules ayant des masses moléculaires comprises entre 30 kDa à 5 MDa peuvent être détectées au cours d'une minute d'acquisition de données. En plus de mesurer la masse moléculaire des objets, cette technique peut servir à déterminer la stœchiométrie de complexes, mais également l'homogénéité et les populations d'espèces présentes dans la solution. De plus, il est possible de suivre des assemblages protéiques par photométrie de masse.

### **3. Diffusion des rayons X aux petits angles et la diffusion des rayons X aux petits angles résolue dans le temps**

La diffusion des rayons X aux petits angles (SAXS), est une méthode d'analyse structurale des macromolécules en solution. Cette approche permet de collecter des données structurales sur des protéines en solution avec une résolution spatiale de l'ordre du nanomètre. De plus, elle offre la possibilité d'étudier les changements de conformation majeurs de protéines en fonction des conditions telles que la température, le pH, la présence de ligands ou petites molécules.

Cette technique n'est pas limitée par la masse moléculaire des objets et ne nécessite ni cristaux, ni préparation spéciale des protéines. De plus, grâce à la rapidité d'acquisition des données, le SAXS s'avère être une technique de caractérisation particulièrement pratique pour l'étude des protéines. Cependant, elle est à combiner avec d'autres approches structurales, dynamiques et biochimiques comme la microscopie électronique qui sera abordée dans la partie 4 de ce chapitre.

Cette section se penchera sur les principes physiques du SAXS (Feigin, Svergun, et Taylor 1987; Svergun et Koch 2003; Daillant et Gibaud 2009), l'instrumentation nécessaire à sa mise en œuvre, ainsi que les méthodes employées pour l'analyse des données collectées. De plus, elle abordera les spécificités de la diffusion des rayons X aux petits angles résolue dans le temps (TR-SAXS).

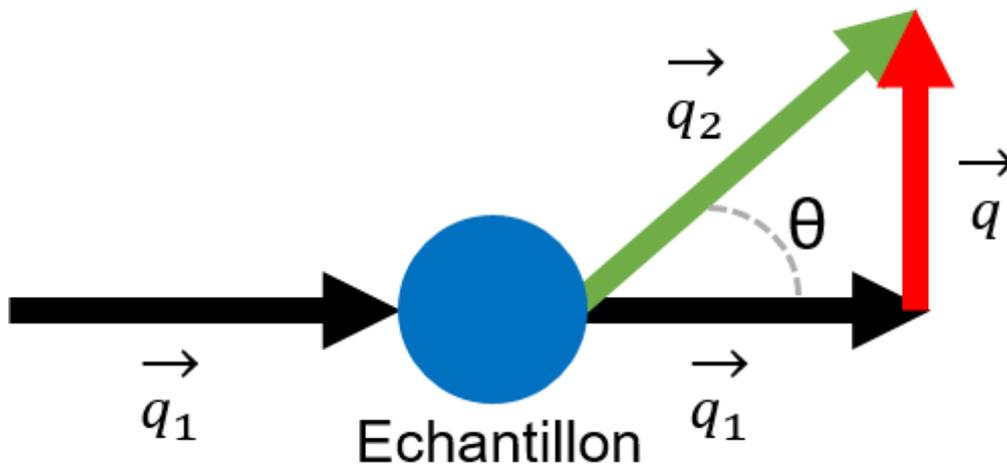
## 3.1 Principe de la diffusion des rayons X aux petits angles

### 3.1.1 Interaction des rayons X avec l'échantillon

En 1895, Wilhelm Conrad Röntgen fit une découverte révolutionnaire en mettant en évidence les rayons X. Cette découverte lui valut la distinction d'obtenir le tout premier prix Nobel de physique en 1901. Puis, en 1912, Max Von Laue établit la nature ondulatoire des rayons X, en démontrant que les rayons X sont bien des ondes électromagnétiques.

Il est important de noter que la longueur d'onde des photons de rayons X, qui peut être déterminée par la relation de Planck-Einstein, est comprise entre 0,1 nm et 10 nm. Elle est donc du même ordre de grandeur que les distances interatomiques. En conséquence, les rayons X se révèlent être un outil de premier choix pour les études structurales.

Lorsqu'un faisceau de rayons X interagit avec l'échantillon, les rayons X incidents sont diffusés dans toutes les directions de manière élastique, c'est-à-dire sans perte d'énergie. Les motifs de diffusion sont alors définis comme étant des fonctions du vecteur de diffusion noté  $\vec{q}$  (Figure 3.7).



**Figure 3.7 : Vecteur d'onde.**  $\vec{q}_1$  correspond au vecteur d'onde de l'onde incidente,  $\vec{q}_2$  correspond au vecteur d'onde de l'onde émise, et  $\vec{q}$  correspond au vecteur de diffusion. L'angle de diffusion est schématisé par la lettre  $\theta$ .  $\vec{q} = \vec{q}_2 - \vec{q}_1$

Ainsi le module du vecteur de diffusion  $q$ , s'écrit :

$$q = \frac{4\pi \sin \frac{\theta}{2}}{\lambda} \quad (1.6)$$

où  $\lambda$  correspond à la longueur d'onde et  $\theta$  à l'angle de diffusion.

### 3.1.2 Intensité de diffusion

L'amplitude de diffusion d'une molécule,  $A(\vec{q})$  est la transformée de Fourier de la distribution du contraste de densité de longueur de diffusion  $\Delta\rho(\vec{r})$  de cette molécule.

L'amplitude de diffusion s'écrit :

$$A(\vec{q}) = F[\Delta\rho(\vec{r})] = \int_{V_p} \Delta\rho(\vec{r}) \exp(i\vec{q}\vec{r}) d\vec{r} \quad (1.7)$$

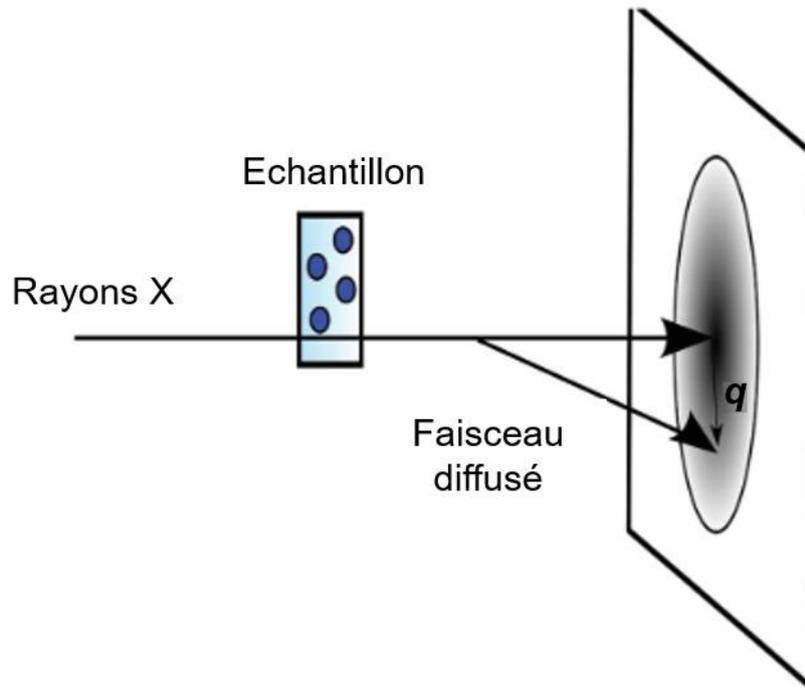
L'intensité de diffusion mesurée contient des informations dans l'espace réciproque sur la structure de cette molécule en solution. Elle correspond au carré de l'amplitude de diffusion, et s'écrit :

$$I(\vec{q}) = |A(\vec{q})|^2 \quad (1.8)$$

$$I(\vec{q}) = \iint \Delta\rho(\vec{r}) \Delta\rho(\vec{r}') \exp(i\vec{q}(\vec{r} - \vec{r}')) d\vec{r} d\vec{r}' \quad (1.9)$$

Dans une solution de N particules monodisperses, l'échantillon est dit isotrope. En effet, les particules présentes en solution ont toutes des orientations aléatoires. Ainsi, seule la moyenne sur toutes les orientations de l'intensité diffusée peut être accessible (Figure 3.8). Par la suite, nous écrivons donc :

$$I(\vec{q}) = I(q) \quad (1.10)$$



**Figure 3.8 : Motif de diffusion.** Motif de diffusion d'un échantillon traversé par le faisceau de rayons X. Adaptée de (Tuukkanen, Spilotros, et Svergun 2017).

L'intensité de diffusion s'écrit donc :

$$I(q) = \frac{N}{V} \overline{\Delta\rho^2} V_p^2 P(q) \quad (1.11)$$

où N représente le nombre de particules, V le volume de la solution,  $V_p$  le volume de la particule,  $\Delta\rho$  le contraste et,  $P(q)$  le facteur de forme.

Le facteur de forme contient toutes les informations sur la morphologie de la particule. Il s'écrit :

$$P(q) = \left\langle \frac{1}{V_p^2} \iint \Delta\rho(\vec{r}) \Delta\rho(\vec{r}') \exp(i\vec{q}(\vec{r} - \vec{r}')) d\vec{r} d\vec{r}' \right\rangle \quad (1.12)$$

Cependant quand la concentration en particules augmente, des interactions entre elles peuvent être présentes. Dans ce cas, l'intensité de diffusion s'écrit :

$$I(q) = \frac{N}{V} \overline{\Delta\rho^2} V_p^2 P(q) S(q) \quad (1.13)$$

où  $S(q)$ , correspond au facteur de structure. Il décrit l'arrangement dans l'espace de l'ensemble des particules. Pour calculer le facteur de structure, il suffit de diviser l'intensité mesurée par le facteur de forme. Dans le cas d'une solution idéale, où il n'y a pas d'interaction intermoléculaire, le facteur de structure est égal à 1.

Il convient de souligner que les faibles valeurs de  $q$  dans l'espace réciproque sont associées à de grandes distances  $r$  au sein de la particule dans l'espace réel, tandis que les grandes valeurs de  $q$ , correspondent à de courtes distances  $r$  au sein de la particule.

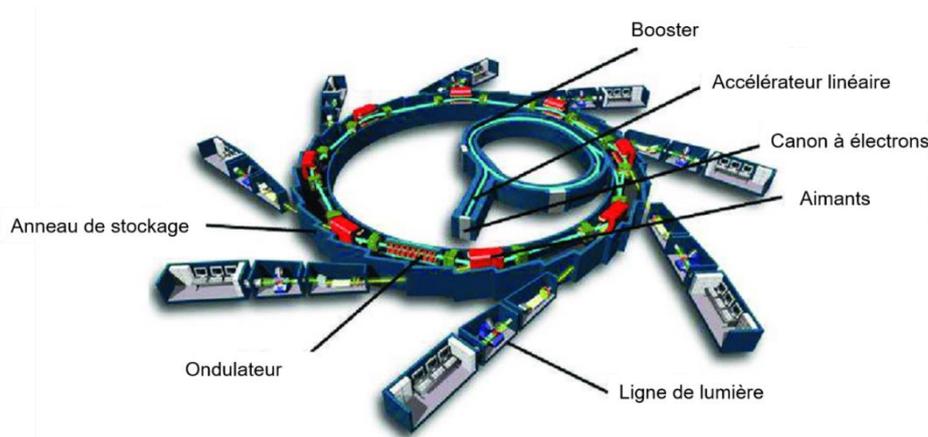
### 3.2 Instrumentation

Les expériences de SAXS et TR-SAXS ont été effectuées dans deux synchrotrons : Sur la ligne de lumière SWING à SOLEIL (Source Optimisée de Lumière d'Énergie Intermédiaire du LURE) et sur la ligne ID02 à l'ESRF (European Synchrotron Radiation Facility). Ces lignes de lumière sont utilisées pour répondre aux questions liées aux macromolécules biologiques en solution. Dans cette section, leurs caractéristiques ainsi que les différents équipements utilisés lors d'une expérience de SAXS seront brièvement présentés.

#### Le synchrotron : La source d'électron

Le synchrotron (Figure 3.9) est un accélérateur de particule composé de deux anneaux et plusieurs lignes de lumière. Un dispositif de type canon à électrons injecte des électrons dans un premier anneau appelé « booster ». A l'intérieur du booster, ces électrons sont propulsés à des vitesses proches de celle de la lumière. Ensuite, ils sont injectés dans un second anneau appelé « anneau de stockage », où ils continuent leur trajet. L'anneau de stockage est conçu avec une alternance de sections droites et

de virages. A chaque virage, grâce à un champ magnétique, les électrons perdent un peu d'énergie sous forme de rayonnement.



**Figure 3.9 : Schéma d'un synchrotron.** Adaptée de (Narayanan Nair 2013).

Pour l'expérience de SAXS, le rayonnement résultant, capté sur la ligne de lumière est filtré à l'aide de dispositifs optiques tels que le monochromateur et les fentes. Cette étape de filtrage permet un contrôle précis sur des paramètres tels que la longueur d'onde, la taille et l'intensité du faisceau.

	ID02 (ESRF)	SWING (SOLEIL)
Détecteur	Eiger2 4M	EigerX 4M
Gamme d'énergie du faisceau (en keV)	8 – 20	5 – 16
Résolution (en eV)	Environ 2,5	Environ 2
Distance de l'échantillon au détecteur	0,8 m à 31 m	0,5 m à 6,5 m
Flux de photons sur l'échantillon	$10^{13}$ ph/s à 12,5 keV	$10^{12}$ ph/s à 16 keV

**Tableau 3.1 : Récapitulatif des caractéristiques des lignes de lumière ID02 et SWING.**

### L'environnement échantillon

Il se compose de plusieurs éléments nécessaires à la mesure de l'échantillon. Il comprend une cellule de mesure en quartz, un échantillonneur automatique qui permet de mesurer plusieurs échantillons au cours d'une seule expérience, ainsi qu'un robot d'injection qui prélève l'échantillon de l'échantillonneur automatique et l'injecte dans la cellule de mesure. Tout ce système fonctionne sous-vide afin d'éviter l'absorption et la

diffusion par l'air. De plus, la présence d'un bain thermostaté permet de maintenir la température de l'environnement échantillon constante tout au long de l'expérience.

#### Le beam stop

Son rôle est d'empêcher que l'intensité du faisceau direct ne soit dirigée vers le détecteur.

#### Le détecteur

Il est utilisé pour mesurer l'intensité des rayons X diffusés par l'échantillon. La distance entre l'échantillon et le détecteur est variable et s'ajuste en fonction de la taille des objets. A mesure que cette distance augmente, on se dirige vers des plus petites valeurs de  $q$ , ce qui correspond à des angles de diffusion plus petits.

### **3.3 Analyse des courbes SAXS**

Pour obtenir une mesure correcte d'un échantillon protéique, il faut :

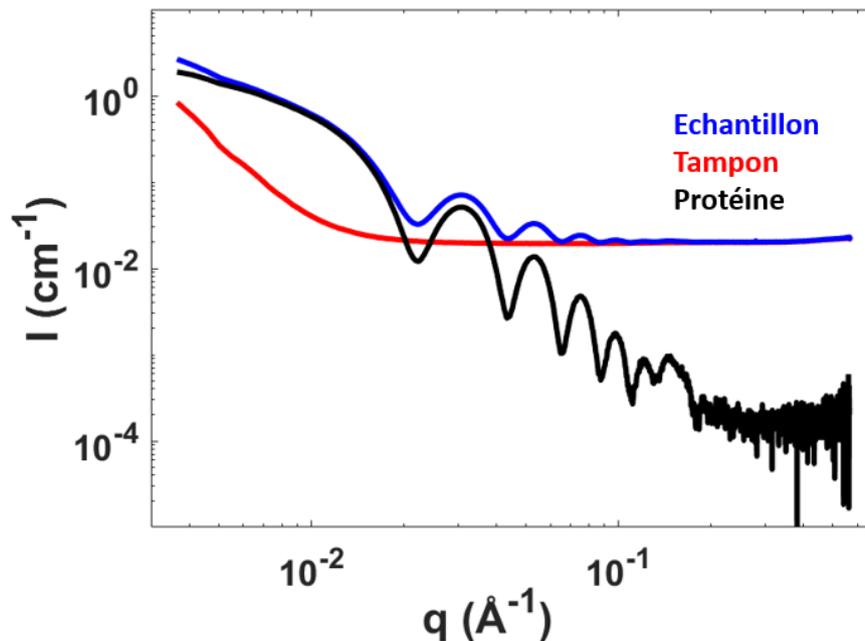
- Un échantillon le moins polydisperse possible, car le SAXS ne peut pas distinguer les différents objets présents en solution. La polydispersité de l'échantillon doit être contrôlée en amont de la mesure.
- Un volume d'échantillons de l'ordre du microlitre minimum, avec une concentration en protéines comprise entre 0.1 mg/ml et 10 mg/ml.
- L'absence d'interactions intermoléculaires. Elle peut être obtenue en modifiant la composition du tampon, en ajustant la concentration saline par exemple.
- Une bonne correspondance entre le tampon de référence et le tampon de l'échantillon. Il est donc préférable de dialyser les échantillons contre le tampon de référence avant la mesure.

#### **3.3.1 Premier traitement**

Lors d'une mesure SAXS, l'intensité mesurée correspond à celle de l'échantillon. Ce dernier est composé des protéines et du tampon. Afin d'avoir accès à l'intensité diffusée par les protéines, il est nécessaire de soustraire le tampon. Cela implique également la mesure de l'intensité de diffusion du tampon seul, en utilisant la même cellule de mesure. La soustraction permet aussi de réduire la contribution du bruit parasite provenant des éléments du montage expérimental, tels que les fentes et le

porte-échantillon. Ainsi, l'intensité diffusée exclusivement par les protéines s'écrit (Figure 3.10) :

$$I_{\text{protéine}}(q) = I_{\text{échantillon}}(q) - I_{\text{tampon}}(q) \quad (1.14)$$



**Figure 3.10 : Soustraction du tampon.** La courbe bleue représente la courbe de diffusion de l'échantillon, et la courbe rouge, celle du tampon. La courbe noire représente l'intensité de la diffusion provenant uniquement de la protéine, après la soustraction du tampon.

Enfin, il est essentiel d'examiner les aspects suivants lors de l'analyse :

- Les effets de l'irradiation : les rayons X peuvent provoquer la formation de radicaux libres qui déstabilisent la protéine. Cela se traduit par une augmentation de la valeur de l'intensité diffusée aux petits angles. Il est donc important de mesurer l'échantillon avec différents temps d'exposition afin de se rendre compte des effets de l'irradiation et d'ainsi choisir un temps d'exposition qui n'endommagera pas la protéine mais qui maximisera le signal mesuré.
- Les interactions moléculaires : Les courbes de diffusion de l'échantillon mesuré doivent être superposables. Si l'intensité aux petits angles diminue alors que la concentration en protéines augmente, cela indique la présence de répulsions

intermoléculaires. En revanche, si l'intensité aux petits angles augmente, cela suggère la présence d'attractions intermoléculaires.

- La présence d'une éventuelle agrégation ou oligomérisation : Une remontée de l'intensité aux petits angles indique la présence d'oligomères voire d'agrégats dans l'échantillon.

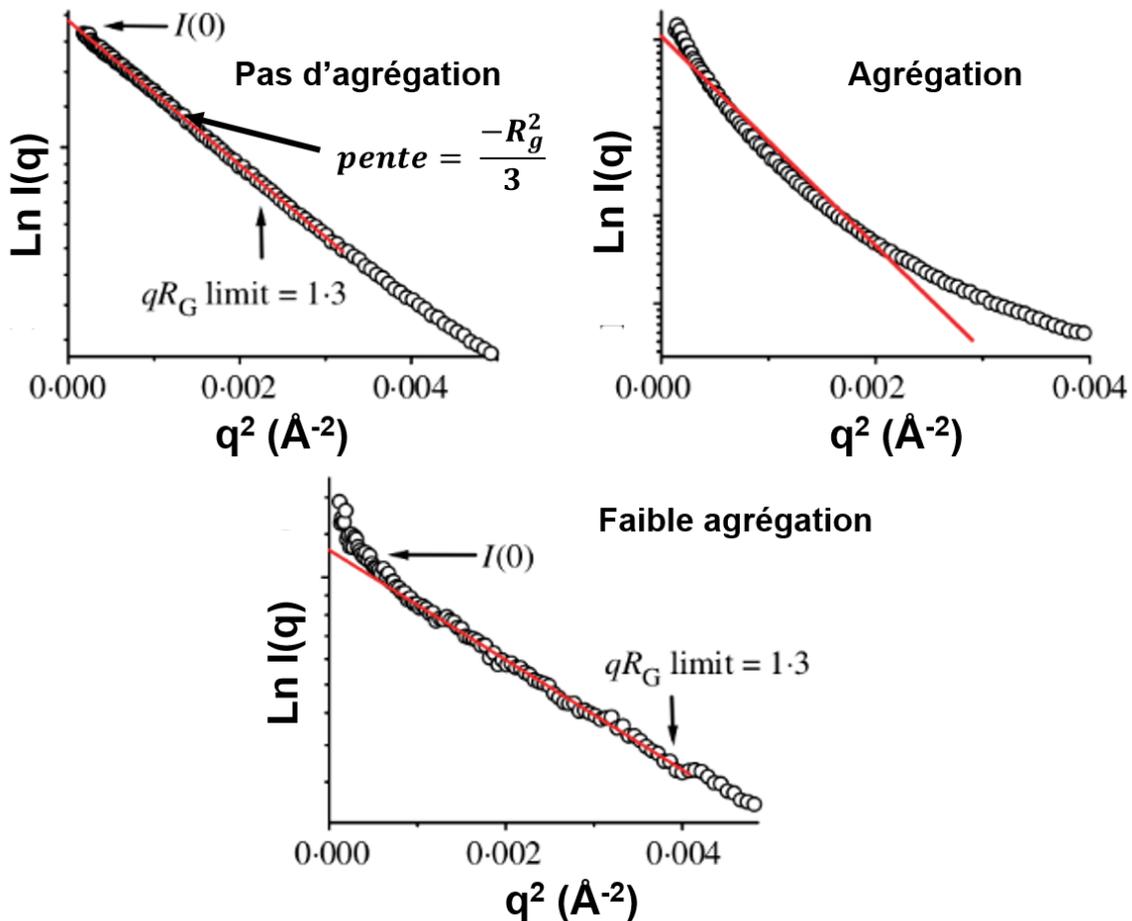
### 3.3.2 Deuxième traitement

#### Approximation de Guinier

L'approximation de Guinier (Figure 3.11) est une méthode utilisée pour déduire la taille moyenne de la protéine à partir de la portion de la courbe de diffusion aux petites valeurs de  $q$  (Guinier et al. 1956). D'une façon très générale, l'intensité diffusée peut être approximée par l'équation suivante :

$$I(q) \approx I(0)e^{\frac{-q^2 R_g^2}{3}} \quad (1.15)$$

où  $I(0)$  correspond à l'intensité à  $q = 0$  et  $R_g$  correspond au rayon de giration de la protéine. Le carré du rayon de giration représente la moyenne quadratique des distances atomiques depuis le barycentre pondéré par la densité électronique. Il renseigne uniquement sur la taille de la protéine.



**Figure 3.11 : Représentation de Guinier.** Elle permet de déterminer le  $I(0)$ , le rayon de giration de la particule et la présence d'agrégat. Adaptée de (Quijada et Marileth 2010).

La représentation de Guinier consiste à tracer  $\text{Ln } I(q)$  en fonction de  $q^2$ . Cette représentation permet de déterminer à la fois le rayon de giration et le  $I(0)$  de la protéine étudiée. Cependant, il est important de noter que le diagramme de Guinier n'est linéaire que dans une région restreinte du spectre de diffusion respectant la condition suivante :  $qR_g < 1,3$ . Si les données sont linéaires dans cette région, cela indique que l'approximation de Guinier est valide et que l'estimation obtenue pour le rayon de giration est fiable. De plus, la linéarité des données dans cette région peut donner des informations sur le degré d'agrégation de la protéine, car des déviations par rapport à la linéarité pourraient indiquer des interactions ou des agrégats dans la solution.

#### La $P(r)$ : La fonction de distribution des distances

Cette fonction fournit des informations sur toutes les distances interatomiques  $r$  au sein de la protéine dans l'espace réel (Receveur-Bréchet et Durand 2012). Elle nous

renseigne sur la structure de la protéine, qu'elle soit compacte, étalée, ou si elle présente des domaines séparés (Figure 3.12). Cette fonction correspond à la transformée de Fourier inverse de l'intensité de diffusion  $I(q)$ , et s'écrit :

$$P(r) = T^{-1} [I(q)] \quad (1.16)$$

Et,

$$I(q) = 4\pi \int_0^{\infty} P(r) \frac{\sin qr}{qr} dr \quad (1.17)$$

Donc,

$$P(r) = \frac{r^2}{2\pi^2} \int_0^{\infty} I(q) \frac{\sin qr}{qr} q^2 dq \quad (1.18)$$

Il est également possible de déterminer le  $I(0)$  et le rayon de giration à partir de la  $P(r)$ .

Le  $I(0)$  s'écrit :

$$I(0) = 4\pi \int_0^{Dmax} P(r) dr \quad (1.19)$$

Le  $R_g$  s'écrit :

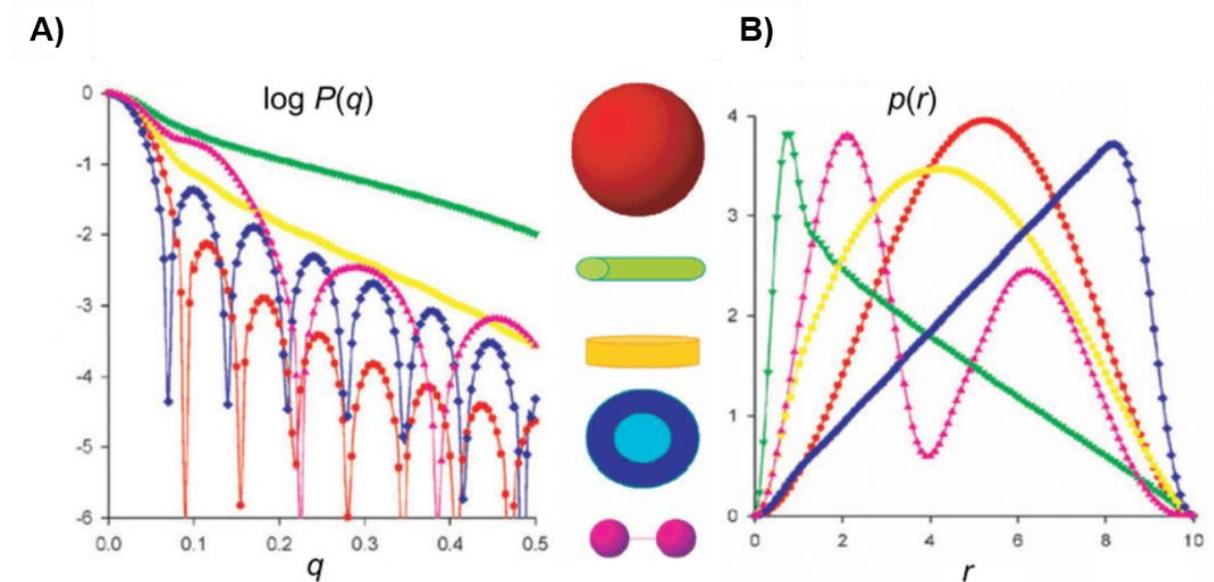
$$R_g^2 = \frac{\int_0^{Dmax} r^2 P(r) dr}{2 \int_0^{Dmax} P(r) dr} \quad (1.20)$$

où  $Dmax$  correspond au diamètre maximal qui sépare les deux électrons les plus éloignés dans la protéine étudiée. C'est le diamètre maximal de la protéine.

Ce mode de calcul du  $R_g$  utilise toute la courbe de diffusion et est donc moins sensible aux interactions ou à la présence de petits oligomères dans l'échantillon. La

comparaison entre le  $R_g$  obtenu par l'approximation de Guinier et le  $R_g$  obtenu grâce à la  $P(r)$  est très utile pour éviter les erreurs d'estimation.

La représentation de la  $P(r)$ , consiste à tracer  $\frac{P(r)}{I(0)}$  en fonction de  $r$ .

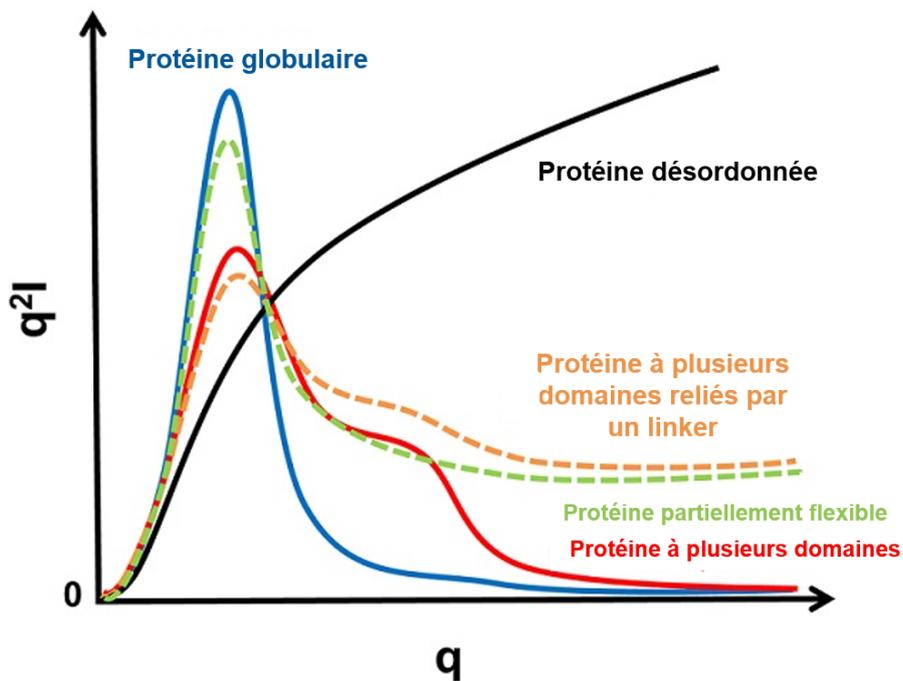


**Figure 3.12 : Fonction de distribution de distance.** A) Intensité de diffusion de plusieurs objets de forme différentes et B) Fonction de distribution de distance associée à chaque objet. Tous ces objets ont le même  $D_{\max}$  (diamètre maximal). Adaptée de (Schroer 2011).

### La représentation de Kratky

La représentation de Kratky (Receveur-Bréchet et Durand 2012) est une représentation de l'intensité de diffusion visant à évaluer le degré de compaction d'une protéine. Cette méthode permet de déterminer si la protéine est compacte, partiellement structurée ou peu structurée.

Elle consiste à représenter graphiquement  $q^2 I(q)$  en fonction de  $q$ . Elle est donc directement déterminée à partir de la courbe expérimentale, contrairement à la fonction de distribution des distances (Figure 3.13).



**Figure 3.13 : Représentation de Kratky.** Représentation de Kratky de plusieurs protéines ayant des niveaux de structuration différents. Adaptée de (« Solution Scattering Data Assessment » s. d.).

Cette représentation peut également être normalisée en traçant  $(qR_g)^2 \frac{I(q)}{I(0)}$  en fonction de  $qR_g$ . La multiplication de  $q$  par le rayon de giration rend les angles de diffusion indépendants de la taille de la protéine, et comme  $I(0)$  est proportionnel au poids moléculaire, la division de  $I(q)$  par  $I(0)$ , rend  $I(q)$  indépendante du poids moléculaire. En conséquence, la représentation de Kratky normalisée permet de mieux mettre en évidence les différences entre les protéines globulaires, celles qui sont complètement désordonnées ou celles qui possèdent des domaines plus rigides.

## **3.4 Diffusion des rayons X aux petits angles résolue dans le temps**

### **3.4.1 Utilité du TR-SAXS**

En SAXS, les protéines se trouvent en solution. Elles peuvent être facilement perturbées ou modifiées et l'effet de ces perturbations peut être mesuré lors d'expériences de SAXS résolue dans le temps (TR-SAXS).

Le TR-SAXS (Chen et al. 2023; Trewhella 2022; Cho et al. 2021) peut être réalisé lorsque des changements mesurables surviennent entre l'état initial et l'état final d'une protéine lors d'une réaction. Ces changements peuvent se manifester par des modifications de la valeur du  $I(0)$  et du rayon de giration de la protéine. Le TR-SAXS est sensible aux espèces intermédiaires présentes en solution au cours de la réaction et cette méthode permet d'obtenir des informations détaillées sur la cinétique d'un phénomène comme l'autoassemblage de capsides par exemple.

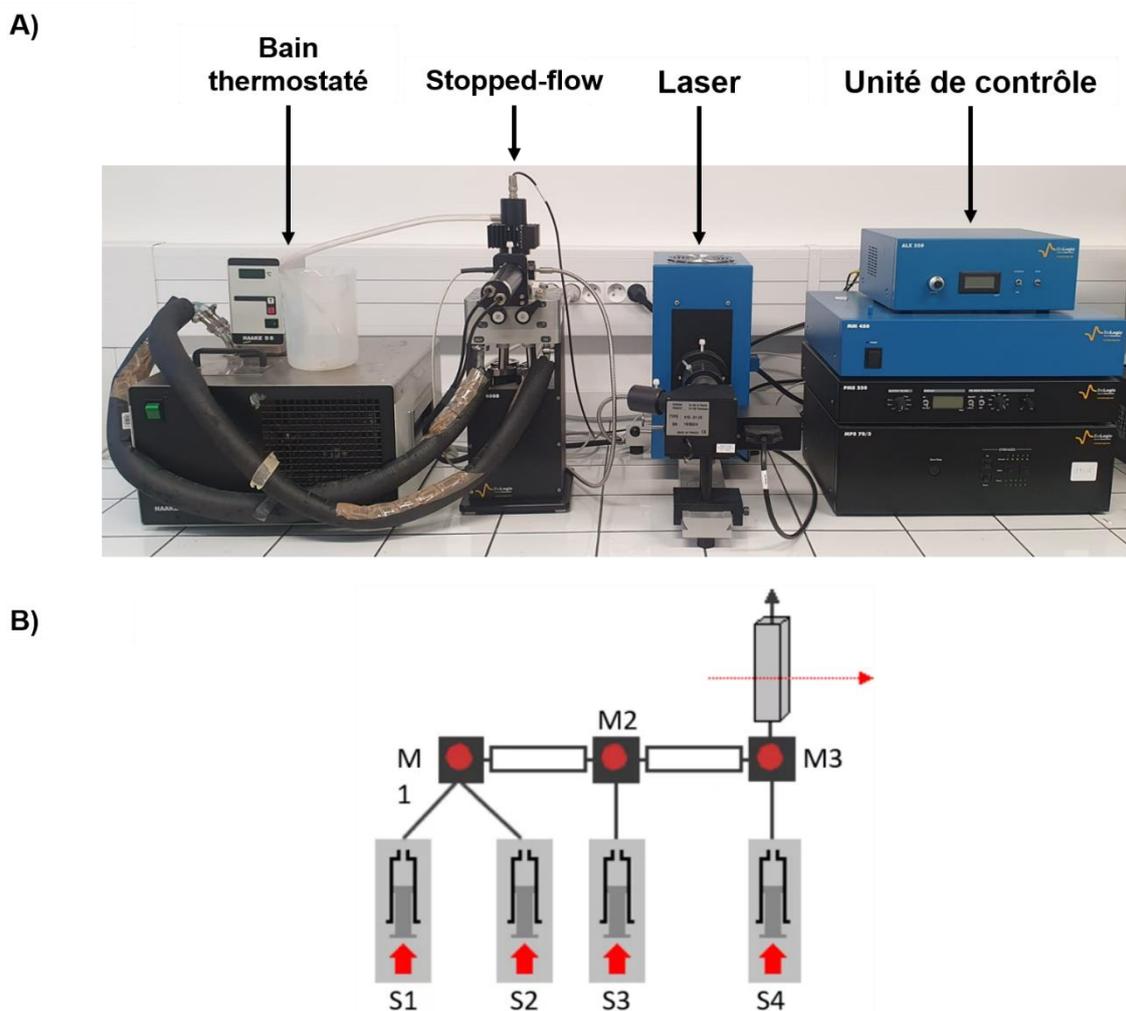
Cette technique permet d'accéder à des réactions ayant une résolution temporelle allant de la milliseconde à l'heure, ce qui signifie qu'elle peut suivre des changements structuraux des protéines au fil du temps, sur une large gamme d'échelles temporelles. Ces changements peuvent donc être très rapides, ou très lents. Cette capacité à étudier la cinétique des réactions protéiques en temps réel, tout en ayant des informations structurales fait du TR-SAXS un précieux outil pour comprendre les mécanismes sous-jacents du processus d'assemblage de la capsid du VHB en présence des modulateurs d'assemblage.

Le déclenchement de la réaction doit être initié de manière contrôlée par un facteur externe, et il est essentiel de pouvoir collecter les données à des moments spécifiques pendant le déroulement de la réaction. Dans notre étude, les données ont été collectées durant plusieurs minutes en utilisant une progression géométrique pour les temps de mesure.

### **3.4.2 Principe du stopped-flow et précautions d'usage**

La réaction suivie par TR-SAXS est déclenchée par mélange rapide à l'aide d'un stopped-flow.

Le stopped-flow (Figure 3.14) est un dispositif capable de mélanger rapidement plusieurs solutions envoyées simultanément puis d'arrêter ce mélange après un intervalle de temps défini. Il offre une résolution temporelle de l'ordre de la milliseconde. Les appareils de stopped-flow utilisés sur ID02 et sur SWING (Biologic® SFM-400) se composent de 4 réservoirs différents et sont équipés d'une cellule de mesure en quartz. Les réservoirs sont contrôlés par 4 moteurs indépendants, ce qui rend possible la modification du débit, du volume d'injection, mais également du rapport de mélange entre les différentes solutions. De plus, l'appareil peut être relié au bain thermostaté afin de contrôler la température lors de la cinétique.



**Figure 3.14 : Stopped-flow.** A) Photo du stopped-flow SFM 400 utilisé au Laboratoire de Physique des Solides, B) Schéma de fonctionnement du SFM 400.

Un inconvénient notable de cet équipement est qu'il consomme énormément de solution. Il est donc important de disposer d'une quantité importante de protéines. En règle générale, plus la réaction est rapide, plus la quantité de protéine nécessaire est

importante. Pour nos expériences, il était généralement nécessaire d'avoir entre 20 et 30 mg de protéines. En outre, lors de l'expérience, plusieurs précautions sont à prendre pour avoir une bonne utilisation de l'appareil et donc un résultat optimal :

- Il est important de dégazer les solutions avant chaque expérience et d'injecter très lentement les solutions dans les réservoirs pour éviter la présence de bulles. Ces bulles conduisent à une augmentation de l'intensité diffusée et rendent impossible l'interprétation des données.
- Des fuites peuvent survenir au niveau des réservoirs, provoquant la migration des solutions, ce qui perturbe le signal de diffusion. Il est donc important de bien choisir le réservoir pour chaque solution. Dans le cadre de nos expériences de cinétique, afin d'éviter une contamination entre nos solutions, nous avons adopté la stratégie suivante : le tampon contenant du sel a été placé dans le premier réservoir, tandis que la solution protéique a été introduite dans le réservoir numéro 4, celui-ci étant le plus proche de la cellule de mesure et étant mélangé en dernier.
- Il est également important de ne laisser aucun réservoir vide lors de la cinétique, pour éviter l'introduction de bulles.

Enfin, après chaque cinétique, il est obligatoire de nettoyer la cellule de mesure et de s'assurer qu'elle est bien propre en effectuant une mesure du tampon. En effet, des protéines peuvent rester collées sur les parois de la cellule de mesure et ainsi polluer la cinétique suivante.

### **3.4.3 Analyse des données obtenues**

Après une expérience de cinétique par TR-SAXS, plusieurs centaines de courbes de diffusion sont obtenues. Afin de traiter ces courbes, il est nécessaire d'utiliser des logiciels et programmes afin d'automatiser certaines analyses.

Le calcul du  $I(0)$  et du rayon de giration pour chaque courbe SAXS, qui nous permet de suivre l'évolution au cours du temps de l'échantillon peut être réalisé à l'aide des programmes de la suite ATSAS. La détermination de ces deux paramètres repose sur l'approximation de Guinier.

Ensuite, les valeurs de l'intensité à  $I(0)$  permettent de déterminer le nombre d'agrégation moyen pour chaque courbe lors de la cinétique. Dans le cadre d'un

assemblage de capsides du VHB, ce nombre correspond au nombre moyen de dimères présents dans chaque objet lors de la cinétique d'assemblage.

Enfin, lors d'une expérience de TR-SAXS, contrairement à nos expériences de SAXS, l'échantillon ne reste pas homogène lors de la cinétique. Selon les conditions d'assemblage, une, deux voire plusieurs espèces sont présentes. Le signal de diffusion correspond donc à une moyenne de l'intensité de diffusion provenant de ce mélange d'espèces. Pour décomposer le signal, une méthode mathématique appelée la décomposition en valeur singulière peut être utilisée. Elle permet de séparer et d'analyser la contribution au signal de chaque espèce en solution. Le nombre de valeurs singulières déterminées correspond au nombre d'espèces qui sont détectables.

## **4. Microscopie électronique et traitement d'images**

La microscopie électronique est une technique permettant d'imager des échantillons à l'aide d'un faisceau d'électrons. Elle offre une résolution plus élevée que la microscopie optique, ce qui offre une observation en détail des éléments de très petites tailles. Les électrons qui ont une longueur d'onde de l'ordre du picomètre peuvent être focalisés. Cependant, il est important de noter que l'utilisation des électrons endommage l'échantillon et que les électrons ont un faible pouvoir de pénétration. Ainsi, l'épaisseur maximum des échantillons observés en microscopie électronique doit être inférieure à 500 nm.

### **4.1 Anatomie d'un microscope électronique**

Un microscope électronique est composé de plusieurs éléments : une source d'électrons, des systèmes de lentilles, la chambre échantillon, les filtres à énergie, le détecteur et les pompes (Figure 3.15).

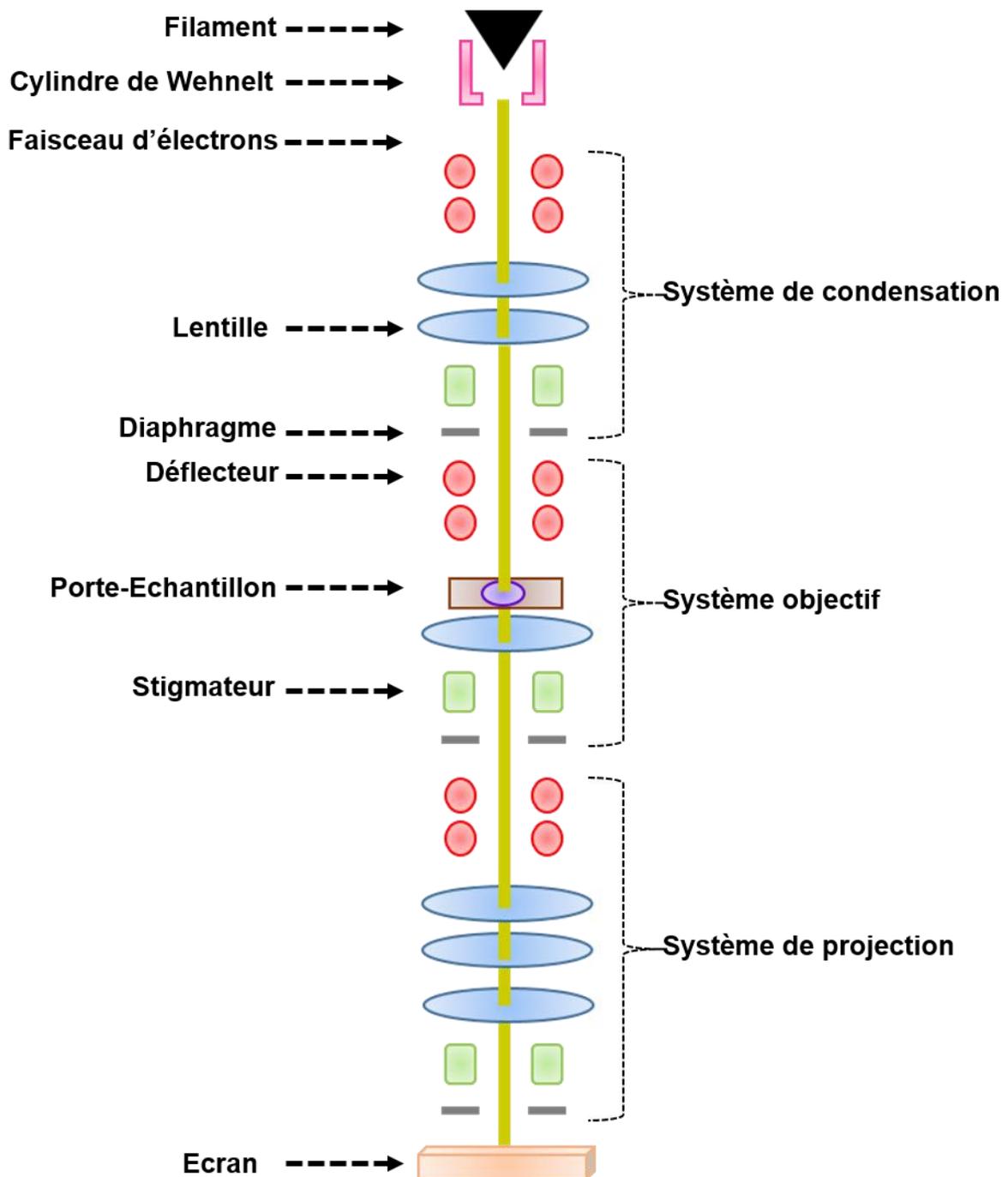


Figure 3.15 : Schéma d'un microscope.

### Le canon à électron

Le faisceau d'électrons est généré à l'aide d'un canon à électrons. Le courant est appliqué et passe à travers le filament. A la pointe de ce filament, les électrons vont s'échapper et être focalisés puis accélérés vers le bas du microscope grâce au cylindre de Wehnelt. Il est important d'obtenir un faisceau cohérent dans le temps et dans l'espace. En d'autres termes, les électrons doivent émerger de la pointe du canon

exactement à la même position et avec la même énergie. Il existe différents types de filaments : le filament de tungstène, le filament LaB<sub>6</sub>, et le canon à émission de champ. Dans le cas du canon à émission de champ, une tension est appliquée au niveau de la pointe du filament. Cette tension permet d'extraire des électrons qui sont plus cohérents. En conséquence, le faisceau devient plus homogène, plus brillant, et un plus grand nombre d'électrons sont émis par rapport aux autres types de filaments.

### **Les lentilles**

Les lentilles électromagnétiques concentrent le rayonnement afin de le focaliser. Elles permettent d'agrandir l'image. Le système de lentilles utilisé en microscopie électronique est composé des lentilles, des défecteurs, des stigmatiseurs et d'une ouverture. Le rayonnement entrant dans le microscope n'est pas obligatoirement aligné le long de l'axe optique. Pour remédier à cette situation, les défecteurs produisent un champ magnétique qui peut courber la trajectoire du faisceau d'électrons et corriger sa direction. Les stigmatiseurs quant à eux, permettent d'ajouter un champ magnétique supplémentaire pour compenser l'asymétrie du champ magnétique des lentilles. Enfin, les ouvertures, qui sont situées en aval des lentilles, bloquent les rayonnements qui s'éloignent de l'axe optique.

### **La colonne**

La colonne du microscope est composée d'un ensemble de lentilles ayant des fonctions spécifiques, ainsi que de la chambre échantillon. Le système de condensation est formé de lentilles permettant de concentrer et focaliser le faisceau sur l'échantillon. Il est situé en amont de l'échantillon. Ensuite, le système de lentilles objectifs est situé juste en dessous de l'échantillon. Enfin, le système de projecteur est composé de lentilles permettant d'agrandir l'image de l'échantillon observé. L'image est agrandie au fur et à mesure de sa traversée des lentilles.

### **Les filtres à énergie**

Ils sont situés en aval du système de projecteurs et permettent de filtrer les électrons diffusés de manière non élastique après avoir traversé l'échantillon.

### **Le détecteur**

La caméra CCD (Charge coupled device pour dispositif à transfert de charge) est un type de caméra qui repose sur le principe du transfert de charges. A l'aide d'un scintillateur qui émet des photons à la suite du dépôt des électrons diffusés, la caméra CCD convertit la lumière en signaux électriques qui peuvent être enregistrés et traités pour créer des images numériques.

### **Le système de pompe**

La colonne est placée sous un vide poussé afin d'éviter que les électrons du faisceau ne diffusent. Les pompes permettent de créer et de maintenir ce vide.

## **4.2 Contraste en microscopie électronique**

Lorsque les électrons traversent l'échantillon, ils sont diffusés. La lentille située en aval de l'échantillon va focaliser les électrons et conduire à la formation d'un spectre de diffraction au niveau du plan focal. Ce plan focal est situé à l'arrière de la lentille et le spectre de diffraction correspond à la transformée de Fourier de la densité électronique de l'échantillon. Chaque point sur le plan focal représente une onde sinusoïdale particulière nécessaire à la reproduction de l'échantillon. Ensuite, les électrons diffusés vont continuer à se propager dans le microscope et conduire à la formation de l'image. Cependant, l'image de l'objet obtenu est floue puisque certains composants de la fonction de Fourier sont manquants. Par exemple, si cette image est composée de plusieurs points, la fonction d'étalement de point (PSF pour "*point spread function*" en anglais), décrit comment chaque point est étalé sur l'image dans l'espace réel. Pour une image, il existe une seule PSF. Chaque point de l'image sera affecté par la même PSF. Ainsi, l'image obtenue en microscopie correspond au produit de convolution de l'échantillon observé par la fonction d'étalement de point.

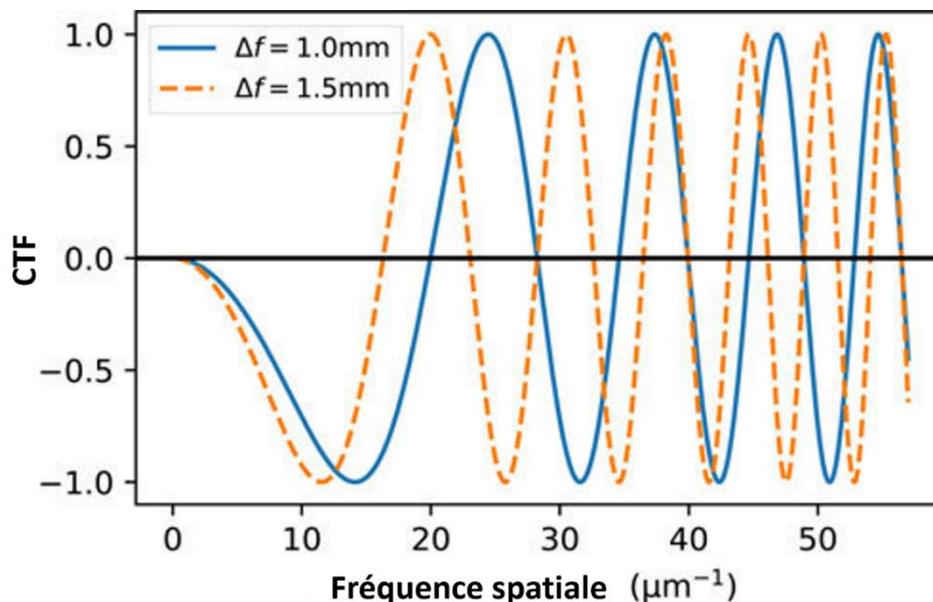
La résolution des images en microscopie électronique est limitée par le contraste. Il dépend de la nature et de l'intensité des interactions qui ont lieu entre le faisceau et le spécimen. Il existe deux types de contrastes : le contraste d'amplitude qui est lié à la perte d'électrons au passage du faisceau dans l'échantillon et les lentilles, et le contraste de phase qui résulte du décalage de phases des différentes portions du faisceau contribuant à l'image.

La fonction de transfert de contraste (CTF) caractérise la proportion et la qualité de contraste provenant à la fois d'effets d'amplitude et de phase transférés dans l'image (Figure 3.16). Elle dépend de la longueur d'onde, de la constante d'aberration de sphéricité et du défaut durant la prise de vue. Elle renseigne sur la qualité et sur la résolution maximale de l'image. Cette fonction est l'équivalent de la PSF, mais dans l'espace réciproque de Fourier. La CTF s'écrit :

$$CTF = \sin\left(-\pi\Delta z\lambda k^2 + \frac{\pi C_s\lambda^3 k^4}{2}\right) \quad (1.21)$$

Où  $\Delta z$  correspond au défaut,  $k$  correspond à la fréquence spatiale,  $C_s$  correspond à la constante d'aberration de sphéricité, et  $\lambda$  correspond à la longueur d'onde des électrons.

Ainsi, changer le défaut, modifie la CTF. Une image prise près du focus aura des détails à haute résolution tandis qu'une image prise loin du focus aura des détails à basse résolution.



**Figure 3.16 : La fonction de transfert de contraste.** CTF calculée à deux défauts différents. Adaptée de (Paterson et al. 2021).

La correction de la CTF permet d'améliorer les images obtenues au microscope électronique.

Elle s'effectue de cette manière :

$$F(\text{Image}) = F(\text{Echantillon}) \times F(\text{PS}) \quad (1.22)$$

$$TF(\text{Image}) = TF(\text{Echantillon}) \times CF \quad (1.23)$$

$$F(\text{Echantillon}) = \frac{F(\text{Image})}{CF} \quad (1.24)$$

Ainsi, si nous prenons la transformée de Fourier de notre image, que nous la divisons par la CTF puis que nous réalisons une Fourier inverse du résultat, nous obtiendrons l'image corrigée par la CTF. Cette image est moins floue et plus proche de la vraie structure de l'objet. Cette image pourra ensuite être analysée.

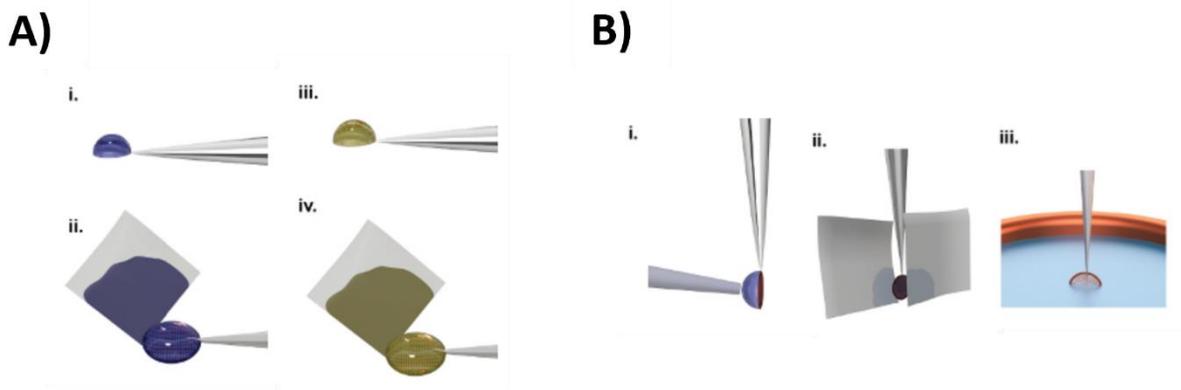
En pratique, ce n'est pas exactement ce qui est fait, car la CTF s'annule pour plusieurs fréquences. Les valeurs nulles de la CTF correspondent à une absence d'information transmise sur l'image et il n'est pas possible de récupérer ces informations perdues. Une première méthode consiste à exclure toutes les fréquences spatiales où la valeur de la CTF est proche de 0. De plus, une inversion de phase dans les zones où le contraste est négatif peut être réalisée. Cela permet de restaurer une partie de l'information dans l'image. Enfin, une correction plus sophistiquée peut être effectuée en utilisant un filtre de Wiener. Il permet une correction des phases, mais également des amplitudes de la CTF.

Il est important de noter également que prendre des images à plusieurs défocus différents permet de générer plusieurs CTF distinctes. Les informations obtenues à un niveau de défocus donné peuvent ensuite être utilisées pour compenser les informations manquantes à d'autres niveaux de défocus.

### 4.3 Préparation de l'échantillon

Afin d'être observé en microscopie électronique, l'échantillon est déposé sur une grille qui sera ensuite placée sur le porte-objet à insérer dans la chambre échantillon du microscope.

Les grilles utilisées en microscopie électronique mesurent environ 3 mm de diamètre, sont très fines et ont un maillage défini. Elles sont fabriquées à partir de matériaux conçus pour permettre aux électrons de passer à travers sans interférences. Généralement en cuivre, elles sont recouvertes d'un film de carbone et/ou de formvar. Ces grilles peuvent également être trouées et le nombre de trous, leur forme et la distance entre ces trous diffèrent selon les modèles de grille. Avant de déposer l'échantillon sur la grille, celle-ci doit être rendue hydrophile et chargée négativement afin de permettre une adsorption optimale de l'échantillon. Pour cela, un effluage est effectué.



**Figure 3.17 : Méthodes de préparation des grilles de microscopie électronique.** A) Coloration négative. B) Vitrification plongeante. Adaptée de (Schmidli et al. 2018).

Deux techniques de préparation de l'échantillon ont été utilisées lors de cette étude. La coloration négative pour l'observation des capsides en microscopie électronique à transmission, et la vitrification plongeante pour l'observation des capsides en cryomicroscopie électronique :

- Coloration négative (Figure 3.17.A) : Après avoir ionisé la grille pour la rendre hydrophile, une goutte de solution contenant l'échantillon est déposée sur la grille (Figure 3.17.A.i). S'ensuit le retrait d'excédent de solution (Figure 3.17.A.ii) et le dépôt du contrastant (Figure 3.17.A.iii). L'excès de contrastant est également retiré (Figure 3.17.A.iv). Ce dernier va se placer sur les bords des particules à observer et créer des zones denses en électrons qui diffusent

fortement. Il augmente ainsi le contraste. Cette technique déshydrate l'échantillon et risque d'entraîner des modifications des structures de l'échantillon observé.

- Vitrification plongeante (Figure 3.17.B) : La vitrification plongeante ou "plunge-freezing" est une technique de congélation utilisée en cryomicroscopie électronique. Elle permet de congeler l'échantillon dans un état hydraté. L'échantillon est déposé sur une grille hydrophile préalablement placée dans un dispositif de congélation automatique (Figure 3.17.B.i). La grille est ensuite épongée entre deux papiers filtres qui absorbent l'excès de matière (Figure 3.17.B.ii) puis elle est plongée très rapidement dans l'éthane liquide (Figure 3.17.B.iii). L'éthane liquide est refroidi à environ 90 K par de l'azote liquide refroidi à 77 K. La congélation très rapide à cette température permet de geler les réarrangements moléculaires. L'échantillon est donc figé et sa structure préservée. Une fois congelé, l'échantillon doit rester dans cet état.

#### **4.4 Traitement de l'image : analyse de particules uniques**

Afin de reconstruire les capsides du VHB en présence de modulateurs, des analyses de particules uniques ont été réalisées à partir des images prises par cryomicroscopie électronique.

L'analyse de particules uniques permet en biologie structurale de résoudre la structure tridimensionnelle de particules à partir d'images en 2 dimensions de ces particules.

Elle peut être décomposée en plusieurs étapes décrites ci-dessous :

- La première étape consiste à collecter des données. Pour cela, des images de particules sont acquises à l'aide du microscope.
- Puis, les images sont ensuite triées afin d'éliminer les images où la glace n'est pas vitreuse. Celles-ci sont ensuite traitées afin de corriger les mouvements apparents sur l'image qui sont dus au faisceau et la CTF est estimée et corrigée afin d'obtenir à la fin de l'analyse, une reconstruction 3D avec la meilleure résolution possible.
- Ensuite, les particules sont identifiées, alignées entre elles et classées. Chaque classe contient des particules homogènes.
- Après avoir sélectionné les bonnes classes 2D, les particules appartenant à ces classes sont extraites des micrographies.
- Puis, un premier modèle 3D est reconstruit à partir des particules isolées.

- La dernière étape consiste à affiner le modèle. Des reprojctions dans toutes les directions à partir du modèle sont ensuite effectuées. Elles permettent d'affiner le modèle. A chaque itération lors de l'affinement, la résolution du modèle augmente jusqu'à atteindre la résolution maximale.

## Bibliographie

Chen, Ruixin, Yang Song, Zhichun Wang, Hang Ji, Zhongyao Du, Qingwen Ma, Ying Yang, Xingxun Liu, Na Li, et Yang Sun. 2023. « Developments in Small-Angle X-Ray Scattering (SAXS) for Characterizing the Structure of Surfactant-Macromolecule Interactions and Their Complex ». *International Journal of Biological Macromolecules* 251 (novembre): 126288. <https://doi.org/10.1016/j.ijbiomac.2023.126288>.

Cho, Hyun Sun, Friedrich Schotte, Valentyn Stadnytskyi, et Philip Anfinrud. 2021. « Time-Resolved X-Ray Scattering Studies of Proteins ». *Current Opinion in Structural Biology* 70 (octobre): 99-107. <https://doi.org/10.1016/j.sbi.2021.05.002>.

Daillant, Jean, et Alain Gibaud, éd. 2009. *X-ray and Neutron Reflectivity*. Vol. 770. Lecture Notes in Physics. Berlin, Heidelberg: Springer. <https://doi.org/10.1007/978-3-540-88588-7>.

Einstein, Albert. 1905. « -INVESTIGATIONS O N THE THEORY .OF ,THE BROWNIAN MOVEMENT ».

Falke, Sven, et Christian Betzel. 2019. « Dynamic Light Scattering (DLS) ». *Radiation in Bioanalysis* 8 (juillet): 173-93. [https://doi.org/10.1007/978-3-030-28247-9\\_6](https://doi.org/10.1007/978-3-030-28247-9_6).

Feigin, L. A., D. I. Svergun, et George W. Taylor. 1987. « Principles of the Theory of X-Ray and Neutron Scattering ». In *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*, édité par L. A. Feigin, D. I. Svergun, et George W. Taylor, 3-24. Boston, MA: Springer US. [https://doi.org/10.1007/978-1-4757-6624-0\\_1](https://doi.org/10.1007/978-1-4757-6624-0_1).

Friskén, B. J. 2001. « Revisiting the Method of Cumulants for the Analysis of Dynamic Light-Scattering Data ». *Applied Optics* 40 (24): 4087-91. <https://doi.org/10.1364/ao.40.004087>.

Guinier, André, Gérard Fournet, C. B. Walker, et George H. Vineyard. 1956. « Small-Angle Scattering of X-Rays ». *Physics Today* 9 (8): 38-39. <https://doi.org/10.1063/1.3060069>.

Hall, Martin. 2018. « Chapter 21 - Size Exclusion Chromatography (SEC) ». In *Biopharmaceutical Processing*, édité par Günter Jagschies, Eva Lindskog, Karol Łacki, et Parrish Galliher, 421-32. Elsevier. <https://doi.org/10.1016/B978-0-08-100623-8.00021-9>.

Hassan, Puthusserickal A., Suman Rana, et Gunjan Verma. 2015. « Making Sense of Brownian Motion: Colloid Characterization by Dynamic Light Scattering ». *Langmuir* 31 (1): 3-12. <https://doi.org/10.1021/la501789z>.

Mesa-Pereira, Beatriz, Mary C. Rea, Paul D. Cotter, Colin Hill, et R. Paul Ross. 2018. « Heterologous Expression of Biopreservative Bacteriocins With a View to Low Cost Production ». *Frontiers in Microbiology* 9. <https://www.frontiersin.org/articles/10.3389/fmicb.2018.01654>.

Narayanan Nair, Maya. 2013. « Functionalization of epitaxial graphene by metal intercalation and molecules », septembre.

Paterson, Gary W, Gavin M Macauley, Stephen McVitie, et Yoshihiko Togawa. 2021. « Parallel Mode Differential Phase Contrast in Transmission Electron Microscopy, I: Theory and Analysis ». *Microscopy and Microanalysis* 27 (5): 1113-22. <https://doi.org/10.1017/S1431927621012551>.

Provencher, Stephen W. 1982. « CONTIN: A general purpose constrained regularization program for inverting noisy linear algebraic and integral equations ». *Computer Physics Communications* 27 (3): 229-42. [https://doi.org/10.1016/0010-4655\(82\)90174-6](https://doi.org/10.1016/0010-4655(82)90174-6).

Quijada, Lugo, et Dersy Marileth. 2010. « Adsorption of Surfactants on Colloidal Silica: Effects of Surface Curvature on the Structure of Surface Aggregates », avril. <https://depositonce.tu-berlin.de/items/urn:nbn:de:kobv:83-opus-26292>.

Receveur-Bréchet, Véronique, et Dominique Durand. 2012. « How Random are Intrinsically Disordered Proteins? A Small Angle Scattering Perspective ». *Current Protein & Peptide Science* 13 (1): 55-75. <https://doi.org/10.2174/138920312799277901>.

Schmidli, Claudio, Luca Rima, Stefan A. Arnold, Thomas Stohler, Anastasia Syntychaki, Andrej Bieri, Stefan Albiez, et al. 2018. « Miniaturized Sample Preparation for Transmission Electron Microscopy ». *JoVE (Journal of Visualized Experiments)*, n° 137 (juillet): e57310. <https://doi.org/10.3791/57310>.

Schroer, Martin A. 2011. « Small Angle X-Ray Scattering Studies on Proteins under Extreme Conditions », novembre. <https://doi.org/10.17877/DE290R-2928>.

« Solution Scattering Data Assessment ». s. d. Consulté le 11 octobre 2023. <https://www-ssrl.slac.stanford.edu/~saxs/analysis/assessment.htm>.

Stetefeld, Jörg, Sean A. McKenna, et Trushar R. Patel. 2016. « Dynamic light scattering: a practical guide and applications in biomedical sciences ». *Biophysical Reviews* 8 (4): 409-27. <https://doi.org/10.1007/s12551-016-0218-6>.

Svergun, Dmitri I., et Michel H. J. Koch. 2003. « Small-Angle Scattering Studies of Biological Macromolecules in Solution ». *Reports on Progress in Physics* 66 (10): 1735. <https://doi.org/10.1088/0034-4885/66/10/R05>.

Trewhella, Jill. 2022. « Recent Advances in Small-Angle Scattering and Its Expanding Impact in Structural Biology ». *Structure* 30 (1): 15-23. <https://doi.org/10.1016/j.str.2021.09.008>.

Tuukkanen, A. T., A. Spilotros, et D. I. Svergun. 2017. « Progress in Small-Angle Scattering from Biological Solutions at High-Brilliance Synchrotrons ». *IUCr* 4 (5): 518-28. <https://doi.org/10.1107/S2052252517008740>.

Wu, Di, et Grzegorz Piszczek. 2021a. « Rapid Determination of Antibody-Antigen Affinity by Mass Photometry ». *Journal of visualized experiments : JoVE*, n° 168 (février): 10.3791/61784. <https://doi.org/10.3791/61784>.

———. 2021b. « Standard Protocol for Mass Photometry Experiments ». *European biophysics journal : EBJ* 50 (3-4): 403-9. <https://doi.org/10.1007/s00249-021-01513-9>.

Young, Gavin, Nikolas Hundt, Daniel Cole, Adam Fineberg, Joanna Andrecka, Andrew Tyler, Anna Olerinyova, et al. 2018. « Quantitative Mass Imaging of Single Biological Macromolecules ». *Science (New York, N.Y.)* 360 (6387): 423-27. <https://doi.org/10.1126/science.aar5839>.

# **Chapitre 4 : Effets des modulateurs d'assemblage sur la cinétique et la structure des capsides formées à partir de la protéine Cp149**

## **1. Introduction**

Comme nous avons pu le voir dans le chapitre 2, le virus de l'Hépatite B est responsable chaque année de l'infection et de la mort de milliers de personnes à travers le monde. Les traitements actuellement approuvés peuvent être contraignants pour les malades. En effet, la durée des traitements est très longue, généralement à vie. De plus, ces traitements permettent uniquement de contrôler l'infection sans la guérir. Il est donc important de développer de nouvelles molécules antivirales qui peuvent être utilisées plus facilement par les patients et qui idéalement conduiraient à une guérison complète du patient.

Nous avons choisi d'étudier l'influence d'une nouvelle classe de molécules virales prometteuses, les modulateurs d'assemblage de la capsid. La capsid du virus de l'Hépatite B qui protège son génome viral et assure ainsi sa transmission efficace d'un hôte à l'autre, est composée de 120 dimères de la protéine Core. L'assemblage de cette capsid, qui joue un rôle crucial dans le cycle viral du virus de l'Hépatite B, peut être perturbée par ces modulateurs. Ainsi, il est fondamental de comprendre comment ces modulateurs interagissent avec la protéine Core et affectent la cinétique d'assemblage ainsi que la structure des capsides, afin de développer de nouvelles stratégies thérapeutiques contre l'infection par le VHB.

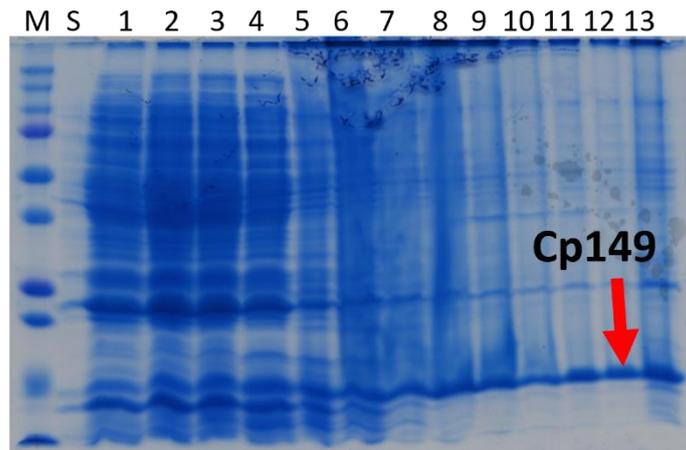
Dans ce chapitre, nous examinons en détail nos résultats obtenus sur les effets des modulateurs d'assemblage sur les capsides du virus de l'Hépatite B formées à partir de la protéine dimérique tronquée Cp149. Nous discutons de l'influence de ces modulateurs sur la cinétique d'assemblage des capsides du virus de l'Hépatite B, notamment de leur capacité à accélérer ou non la formation des capsides. Nous examinons également l'influence de ces molécules sur la structure des capsides du VHB. Enfin, nous analysons de la manière dont ces molécules influencent l'énergie élastique de la capsid.

Afin de préserver le contraste dans la diffusion des rayons X entre la protéine et le solvant, nous avons utilisé un sel plus léger que le chlorure de sodium, qui est le sel utilisé dans la plupart des études précédentes sur l'assemblage de Cp149. Dans un premier temps, nous présenterons une étude sur l'assemblage de la protéine Cp149 en présence de ce nouveau sel, l'acétate d'ammonium (AmAc). Puis, nous présenterons les résultats de notre étude réalisée par TR-SAXS, cryomicroscopie électronique et simulations gros grains sur l'influence des modulateurs Bay 41-4109 (CAM-A) et JNJ-632 (CAM-E) sur la cinétique d'assemblage et la structure des capsides. Ces résultats ont par ailleurs fait l'objet d'un article publié en juin 2023. Dans une troisième partie, nous exposerons les résultats de notre étude portant sur deux nouvelles molécules, GLP-26 et RS-3427, et leur impact sur les capsides du VHB. Ces résultats seront comparés à ceux obtenus lors de l'étude de l'effet du modulateur de type CAM-E, JNJ-632. Enfin, la dernière partie de ce chapitre sera consacrée à la conclusion qui résumera les principales découvertes et implications de ces études.

## **2. Assemblage de la protéine cp149 en absence de modulateurs**

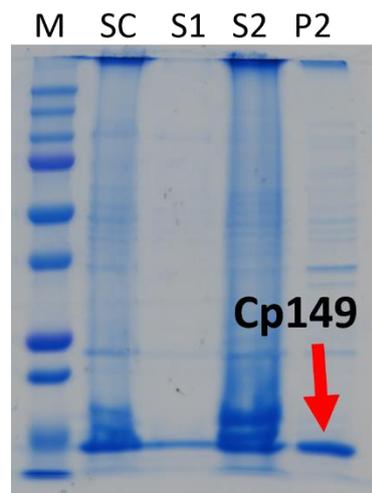
### **2.1 Expression et purification de la protéine dimérique cp149**

La Cp149 est surexprimée dans des bactéries d'*E. coli*, où elle s'assemble en capsides vides, ne contenant pas d'ARN. La première étape de la purification des capsides après la lyse bactérienne est une ultracentrifugation sur gradient de sucrose. Cette étape permet de séparer les dimères non assemblés, et les capsides du VHB d'une grande majorité des protéines bactériennes.



**Figure 4.1 : Gel SDS-PAGE 14% issu de l'étape de centrifugation sur un gradient de sucrose des capsides de Cp149.** Les puits M, S et 1-13 correspondent respectivement au marqueur de poids moléculaire, au surnageant obtenu après la lyse bactérienne et aux 13 fractions prélevées (La fraction 1 correspond au haut du gradient et la 13 au bas du gradient) après la centrifugation sur un gradient de sucrose. 10  $\mu$ l de protéines sont déposés dans chaque puit et 5  $\mu$ l pour le marqueur. La flèche rouge indique la bande protéique Cp149.

Ce gel SDS-PAGE (Figure 4.1) des différentes fractions obtenues après la centrifugation sur un gradient de sucrose montre que Cp149 est présente dans toutes les fractions. Nous décidons de poursuivre la purification des capsides en utilisant les fractions 8 à 12 car ce sont les fractions contenant le moins de protéines contaminantes.

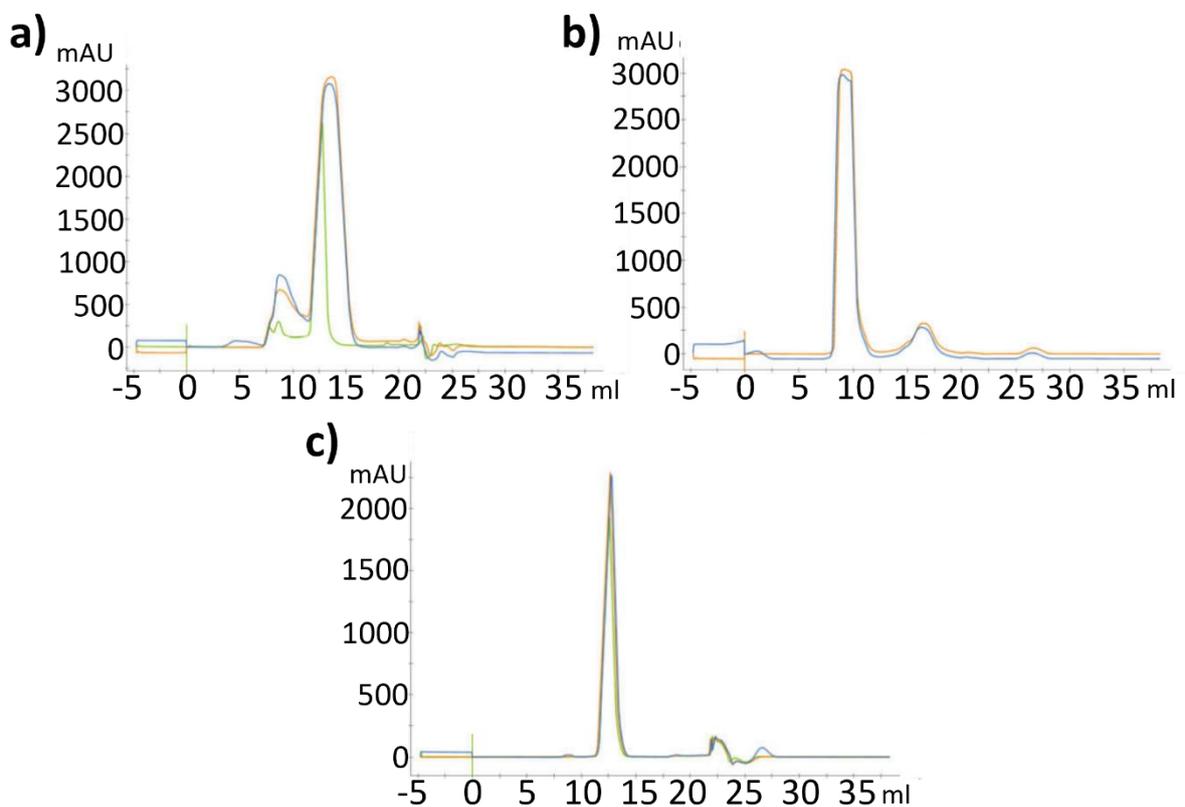


**Figure 4.2 : Gel SDS-PAGE 14% issu de la purification des capsides de Cp149.** Les puits M, SC, S1, S2, P2 correspondent respectivement au marqueur de poids moléculaire, à l'échantillon après la précipitation au sulfate d'ammonium et non centrifugé, au surnageant issu de la première centrifugation après la précipitation au sulfate d'ammonium, au surnageant et au culot issu de la deuxième

centrifugation après re-solubilisation du culot de précipitation au sulfate d'ammonium. 10  $\mu$ l de protéines sont déposés dans chaque puit et 5  $\mu$ l pour le marqueur. La flèche rouge indique la bande protéique confirmant la présence de Cp149.

Les fractions 8 à 12 ont été rassemblées et les protéines ont été précipitées avec du sulfate d'ammonium. Ce gel SDS-PAGE (Figure 4.2) montre les différentes fractions protéiques obtenues dans l'étape de précipitation avec 40 % de sulfate d'ammonium. Les capsides sont présentes dans toutes les fractions et en grande quantité dans la fraction S2 (re-solubilisation du culot de précipitation). Pour la suite des expériences, nous utilisons donc la fraction S2.

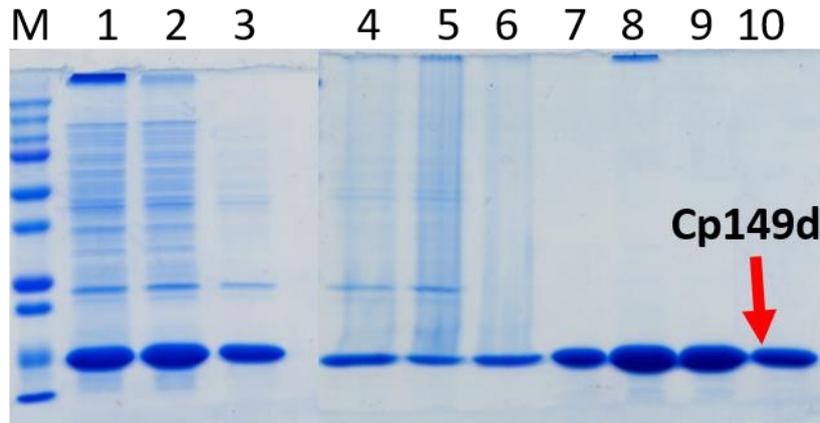
À la suite de la purification des capsides, nous obtenons 6 ml de protéines à 7 mg/ml à partir d'un litre de culture contenant 6 g de bactéries.



**Figure 4.3 : Chromatogrammes de purification par exclusion de taille de la Cp149.** Les 3 chromatogrammes sont issus (a) de la purification des dimères de la Cp149 après la première étape de dissociation, (b) de la purification des capsides de Cp149 réassemblées (en amenant la solution à 500 mM de chlorure de sodium) du matériel non réassemblé, et (c) de la seconde purification des dimères de Cp149 après une nouvelle dissociation. Le volume de la solution a été divisé en 3 avant d'être purifié sur la colonne de chromatographie pour les étapes (a) et (c) (3 chromatogrammes) et par 2 pour l'étape (b) (2 chromatogrammes).

La moitié de la fraction S2 a été dialysée dans un tampon à basse salinité et pH élevé et les capsides dissociées en dimères par addition d'urée. Les dimères ont été purifiés par chromatographie d'exclusion de taille. Les trois chromatogrammes de la figure 4.3 sont issus des chromatographies successives réalisées lors de la purification des dimères. Le premier (a), correspond à la première étape de purification des dimères. Ceux-ci sont élués à un volume d'environ 14 ml de tampon et sont majoritaires dans notre solution. Nous remarquons cependant la présence d'un autre pic moins élevé éluant à partir d'environ 8 ml (volume mort de la colonne) et qui correspond aux différents agrégats et/ou aux capsides non dissociées. Le deuxième chromatogramme (b), est issu de la purification des capsides réassemblées. Nous pouvons également observer deux pics sur ce chromatogramme. Le premier pic correspond à celui des capsides éluées à partir d'environ 8 ml de tampon. Le deuxième pic, plus petit, à environ 16 ml correspond à du matériel non assemblé. Enfin, le troisième chromatogramme (c), est issu de la deuxième et dernière purification des dimères après une nouvelle dissociation des capsides purifiées en (b). Nous observons uniquement un pic. Il correspond aux dimères élués avec environ 13 ml de tampon. La présence d'un seul pic indique que les dimères issus de cette étape sont moins contaminés que ceux issus de la première purification.

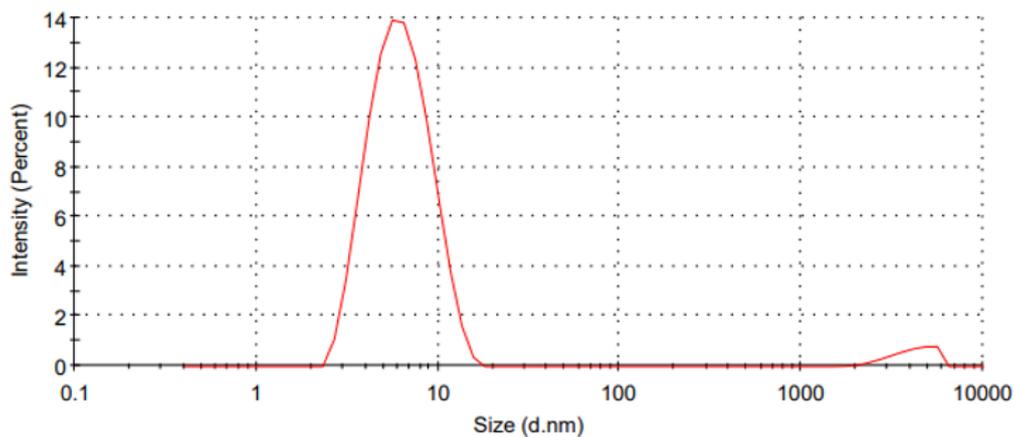
A la suite de la purification des dimères, nous obtenons 3 ml de dimères de Cp149 à 0,88 mg/ml. Soit un rendement de purification de 13%. Cependant, de manière générale, à partir d'un litre de culture contenant 6 à 7 g de bactéries, le rendement de purification est d'environ 20%.



**Figure 4.4 : Gel SDS-PAGE 14% issu de la purification des dimères de Cp149.** 10 µl de protéines sont déposés dans chaque puit et 5 µl pour le marqueur. La flèche rouge indique la bande protéique confirmant la présence de Cp149. Les puits M et 1 à 10 correspondent respectivement à :

- (M) marqueur de poids moléculaire.
- (1), (2), et (3) Echantillon après la dialyse à pH 9, après la dissociation avec 3 M d'urée et après la première purification par chromatographie d'exclusion de taille (SEC).
- (4), (5), et (6) Echantillon après la dialyse d'assemblage des capsides, après la concentration et après la deuxième purification par SEC.
- (7), (8), (9), et (10) Echantillon après la dialyse à pH 9, après la concentration, après la dissociation avec 3 M d'urée et après la troisième purification par SEC.

Ce gel SDS-page (Figure 4.4) réalisé à la fin de la purification des dimères, permet de contrôler la présence de contaminants à chaque étape de la purification et dans notre échantillon final. Nous remarquons la présence de la protéine Cp149 dans tous les puits. Après la première purification de dimères, nous remarquons dans le puit 3 la présence de la Cp149, mais également de protéines contaminantes. Pour les expériences d'assemblage, obtenir une solution pure de dimères compétents pour l'assemblage en capsides est très important afin d'éviter toutes altérations des résultats. En effet, les quantités de sels, mais également de modulateurs ajoutés lors de l'assemblage dépendent de la quantité de dimères capables de s'assembler réellement présents dans la solution. Plus la solution est impure, plus la concentration en dimères sera fausse. Les résultats obtenus seront donc difficilement interprétables. Cependant, après la deuxième et dernière purification des dimères, nous pouvons voir que la protéine dimérique Cp149 est bien présente, qu'il n'y a pas de capsides ni de protéines contaminantes dans notre échantillon. La solution de dimères est donc pure.



**Figure 4.5 : Mesure par DLS des dimères de la Cp149 obtenus après la purification.**

La mesure par DLS représentée sur la figure 4.5, indique que le diamètre hydrodynamique des objets mesuré est égal à 6,5 nm et que l'indice de polydispersité est de 0,218. Ces résultats confirment donc que seuls des dimères de Cp149 sont présents dans notre échantillon.

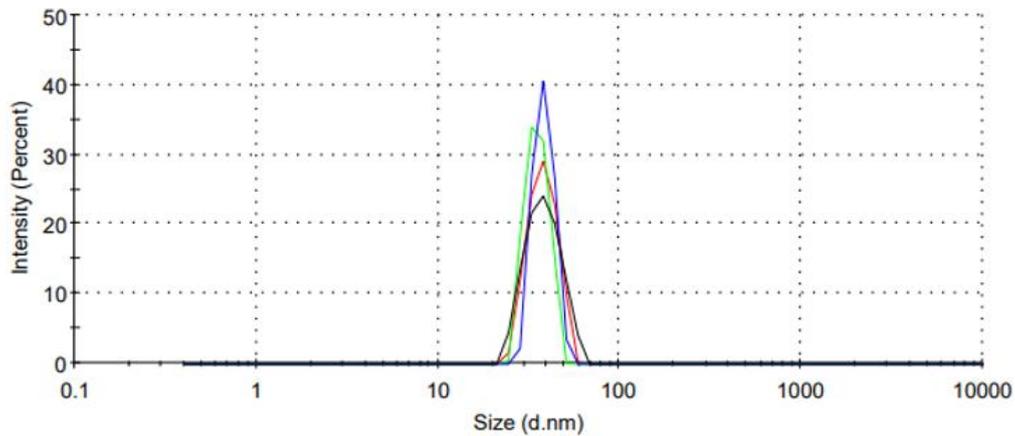
## **2.2 En présence d'acétate d'ammonium, la protéine dimérique cp149 s'autoassemble en capsid T=4**

L'assemblage des capsides de Cp149 *in vitro* est généralement déclenché par ajout de chlorure de sodium. Pour des raisons techniques, nous avons décidé d'assembler les capsides de Cp149 avec un autre sel, l'acétate d'ammonium (AmAc). Ce sel monovalent réduit le bruit sur les courbes de SAXS et permet d'obtenir des courbes de meilleure qualité. Nous avons donc réalisé des assemblages avec ces deux sels à différentes concentrations. Nous avons également étudié l'effet du pH sur l'assemblage des capsides. En effet, afin d'être plus proche des conditions physiologiques, nous avons choisi de réaliser les assemblages à pH 7,5 et à faible concentration saline. Nos résultats pourront ensuite être comparés à l'étude réalisée par Chevreuil *et al.*, sur l'assemblage de la Cp149 à pH 9 ainsi qu'à l'étude d'Asor *et al.*, sur l'assemblage de la Cp149 en présence d'acétate d'ammonium.

Ainsi, l'assemblage par dialyse à 4°C de 26 µM (0,88 mg/ml) de protéines Cp149 dimériques a été testé avec les tampons ci-dessous :

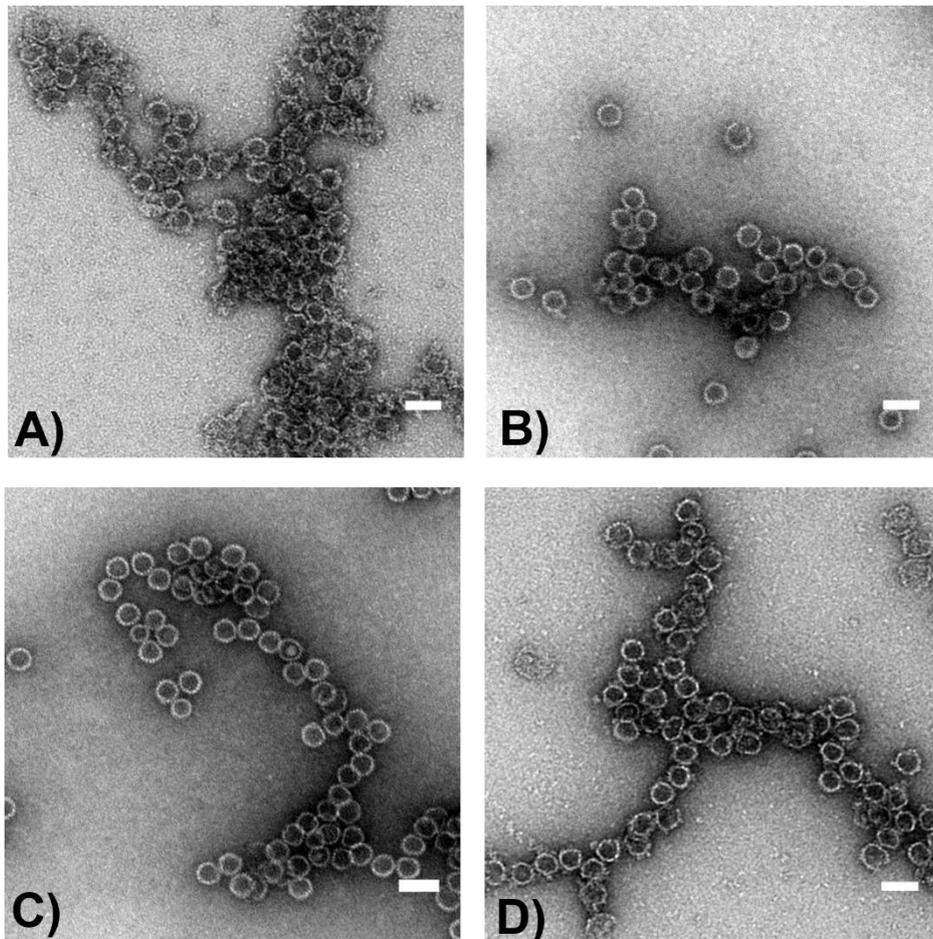
- Tampon A : 50 mM CHES pH 9 ; 1 mM DTT ; 500 mM NaCl
- Tampon B : 50 mM Tris HCl pH 7,5 ; 1 mM DTT ; 500 mM NaCl

- Tampon C : 50 mM Tris HCl pH 7,5 ; 1 mM DTT ; 500 mM AmAc
- Tampon D : 50 mM Tris HCl pH 7,5 ; 1 mM DTT ; 150 mM AmAc



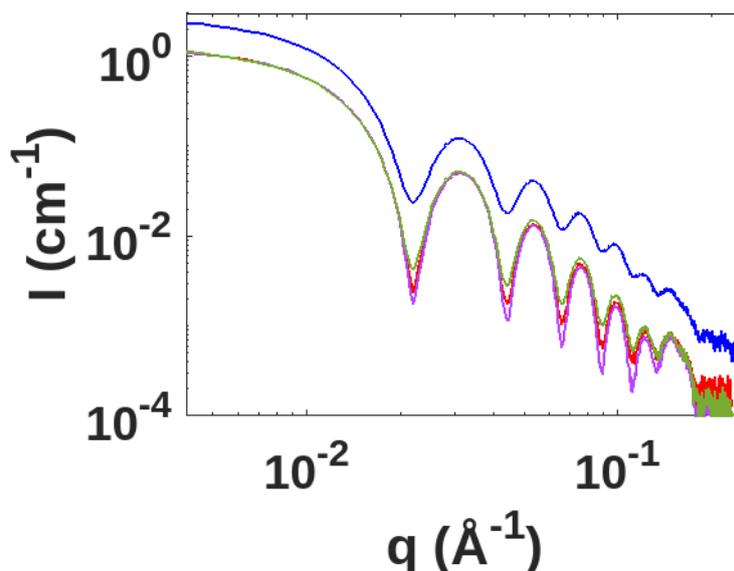
**Figure 4.6 : Mesures par DLS des objets formés après l’assemblage de la Cp149 dans différents tampons.** La ligne rouge correspond à un assemblage avec le tampon A, la ligne verte à un assemblage avec le tampon B, la ligne bleue à un assemblage avec le tampon C et la ligne noire à un assemblage avec le tampon D.

Les résultats des mesures de DLS (Figure 4.6) des capsides formées avec les différents tampons indiquent que leur diamètre hydrodynamique est situé entre 34 et 41 nm et leur indice de polydispersité entre 0,01 et 0,09. Les diamètres et les indices de polydispersité étant très similaires, nous pouvons donc en conclure que les capsides formées dans ces différents tampons sont similaires et homogènes. Cependant, la technique d’analyse utilisée ici ne permet pas de distinguer d’éventuelles différences structurales entre les échantillons.



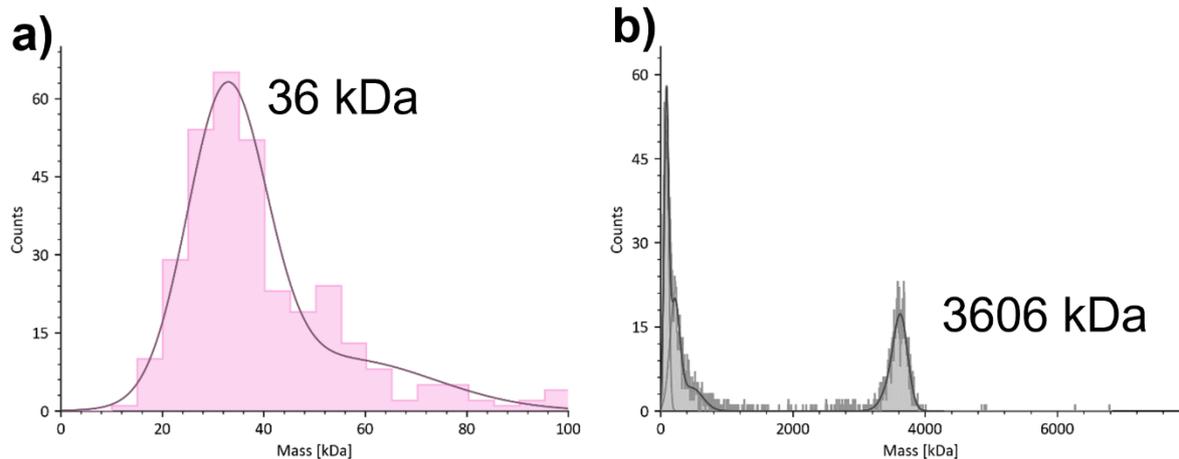
**Figure 4.7 : Micrographies des capsides assemblées dans différents tampons et observées en coloration négative au microscope électronique.** (A) Capsides assemblées dans le tampon A. (B) Capsides assemblées dans le tampon B. (C) Capsides assemblées dans le tampon C. (D) Capsides assemblées dans le tampon D. Les barres d'échelle sont égales à 50 nm. Les échantillons sont observés sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN Ultrascan 4K CCD.

Les capsides formées dans les différents tampons ont ensuite été diluées dans leur tampon respectif afin d'être observées à une concentration de 0,3 mg/ml sur des grilles de microscopie électronique colorées avec de l'acétate d'uranyle à 1%. Les images de microscopie électronique (Figure 4.7) révèlent que les capsides formées dans les cinq conditions présentent une morphologie sphérique. Seules très peu de capsides sont ouvertes et nous n'observons pas de capsides endommagées. Ces observations montrent donc que les capsides formées à pH 9 ou pH 7,5, en présence d'acétate d'ammonium ou du chlorure de sodium à 500 ou 150 mM ont la même morphologie.



**Figure 4.8 : Courbes SAXS des capsides assemblées dans différents tampons à pH 9 ou à pH 7,5.** Courbes SAXS expérimentales des capsides assemblées à partir de 26  $\mu\text{M}$  de dimères de Cp149 dans le tampon A (courbe bleu), le tampon B (courbe rouge), le tampon C (courbe violette) et le tampon D (courbe verte).

Enfin, l'analyse des courbes SAXS présentées dans la figure 4.8, permet de distinguer des altérations de la structure des capsides formées en présence des différents tampons. La présence d'oscillations, observées dans toutes les conditions indique que des capsides sont bien formées dans chacune d'elles. En effet, ces oscillations sont caractéristiques des courbes SAXS associées à des objets de symétrie sphérique. L'allure des courbes dans la région de Guinier suggère que ces objets ne sont pas agrégés et le premier minimum à  $q = 0,022 \text{ \AA}^{-1}$  indique que les capsides formées sont des capsides T=4. Nous remarquons cependant que selon les conditions d'assemblage, les premières oscillations varient en profondeur. Lors d'un assemblage à pH 9 (courbe bleue), ces oscillations peu profondes, indiquent la présence de capsides incomplètes et d'une plus grande concentration de dimères non assemblés. Ce phénomène n'est pas observé à pH 7,5. Par conséquent, il est préférable de réaliser des assemblages dans un tampon à pH 7,5. Ensuite lors d'un assemblage à pH 7.5 en présence de chlorure de sodium ou d'acétate d'ammonium à 150 ou 500 mM, les oscillations sont plus profondes et peu de différences entre les courbes SAXS sont observables. Ainsi, pour la suite des expériences d'assemblage de la Cp149, nous avons choisi de travailler à pH 7.5 en présence de 150 mM d'acétate d'ammonium (courbe verte), sauf indication contraire.



**Figure 4.9 : Analyse par photométrie de masse de la distribution de masse de la protéine Cp149.** (a) Distribution de masse des dimères de Cp149. (b) Distribution de masse des capsides de Cp149.

La distribution de masse des dimères de Cp149 ainsi que des capsides formées à partir de  $30 \mu\text{M}$  ( $1 \text{ mg/ml}$ ) de dimères de Cp149 en présence de  $150 \text{ mM}$  d'acétate d'ammonium a été analysée par photométrie de masse (Figure 4.9). En présence du dimère (a), nous observons un seul pic légèrement asymétrique de  $36 \text{ kDa}$ . La masse mesurée par cette méthode des dimères de Cp149 est donc très proche de la masse théorique qui est de  $34 \text{ kDa}$ . En présence des capsides (b), nous observons un pic à  $3606 \text{ kDa}$  qui correspond donc à la masse des capsides. Cependant, cette mesure est inférieure à la masse théorique attendue de la capside de Cp149, établie à  $4042 \text{ kDa}$ . Cette différence pourrait être due à la présence de capsides incomplètes ou bien à une calibration de l'appareil de masse photométrie utilisant des protéines dont la masse moléculaire est nettement inférieure à celle des capsides. De plus, nous observons deux autres pics de  $89$  et  $252 \text{ kDa}$ . Il pourrait s'agir d'oligomères de Cp149.

### **3. Article : *Energetics and Kinetics Assembly Pathway of Hepatitis B Virus Capsids in the Presence of Antivirals***

#### **3.1 Introduction**

Notre article "*Energetics and kinetics assembly pathway of Hepatitis B virus capsids in the presence of antivirals*" a été accepté en juin 2023 dans le journal ACS Nano (Kra et al. 2023). Cet article porte sur l'étude des effets de deux modulateurs, Bay 41-4109, un modulateur de type CAM-A et JNJ-632, un modulateur de type CAM-E sur la cinétique d'assemblage des capsides du virus de l'Hépatite B. Une étude structurale ainsi que des simulations ont également été réalisées. La préparation des échantillons a été effectuée par Laetitia Poncet et moi-même et les expériences de cryoTEM ont été réalisées avec l'aide de Jéril Degrouard. L'analyse des résultats de SAXS, TR-SAXS et de cryoTEM de l'assemblage de 30  $\mu\text{M}$  de sous-unités de Cp149 en présence de 10/100  $\mu\text{M}$  de Bay 41-4109 ou 20/200  $\mu\text{M}$  JNJ-632 a été réalisée par moi-même, avec l'aide de Stéphane Bressanelli et de Guillaume Tresset mes deux superviseurs. Puis, les modèles cinétiques ont été réalisés par Guillaume Tresset. Enfin, les simulations gros grains des capsides ont été réalisées par nos collaboratrices Siyu Li et Roya Zandi à l'Université de Californie Riverside.

Des expériences ayant été réalisées en amont de celles décrites dans cet article seront présentées dans la partie 3.3. Il s'agit des mesures par DLS et des images de microscopie électronique des capsides assemblées en présence de ces deux modulateurs. Une expérience de cinétique par cryomicroscopie électronique des objets formés en présence de la CAM-A sera également présentée.

This article is reprinted with permission from ACS Nano 2023, 17, 12723–12733. Copyright 2023 American Chemical Society.

# Energetics and kinetic assembly pathways of Hepatitis B virus capsids in the presence of antivirals

*Kalouna Kra,<sup>1,2</sup> Siyu Li,<sup>3</sup> Laetitia Gargowitsch,<sup>1</sup> Jéril Degrouard,<sup>1</sup> Javier Pérez,<sup>4</sup> Roya Zandi,<sup>3</sup> Stéphane Bressanelli<sup>2\*</sup> and Guillaume Tresset<sup>1\*</sup>*

<sup>1</sup>Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France.

<sup>2</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198 Gif-sur-Yvette, France.

<sup>3</sup>Department of Physics and Astronomy, University of California, Riverside, United States.

<sup>4</sup>SOLEIL Synchrotron, 91192 Gif-sur-Yvette, France.

\*Corresponding authors : [stephane.bressanelli@i2bc.paris-saclay.fr](mailto:stephane.bressanelli@i2bc.paris-saclay.fr);  
[guillaume.tresset@universite-paris-saclay.fr](mailto:guillaume.tresset@universite-paris-saclay.fr)

Capsid assembly modulators (CAMs) are antiviral molecules that disturb the formation of icosahedral viral capsids, in particular, those of the Hepatitis B virus (HBV). We report an integrated, physics-driven study elucidating quantitatively the effects of two classes of CAMs on HBV capsid assembly. Time-resolved small-angle X-ray scattering measurements revealed accelerated self-assembly processes that implied the increase of subunit binding energy from 9- up to 18-fold the thermal energy due to CAMs. Cryotransmission electron microscopy images showed that both classes induce various changes in capsid morphology: from a slight elongation, unrecognized in previous work, to a strong deformation with a capsid size more than twice as large. The observed capsid morphologies were closely reproduced in coarse-grained simulations by varying the Föppl-von-Kármán number, thus pointing out the role of CAMs in altering the capsid elastic energy. Our results illuminate the mechanisms of action of CAMs on HBV capsid assembly at high spatiotemporal resolution, and may open up perspectives on virus-derived nanocapsules with tunable morphologies.

**KEYWORDS:** Hepatitis B virus, capsid assembly, assembly modulators, kinetic modelling, capsid mechanics, time-resolved small-angle X-ray scattering, cryotransmission electron microscopy

The Hepatitis B virus (HBV) is a major health problem worldwide. Despite the existence of a safe and effective vaccine, more than 292 million people are chronically infected by HBV, mainly in Africa and South Asia, due to a poor vaccination rate. This virus causes serious pathologies such as liver cirrhosis in some patients, but also hepatocellular cancer that can lead to death.<sup>1</sup> The treatments currently approved use nucleos(t)ide analogues. However, the duration of the treatments is very long, usually lifelong,<sup>2</sup> and they only permit to control the infection but not to cure it. Therefore, it is important to develop new antiviral compounds with alternative mechanisms of action that may be used more easily, in combination or not with other drugs, in order to completely treat the infection.

HBV is an enveloped virus with an icosahedral nucleocapsid that contains the viral genome in the form of a 3.2 kb partially double-stranded DNA. The HBV capsid, which plays an essential role in the HBV viral cycle, is composed of a homodimeric Core protein.<sup>3</sup> Core is a 183-residue-long protein composed of two domains. The 149 N-terminal residues are an alpha-helical assembly domain called N-terminal domain (NTD) referred to as Core protein 149 (Cp149).<sup>4</sup> The 34 remaining residues are the intrinsically disordered C-terminal domain (CTD) which interacts with RNA and is required for its encapsidation. Cp149 can self-assemble<sup>5,6</sup> both *in vivo* and *in vitro* mainly in  $T = 4$  icosahedral capsid (120 dimers) but also in  $T = 3$  icosahedral capsid (90 dimers).

Early studies suggested that Cp149 capsid assembly follows a nucleation-elongation process.<sup>7-9</sup> Nucleation consists of the formation of a critical nucleus made of dimers, the latter being the building blocks of the assembly and hereafter called subunits. Free subunits are sequentially added to the nucleus during the elongation phase to form the final capsid. There are at most a few long-lived intermediate species present during this step.<sup>10</sup> Many distinct factors, including the ionic strength,<sup>11</sup> the stability of the intra-subunit interface<sup>12</sup> and/or the contact region of the subunits,<sup>13</sup> influence HBV capsid assembly and its duration, but this regulation, however, occurs without the assistance of chaperone proteins.

Since the intermediate steps can occur very fast, following and solving the structures involved in the assembly are difficult. With a spatial sensitivity of a few nanometers and a temporal precision of a few milliseconds, time-resolved small-angle X-ray scattering (TR-SAXS) enables to track complex reactions,<sup>14</sup> notably those pertaining to virus self-assembly.<sup>10,15-22</sup> TR-SAXS data revealed that the assembly of Cp149 empty capsids appeared to be a two-state process involving solely dimeric subunits and  $T = 4$  capsid under low salt conditions – *i.e.*, physiological pH  $\sim 7.5$  and monovalent salt concentration  $\sim 150$  mM.<sup>18</sup> Nevertheless, the development of  $T = 3$  capsids and the brief accumulation of intermediates containing between 7 and 35 subunits were reported under high salt conditions<sup>18</sup> – *i.e.*,  $> 300$  mM. In another study,<sup>20</sup> under conditions where pH was far from being physiological – *i.e.*, pH 9.0 –, neither the intermediates nor the  $T = 3$  minor species were detected

at high salt concentration. However, an additional assembly step, referred to as relaxation, was observed, during which misassembled capsids rearranged themselves via a disorder-order transition into well-ordered capsids.

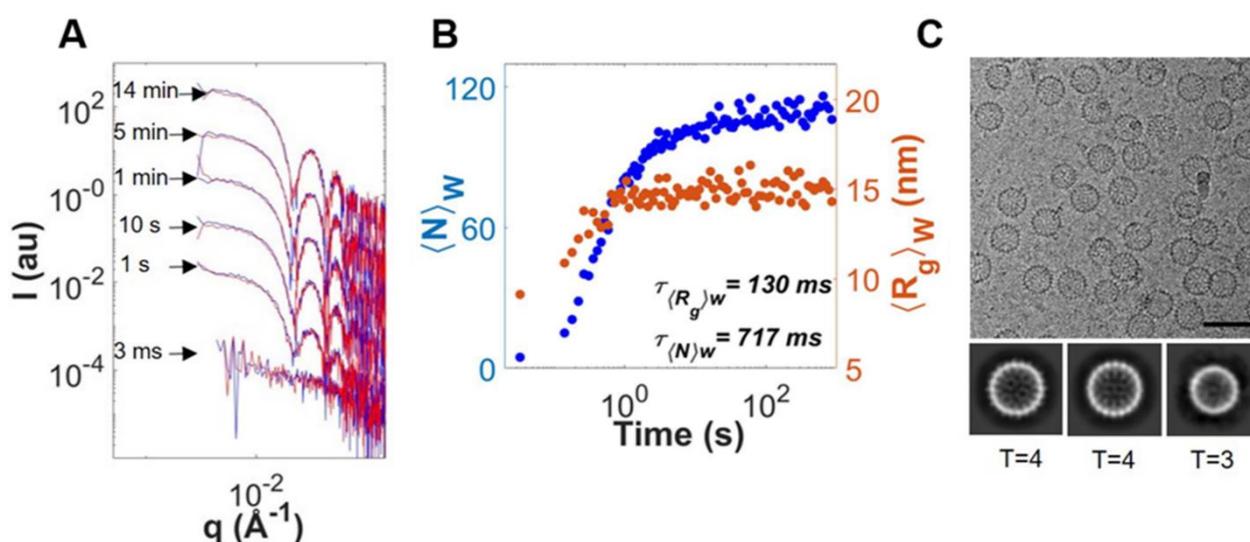
Even though the HBV assembly to icosahedral shell is robust under many different *in vitro* and *in vivo* conditions, small molecules known as capsid assembly modulators (CAMs) have been developed to target the HBV Core protein, which has proven to be a successful new antiviral target.<sup>2</sup> Through a number of ways, they can inhibit HBV replication and possibly even combat virus persistence:<sup>3,23</sup> Their main effect is to disrupt capsid assembly, but some CAMs have been shown to be capable also of blocking the recycling pathway of nucleocapsids to the nucleus. CAMs actually accelerate capsid assembly in a dose-dependent manner by binding to a hydrophobic pocket that mediates contacts between dimers of the Core protein in the final capsid.<sup>24,25</sup> According to their mode of action, two classes of modulators have been defined. Heteroaryldihydropyrimidines (HAP) induce the formation of aberrant capsids<sup>26</sup> and are the only molecules belonging to the class CAM-A (for aberrant).<sup>27</sup> Class CAM-E (for empty)<sup>27</sup> is made up of several different types of molecules, e.g., sulfamoylbenzamides (SBA), phenyl propenamides derivatives (PPA),<sup>25,28</sup> and glyoxamide-pyrrolamide (GPA). The latter molecules, unlike HAP, are thought to induce the formation of morphologically intact HBV capsids. However, they inhibit the formation of nucleocapsids containing pregenomic (pg)RNA.<sup>29,30</sup> Several *in vitro* investigations were conducted about the effects of modulators on the capsid morphology, but the kinetic assembly pathways with CAMs have not been explored yet.

We report here a multidisciplinary and integrated study aimed at shedding light quantitatively on the effects of JNJ-632<sup>25,31</sup> and Bay 41-4109,<sup>32,33</sup> two modulators of type CAM-E and CAM-A, respectively, on the assembly of Cp149 capsids at both nano- and millisecond-scale resolutions. We chose physicochemical conditions – *i.e.*, salinity of 150 mM, pH 7.5 and temperature of 37 °C – that reproduced those found in the intracellular milieu. Thanks to the high spatiotemporal sensitivity of TR-SAXS, we were able to probe the assembly timescales and to investigate the capsid morphologies with different stoichiometric ratios of modulator. We further investigated capsid morphologies by cryotransmission electron microscopy (cryoTEM) to provide nanometer-scale structural details at various time points. By using a kinetic model of phase transition, we were able to measure the effective subunit binding free energy in the presence of modulator. Finally, we related the observed morphologies of final capsids to their elastic properties via coarse-grained simulations, and we propose mechanisms by which modulators might act at the molecular level.

## Results

### Without modulators, a majority of $T = 4$ capsids were self-assembled.

SAXS data measurements of subunits prior to assembly were consistent with Cp149 dimers, albeit slightly aggregated (Figure S1). Capsids were assembled by mixing manually subunits at concentrations of 15  $\mu\text{M}$  to 30  $\mu\text{M}$  with an assembly buffer, bringing the monovalent salt ammonium acetate up to a final concentration of 150 mM, at pH 7.5 and 37°C. The scattering intensities of reassembled capsids were measured after 48 hours. They displayed oscillations with a first minimum at  $q = 0.022 \text{ \AA}^{-1}$ , which indicated that purified subunits led to the formation of a majority of  $T = 4$  capsids (Figure S2).



**Figure 1. TR-SAXS and cryoTEM measurements for the assembly of capsids without modulator. (A)** TR-SAXS patterns of two repeated assembly experiments (blue and red lines, respectively) for 30  $\mu\text{M}$  of subunits. The assembly was triggered with 150 mM of ammonium acetate. The patterns at different time points are shifted for clarity. **(B)** Evolution of the mean aggregation number  $\langle N \rangle_w$  (blue discs) and the mean radius of gyration  $\langle R_g \rangle_w$  (orange discs) as a function of time for the assembly of capsids with 30  $\mu\text{M}$  of subunits. The timescales were obtained by fitting an exponential decay function. **(C)** CryoTEM micrograph (top) and 2D classification (bottom) of manually assembled capsids. Capsids were assembled with a subunit concentration of 30  $\mu\text{M}$  then concentrated 10 times. Scale bar is 50 nm. The box size for 2D classification is 50 nm. Number of  $T = 4$  particles: 3,476; and number of  $T = 3$  particles: 47.

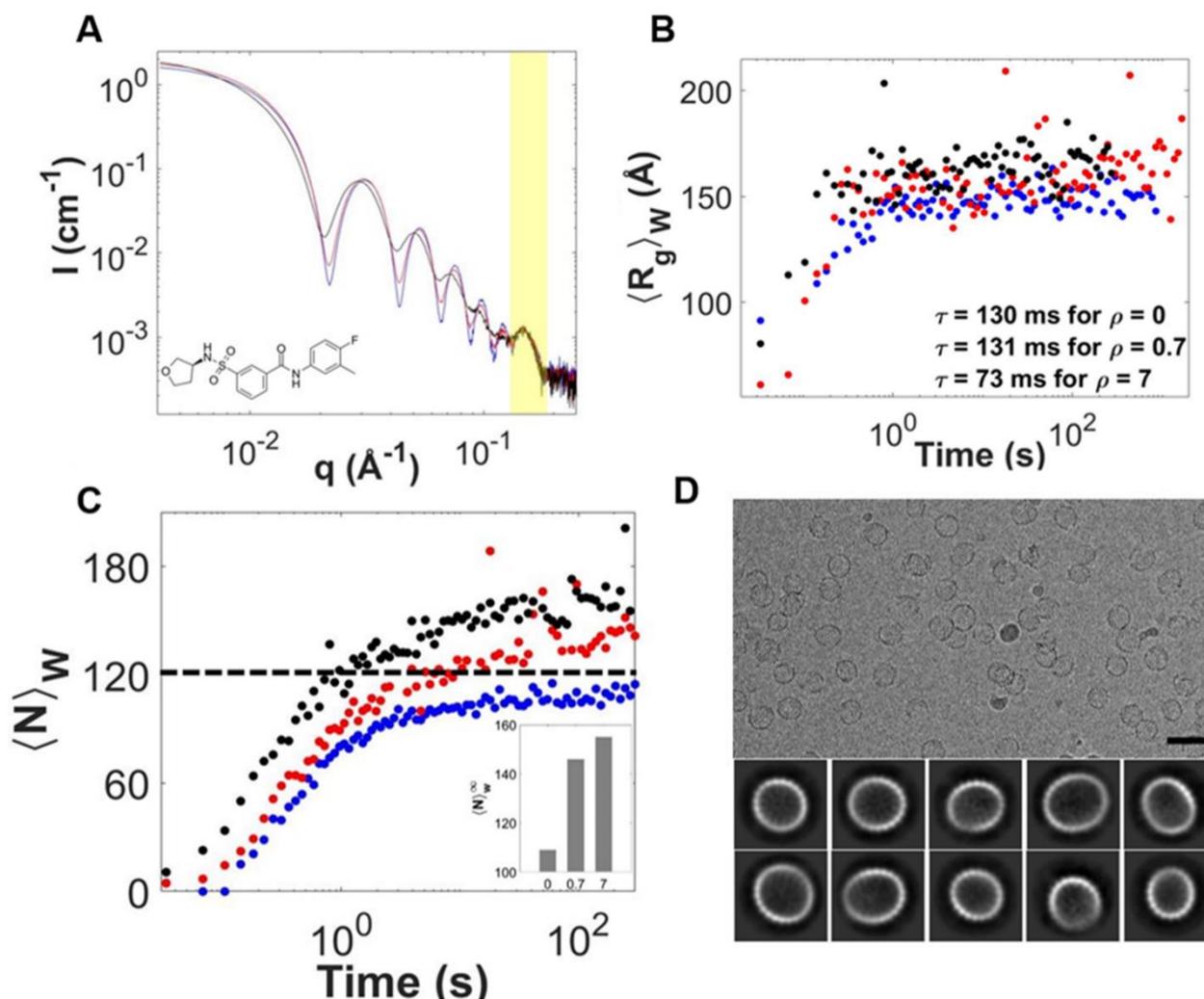
To examine the processes of capsid assembly, TR-SAXS experiments were carried out using a stopped-flow mixer. Capsid assembly was performed at 37°C by rapidly mixing a solution of subunits

with a buffer solution containing a high concentration of ammonium acetate, both at pH 7.5. The final concentration of subunits was 30  $\mu\text{M}$  ( $\sim 1 \text{ g.L}^{-1}$ ), and the final salt concentration was 150 mM. The scattering patterns were collected at different time points for 15 minutes, and each condition was duplicated. SAXS curves are presented in [Figure 1A](#). The SAXS pattern at the earliest time point (3 ms) exhibited a monotonic decrease of the scattering intensity with the wavenumber  $q$ , indicating that initially, mostly dissociated subunits were present in the solution. One second after triggering the assembly, oscillations appeared on the SAXS patterns, and the first minimum at  $q = 0.023 \text{ \AA}^{-1}$  indicated the formation of  $T = 4$  capsids. After 4 s, the patterns remained essentially unchanged, suggesting that the assembly had reached equilibrium. No signs of aggregation were seen during the assembly and capsids were self-assembled with excellent repeatability ([Figure 1A](#), compare blue and red lines). The SAXS curves obtained at 4 s and 850 s after triggering assembly with the stopped-flow mixer, and that measured at equilibrium via manual mixing showed a strong degree of similarity ([Figure S3](#)), demonstrating that the use of a stopped-flow mixer produced eventually the same  $T = 4$  capsids as those assembled manually under the same conditions.

[Figure 1B](#) shows the mean radius of gyration  $\langle R_g \rangle_w$  and the mean aggregation number  $\langle N \rangle_w$  extracted from each SAXS curve of the kinetics (see Methods for details). The extracted timescales for each quantity were calculated by fitting the data with single exponential decay functions. The extracted timescale for  $\langle R_g \rangle_w$  was  $130 \pm 22 \text{ ms}$  (95% confidence interval).  $\langle R_g \rangle_w$  stopped evolving about 1 s after the start of the assembly and plateaued at a value of about 150  $\text{\AA}$ . The mean aggregation number  $\langle N \rangle_w$ , corresponding to the weight-averaged number of subunits found in each object during assembly plateaued slightly below 120, which would have been the value obtained for a pure solution of  $T = 4$  capsids. This was due to the presence of remaining free subunits at the end of assembly as expected from the law of mass action. The extracted timescale observed for the mean aggregation number was  $717 \pm 54 \text{ ms}$  (95% confidence interval), which was higher than that for the mean radius of gyration, *i.e.*, 130 ms.

The capsid morphology of manually assembled capsids was further investigated by cryotransmission electron microscopy (cryoTEM). As shown in the 2D classes ([Figure 1C](#)), a vast majority of  $T = 4$  capsids were formed. Furthermore, a very small number of incomplete capsids was observed and less than 5% of the capsids exhibited a size compatible with the  $T = 3$  symmetry. From this 2D classification, a 3D map ([Figure S4](#)) was obtained by imposing icosahedral symmetry during 3D reconstruction, which matched very well the  $T = 4$  crystal structure.

CAM-E shortened the assembly timescale while inducing structural changes on capsids.



**Figure 2. Capsid assembly in the presence of CAM-E.** **(A)** SAXS patterns at equilibrium with modulator-to-subunit molar ratios  $\rho$  of 0 (blue), 0.7 (red) and 7.0 (black). The shaded area highlights the region where all the patterns were superimposed. The chemical structure of CAM-E is represented at the bottom left. **(B)** Evolution of the mean radius of gyration  $\langle R_g \rangle_w$  as a function of time. The timescales were obtained by fitting exponential decay functions. **(C)** Plot of  $\langle N \rangle_w$  as a function of time. The inset gives the equilibrium mean aggregation numbers  $\langle N \rangle_w^\infty$  as a function of  $\rho$ . For **(A)**, **(B)**, and **(C)**, the subunit concentration was 30  $\mu\text{M}$  and the color code for the  $\rho$  values is the same. **(D)** CryoTEM micrograph (top) and 2D classification (bottom) of capsids assembled with  $\rho = 7.0$  and a subunit concentration of 30  $\mu\text{M}$  before being concentrated about 10 times. The scale bar and the box size for 2D classification are 50 nm.

Capsid assembly was triggered manually with 150 mM of ammonium acetate, 30  $\mu\text{M}$  of subunits, and varying modulator-to-subunit molar ratios  $\rho$  of CAM-E (see chemical structure on [Figure 2A](#)). The scattering intensities measured at equilibrium ([Figure 2A](#)) indicated that the forward intensity  $I_0$

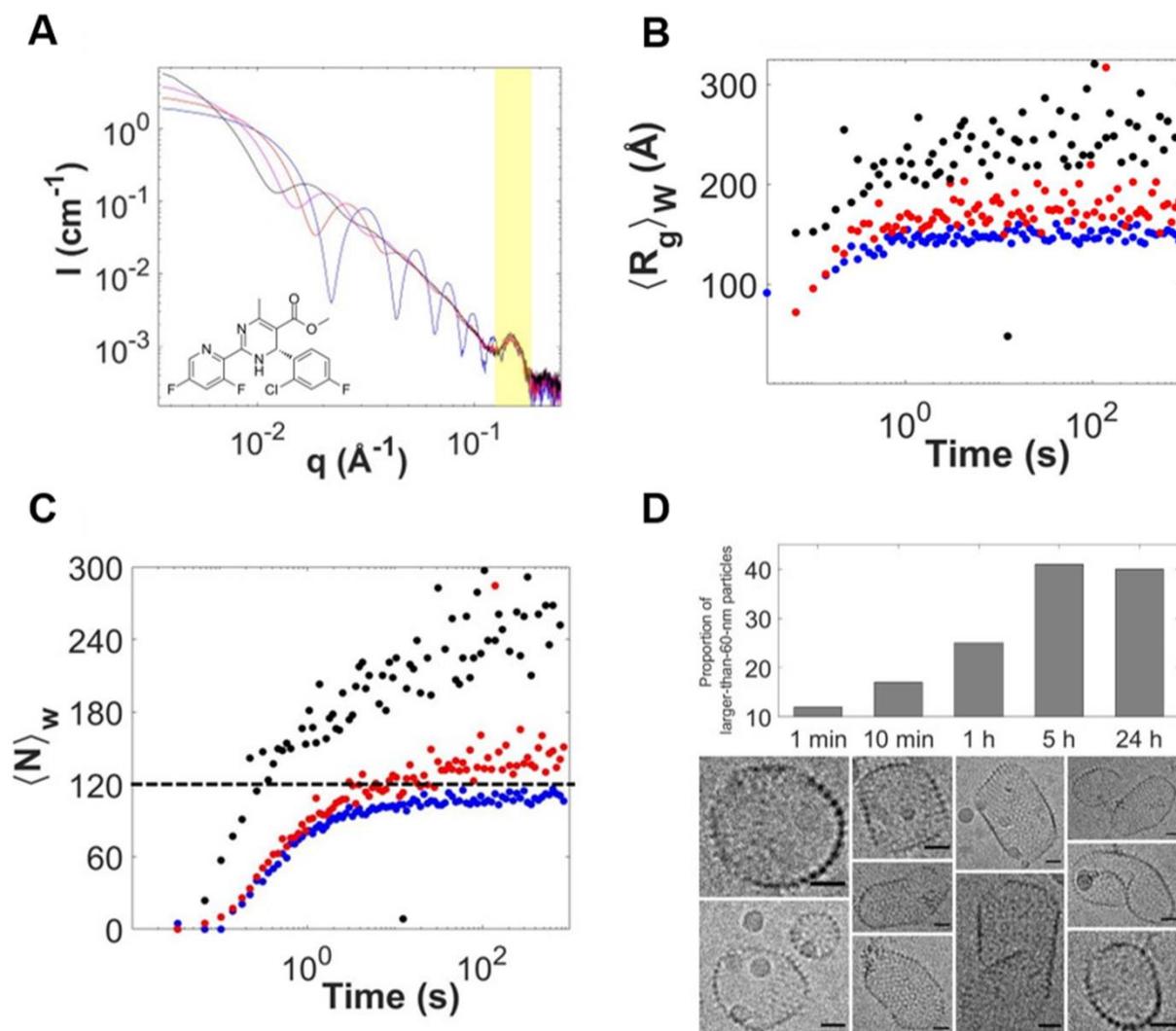
was 10% to 23% larger upon assembly in the presence of CAM-E. Moreover, the oscillations were less pronounced, and the first minimum was slightly shifted to smaller  $q$ -values. These effects increased with  $\rho$  but gradually saturated above  $\rho = 3$  (Figure S5). The differences observed on the SAXS curves show that CAM-E definitely induces structural modifications on capsids. Still, at large  $q$ -values (see the shaded area in Figure 2A), all SAXS curves were superimposed, which means that the local arrangements of subunits seem to be well preserved.

Figure 2B depicts the mean radius of gyration  $\langle R_g \rangle_w$  extracted from TR-SAXS measurements performed for capsid assembly in the presence of CAM-E at  $\rho = 0.7$  and 7.0. The equilibrium value of  $\langle R_g \rangle_w$  increased with  $\rho$  from about 150 Å in the absence of modulator to 170 Å at  $\rho = 7.0$ . Moreover, at the latter ratio, CAM-E accelerated by a factor 1.8 the capsid assembly. The extracted timescale observed for an assembly with  $\rho = 7.0$  for the CAM-E was equal to  $73 \pm 22$  ms (95% confidence interval) and the extracted timescale for an assembly in the absence of modulators was equal to 130 ms. Figure 2C shows that the equilibrium value  $\langle N \rangle_w^\infty$  was over 120 in the presence of CAM-E, which means that the final capsids were larger than the native  $T = 4$  ones. Just like in the absence of modulator,  $\langle N \rangle_w$  evolved more slowly than  $\langle R_g \rangle_w$ .

2D classification from cryoTEM images at  $\rho = 7.0$  (Figure 2D) obtained at equilibrium revealed slightly larger-than-native, ellipsoidal capsids (see Figure S6 for low-resolution 3D reconstructions). No spherical capsids were identified. By contrast, at  $\rho = 0.7$ , 75% of the particles were similar to projections of  $T = 4$  capsids formed without modulator – although the lower quality of the 3D reconstruction indicated that the icosahedral symmetry was likely less perfect –, and 25% appeared to be larger and ellipsoidal (Figure S7).

In summary, under conditions where a vast majority of  $T = 4$  capsids are assembled, CAM-E leads to slightly larger and elongated capsids, the proportion of which increases with the amount of modulator.

**CAM-A modulator had a dramatic effect on capsid morphology.**



**Figure 3. Capsid assembly in the presence of CAM-A.** (A) SAXS patterns at equilibrium with modulator-to-subunit molar ratios  $\rho$  of 0 (blue), 0.34 (red), 1.7 (magenta) and 3.4 (black). The shaded area highlights the region where all the patterns were superimposed. The chemical structure of CAM-A is represented at the bottom left. (B) Evolution of the mean radius of gyration  $\langle R_g \rangle_w$  as a function of time. (C) Plot of  $\langle N \rangle_w$  as a function of time. For (A), (B), and (C), the color code for the  $\rho$  values is the same. (D) Proportion of larger-than-60-nm particles as a function of time estimated from cryoTEM micrographs at  $\rho = 3.4$  (up). CryoTEM micrographs of capsids assembled after 24 h (bottom). Scale bars are 50 nm. In all cases, the subunit concentration was 30  $\mu\text{M}$  and  $\rho = 3.4$ .

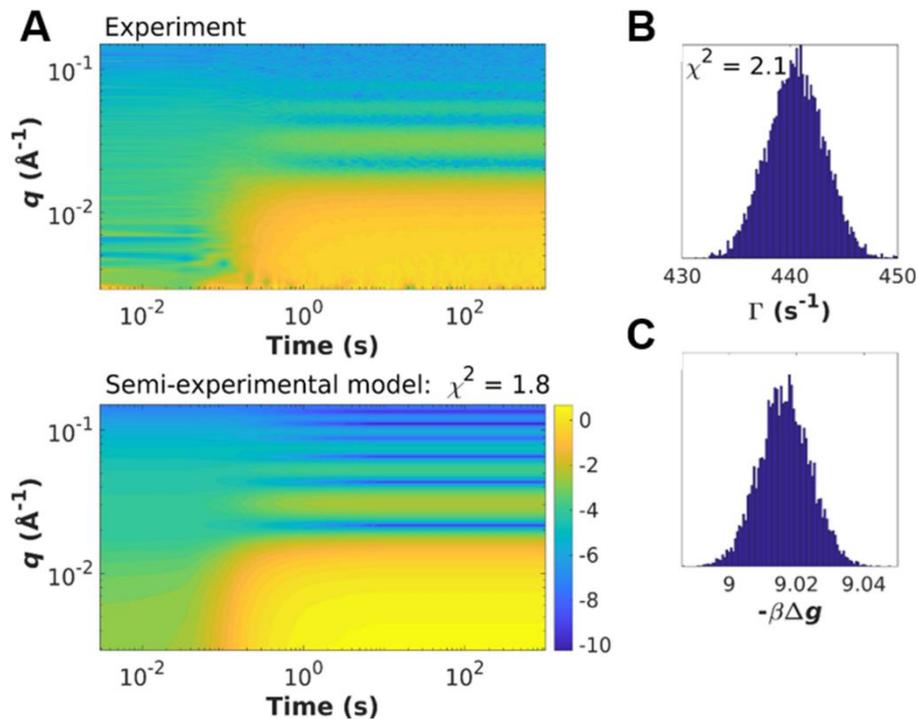
Figure 3A shows SAXS patterns at equilibrium for capsid assembly in the presence of a varying amount of CAM-A. Compared to the patterns in the absence of modulator, almost all the oscillations were strongly damped, and the first minima were shifted to much smaller values of  $q$ . The curves also showed that  $I_0$  increased with  $\rho$ , as was the case with CAM-E. Likewise, all SAXS curves were

superimposed at high  $q$ -values, even in the presence of high ratios of CAM-A (Figure 3A). Thus, the local organization of subunits were conserved in the presence of any of the two modulators, regardless of the modulator-to-subunit molar ratio.

Figures 3B and 3C depict the time evolutions of  $\langle R_g \rangle_w$  and  $\langle N \rangle_w$  during capsid assembly in the presence of CAM-A with  $\rho = 0.34$  and 3.4. Clearly, the presence of CAM-A accelerated the assembly process. At the end of the kinetics, the mean radius of gyration  $\langle R_g \rangle_w$  was about 176 Å and 251 Å, and the mean aggregation number  $\langle N \rangle_w^\infty$  was about 128 and 253, with  $\rho = 0.34$  and 3.4, respectively, which confirmed that the final capsids were larger than native  $T = 4$  capsids. These two quantities kept rising and did not seem to reach a plateau after about 15 min, indicating that the objects were still growing and/or self-organizing.

We investigated the morphologies of the objects by cryoTEM over the course of the assembly in the presence of CAM-A at  $\rho = 3.4$ . Throughout a 24-hours period, several micrographs were collected at different time points. Figure 3D gives the proportion of larger-than-60-nm particles observed at these time points. The proportion of large particles increased steadily and then stabilized after 5 h. The micrographs of Figure 3D obtained after 24 h, reveal elongated and faceted capsids, with a high shape variability, and many of them were uncapped. No spherical capsids other than a few  $T = 4$  ones were seen.

Both modulators strongly increased the subunit binding energy.



**Figure 4. Kinetic modelling of TR-SAXS data without modulator.** (A) Comparison between experimental and semi-experimental scattering intensities for capsid assembly. Intensities are in absolute units ( $\text{cm}^{-1}$ ) and plotted in logarithmic scale. (B), (C) Distribution of parameters  $\Gamma$  and  $-\beta\Delta g$  entering ‘model A’ (Eqs. 1-3) determined by Bayesian inference on experimental data. Subunit concentration was  $30 \mu\text{M}$ .

In order to get a quantitative insight into the effects of modulators on capsid assembly, we carried out kinetic modeling of our TR-SAXS measurements. Icosahedral capsid assembly can be regarded as a spherical cap growing by sequential additions of free subunits until a shell is completed. In that simple picture, it is often reasonable to consider only free subunits and complete capsids at each time point of the process, the intermediate objects being not sufficiently long-lived to accumulate in detectable amounts. Indeed, singular value decomposition (SVD) analysis confirmed that the scattering intensities could be reconstructed with only two singular values within noise level (Figure S8). The scattering intensities  $I(t, q)$  during the capsid assembly can then be reduced to:

$$I(t, q) \propto \phi(t)NP_N(q) + [1 - \phi(t)]P_1(q) \quad (1)$$

where  $\phi(t)$  denotes the fraction of subunits in capsids, while  $P_1(q)$  and  $P_N(q)$  are the form factors of subunits and assembled capsids made up of  $N$  subunits, respectively. Note that  $P_1(0) = P_N(0) =$

1. [Figure 4A](#) shows a comparison between experimental and semi-experimental scattering intensities for capsid assembly without modulator. The semi-experimental intensities were computed with the two-state assumption ([Eq. 1](#)), in which the form factors were obtained from the crystal structures of a subunit and a  $T = 4$  capsid. The semi-experimental fractions  $\phi_{\text{se}}(t)$  were determined by fitting the experimental intensities in the least-squares sense with bound constraints ( $0 \leq \phi_{\text{se}}(t) \leq 1$ ; see Methods for details). The semi-experimental model succeeded in reproducing the main features of the experimental data, which was quantitatively supported by a goodness-of-fit parameter  $\chi^2$  of 1.8.

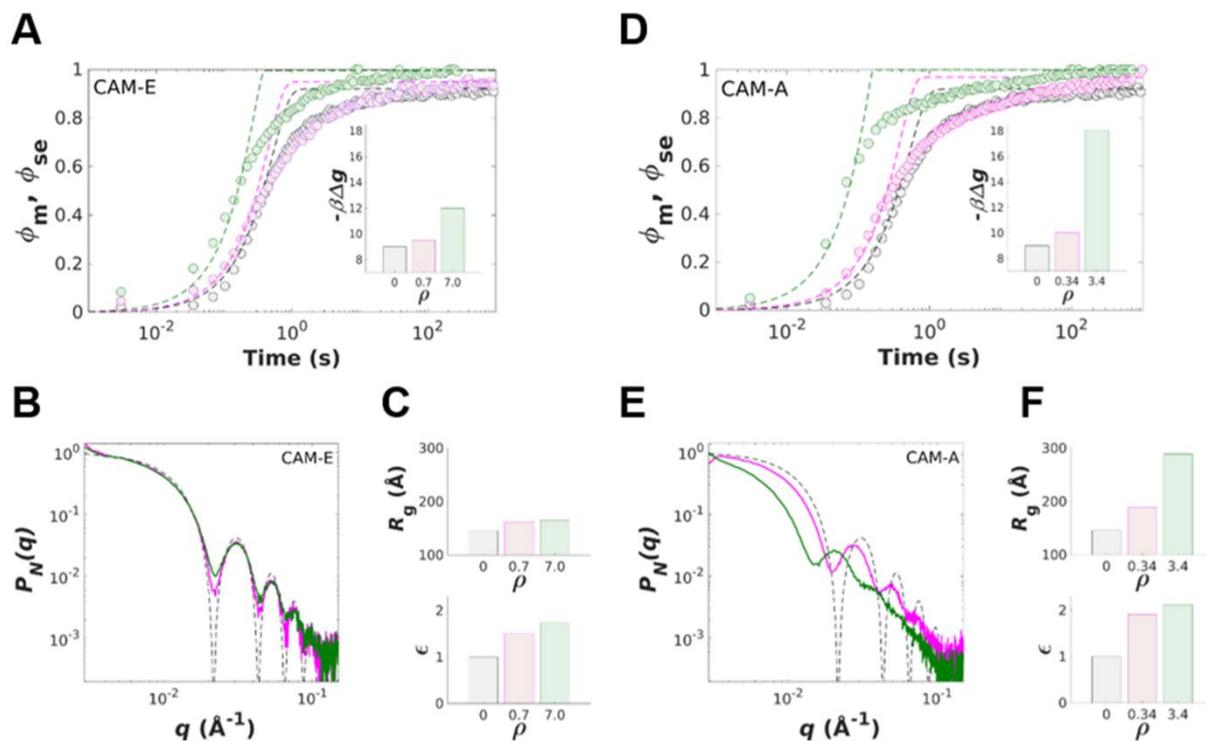
The self-assembly of empty capsids is akin to a thermodynamic phase transition, in which free subunits are suddenly brought to a nonequilibrium state after a so-called quench, and subsequently relax towards assembled capsids. In our experiments, the quench was carried out by raising the salinity and adding modulators via rapid mixing. A kinetic theory adapted to quenched capsid assembly was proposed by van der Schoot and Zandi.<sup>34</sup> The so-called ‘model A’ is based on the relaxation of a non-conserved quantity, here  $\phi$ , at a phenomenological rate  $\Gamma$  assumed constant, until an equilibrium value that minimizes the Helmholtz free energy of the system is reached. The kinetic equation reads:

$$\frac{d\phi}{dt} = -\Gamma c \left[ -\ln\left(\frac{1-\phi}{1-\phi_\infty}\right) + \frac{1}{N} \ln\left(\frac{\phi}{\phi_\infty}\right) \right] \quad (2)$$

where  $c$  stands for the dimensionless concentration of subunits before quench and  $\phi_\infty$  is the fraction of subunits in capsids at equilibrium. The latter is related to the subunit binding energy  $\Delta g$  through

$$\beta\Delta g = \ln[(1 - \phi_\infty)c] - \frac{1}{N} \ln\left(\frac{\phi_\infty}{N} c\right) \quad (3)$$

with  $1/\beta = k_B T$ ,  $k_B$  being the Boltzmann constant and  $T$  the temperature. ‘Model A’ contained solely two free parameters, *i.e.*, the reaction rate  $\Gamma$ , which was considered as an invariant of time and of  $c$ , and the subunit binding free energy  $-\beta\Delta g$  in  $k_B T$  units, which varied with the amount of modulator. [Figure 4B](#) depicts the distribution of the two parameters determined by a Bayesian approach on TR-SAXS data for capsid assembly without modulator. The form factors  $P_1(q)$  and  $P_N(q)$  ( $N = 120$ ) were calculated from the crystal structures. The mean value of  $\Gamma$  was found to be close to  $440 \text{ s}^{-1}$ , that of  $-\beta\Delta g$  was around 9.0, and the standard deviation in both cases was weak.  $\chi^2$  calculated with the mean values of  $\Gamma$  and  $-\beta\Delta g$  was 2.1, which is a good score for so minimalistic a model.

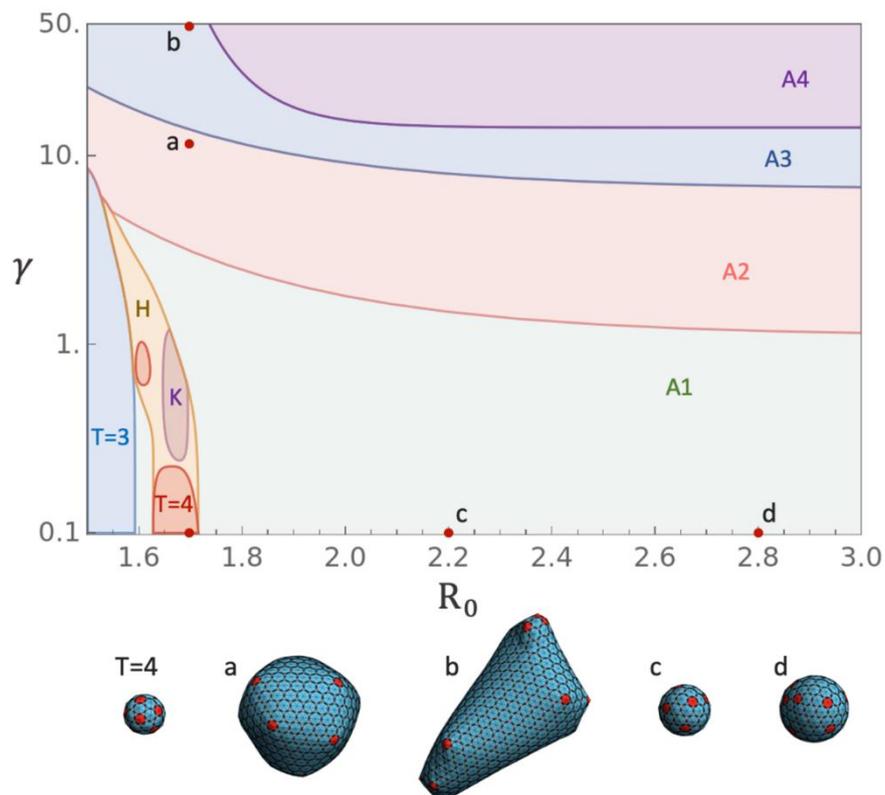


**Figure 5. Kinetic modelling of TR-SAXS data in the presence of modulators.** (A) Fraction of subunits in capsids  $\phi$  versus time for capsid assembly in the presence of CAM-E at various modulator-to-subunit molar ratios  $\rho$ : 0 (black), 0.7 (magenta) and 7.0 (green). Discs are the fractions  $\phi_{se}$  obtained with a semi-experimental model (Eq. 1; see Methods for details) while dashed lines represent the fractions  $\phi_m$  estimated through ‘model A’ (Eqs. 1-3) with  $\Gamma = 440 \text{ s}^{-1}$ . The inset gives the subunit binding energies  $-\beta\Delta g$  used in ‘model A’. (B) Form factors  $P_N(q)$  used in the kinetic model for capsid assembly with CAM-E. The dashed black line was calculated from the crystal structure of a  $T = 4$  capsid. (C) Radius of gyration  $R_g$  and axial ratio  $\epsilon$  corresponding to the form factors shown in (B). (D) Same as (A) but in the presence of CAM-A at ratios  $\rho$  of 0 (black), 0.34 (magenta) and 3.4 (green). (E) Form factors used in ‘model A’ with CAM-A. (F)  $R_g$  and  $\epsilon$  corresponding to the form factors shown in E. In all cases, the subunit concentration was  $30 \mu\text{M}$ .

In the presence of modulators, modelling was performed by fixing  $\Gamma$  to  $440 \text{ s}^{-1}$ ,  $P_1(q)$  was calculated from the crystal structure of a subunit, and only  $-\beta\Delta g$  and  $P_N(q)$  were fitting parameters. Figure 5A shows the fraction of subunits in capsids  $\phi_{se}$  and  $\phi_m$  obtained with either a semi-experimental model (discs) or ‘model A’ (dashed lines), respectively, for different ratios  $\rho$  of CAM-E. The goodness-of-fit parameter  $\chi^2$  of ‘model A’ was between 1.6 and 3.2. Like capsid assembly without modulator, the two-state assumption was supported by SVD analysis for capsid assembly in the presence of CAM-E (Figure S9). It can be seen that  $-\beta\Delta g$  increased with  $\rho$ , starting from 9.0 in the absence of modulator up to 12 for  $\rho = 7.0$ . The extracted form factors of the final capsids (Figure 5B) confirmed the formation of structures slightly larger than  $T = 4$  capsids as evidenced by the radius of gyration  $R_g$  ( $\sim 165 \text{ \AA}$  versus  $146 \text{ \AA}$  for a native capsid; Figure 5C), and more elongated as demonstrated by

the axial ratio  $\epsilon$  exceeding 1.0 (Figure 5C; see Methods for the definition of  $\epsilon$ ). These findings were consistent with the cryoTEM reconstructions described earlier. Likewise, we can see on Figure 5D that CAM-A dramatically strengthened the binding between subunits with  $-\beta\Delta g$  reaching 18 for  $\rho = 3.4$ .  $\chi^2$  ranged from 2.1 to 4.4, the latter value being obtained at  $\rho = 0.34$ . This deviation from experimental intensities was ascribed to the coexistence of more than two long-lived species in solution as revealed by SVD analysis (Figure S10). However, because 'model A' was mainly adjusted with the early stage of the assembly, during which intermediate species were still in low concentration, the estimate of the subunit binding energy should be reliable. The final capsids were much larger ( $R_g = 289 \text{ \AA}$  at  $\rho = 3.4$ ; Figures 5E,F) than in the presence of CAM-E, and they were also even more elongated, with  $\epsilon$  exceeding 2 (Figure 5F at the highest value of  $\rho$ ). Once again, these results supported well the observations made by cryoTEM (see above).

**Capsid morphologies are closely related to their elastic properties.**



**Figure 6. Phase diagram of capsid morphologies calculated by coarse-grained simulations as a function of spontaneous radius of curvature  $R_0$  and  $\gamma$ .** At small  $R_0 \approx 1.7$ , the complete shells had icosahedral symmetry with triangulation numbers  $T = 3$  and  $T = 4$ , as illustrated on the bottom of the phase diagram. The  $H$  shape surrounded the  $T = 4$  region and consisted of a rotated  $T = 4$  shell with the same number of triangular subunits (80) and the same radius. Inside the  $H$  region, we also observed the  $K$  shape, which had less triangular subunits (78) and a smaller radius. A1 to A4 denote regions with different ranges of asphericity: 0 to 0.0002 for A1; 0.0002 to 0.01 for A2; 0.01 to 0.1 for A3; and above 0.1 for A4. Below the phase diagram are representative shapes at different points  $(R_0, \gamma)$  in the phase space.

Figure 6 shows coarse-grained simulations of the complete capsid morphologies calculated as a function of the spontaneous radius of curvature  $R_0$  and  $\gamma = k_s l_0^2 / k_b$ , where  $l_0$  stands for the equilibrium bond length of each subunit, whereas  $k_s$  and  $k_b$  are the stretching and bending moduli between two subunits, respectively. Note that it is possible to map  $\gamma$  in our discrete model to the Föppl-von-Kármán number (FvK) defined as  $Y R^2 / \kappa$  in the continuum limit, with  $Y$  the Young modulus,  $R$  the radius of curvature and  $\kappa$  the bending rigidity (Figure S11). The phase diagram clearly shows that the capsids became more elongated (increasing asphericity) as  $\gamma$  was raised (see

also a more comprehensive phase diagram on [Figure S12](#)). Thus, by fixing  $R_0$  to 1.7, we could reproduce the native  $T = 4$  capsids, the larger and slightly elongated capsids obtained with CAM-E (see representative shape **a** on [Figure 6](#)) and the strongly deformed, aberrant capsids formed in the presence of CAM-A (see representative shape **b** on [Figure 6](#)) by tuning  $\gamma$  from 0.1 to 50. Increasing  $R_0$  at low  $\gamma$ -values yielded spherical shells with larger radii, but not deformed shells as observed in the presence of modulators. Therefore, increasing  $\gamma$  was a prerequisite to form aspherical capsids, and since we did not observe any spherical object larger than the native  $T = 4$  capsid, we suggest that modulators influence mostly  $\gamma$ , or equivalently, the FvK number. It should be mentioned that the values of  $\gamma$  in our discrete model were  $\gamma \simeq 1$  for  $T = 4$  and  $\gamma \simeq 5$  for  $H$  capsids, approximately corresponding to  $\text{FvK} \simeq 20$  and  $\text{FvK} \simeq 100$  in the continuum model, respectively. These values are small compared to those reported in the literature for HBV ( $\text{FvK} \simeq 400$ ) using thin-shell approximation.<sup>35</sup> The discrepancy could be due to the limitations employed in the thin-shell theory. We note that recent all-atom simulations revealed that HBV shells are very flexible<sup>36</sup>; thus, the low value of  $\gamma$  seems to be one of the essential features of HBV Core proteins, which need to be flexible and capable of a conformational switch in order to assemble to two different  $T$  numbers.

## Discussion

Chronic infections by Hepatitis B virus are a major public health burden and new drugs allowing HBV cure are being intensively developed. Among these, CAMs have grown into both promising compounds, a dozen of which are currently in clinical trials, and illuminating probes of the molecular steps of the viral cycle involving HBV Core.<sup>27,37</sup> Indeed, the recently adopted nomenclature of CAM-E and CAM-A<sup>27</sup> reflects the results of years of biochemical, biophysical and virological experiments that suggested that most CAMs paradoxically seem to promote capsid assembly, this nomenclature was also related to the theory that triggering (empty) capsid assembly at the wrong time and space was severely detrimental to the virus.<sup>38</sup> Here we show that capsids of the HBV Core NTD assembled in the presence of a prototypical CAM-E, JNJ-632, are actually structurally distinct from regular  $T = 4$  capsids formed under the same conditions. With CAM-E, they depart from icosahedral symmetry and normal stoichiometry, forming larger ( $> 120$  dimeric subunits) and ellipsoidal capsids even at a concentration where assembly is not detectably accelerated (modulator-to-subunit molar ratio  $\rho = 0.7$ ). Thus, there could be modes of actions of CAM beyond accelerating assembly even for CAM-E. All CAMs seem to bind at the same interface between Core dimers as seen by X-ray crystallography.<sup>39,40</sup> This interface is present in 240 copies in regular assembled  $T = 4$  capsids. Here our SAXS data show that local organization is not drastically changed by the CAMs in any of the assemblies formed. Thus, the number of potential CAM sites likely remains two per dimer and below

$\rho = 2$ , not all potential binding sites can be occupied by CAMs. At CAM-E  $\rho = 7.0$ , alterations of the icosahedral structure are more pronounced, but the objects assembled seem to still be closed capsids and their sizes plateau at  $\sim 160$  subunits vs. 120 in regular capsids. This is in sharp contrast to the behavior in assembly with CAM-A. Here we confirm and extend the previous reports that CAM-A induces grossly abnormal assembly (see below).

The kinetic modelling was carried out using 'model A', but in certain circumstances, capsid assembly can be described by the classical nucleation theory.<sup>41,42</sup> The latter theory is well suited to situations where the energy barrier of the formation of closed shells is high with respect to the thermal energy. This mainly occurs when the subunit concentration is near the critical aggregation concentration that itself depends on the subunit binding energy through  $c_* \sim \exp(\beta\Delta g)$ .<sup>43</sup> We assessed the classical nucleation theory and observed that the values of  $-\beta\Delta g$  required to reproduce the fractions of subunits in capsids during the early time steps, were nearly the same as those inferred from 'model A' (see [Supporting Text](#) and [Figure S13](#)). However, and quite importantly, in all cases, the resulting height of the energy barrier at the beginning of capsid assembly was lower than the thermal energy (see [Supporting Text](#)), and consequently, the classical nucleation theory could not be applied here. In the reported experiments, capsid assembly was triggered by a quench, namely, a rapid change of physicochemical conditions bringing free subunits out of equilibrium, and 'model A' was precisely worked out to account for the dynamics of phase transition in quenched systems. Notice that without modulator, we found  $-\beta\Delta g = 9.0$ , a value close to 8.2 as estimated by Asor and coworkers<sup>18</sup> under the same ionic conditions, but with a temperature of 25°C, which may account for the slight difference.

The striking effects of CAM-A cannot be solely ascribed to an enhanced interaction between subunits mediated by the modulator due to its hydrophobic nature. Although CAM-E at  $\rho = 0.7$  and CAM-A at  $\rho = 0.34$  induced similar subunit binding energies ( $-\beta\Delta g \simeq 10$ ), the capsid morphologies were significantly different, capsids in the presence of CAM-A being larger and more elongated. Additionally, assembly with CAM-A exhibited a long timescale of several hours during which the objects were still growing and most of them eventually remained uncapped. This is characteristic of a kinetic trap:<sup>43</sup> large, uncapped objects are formed within too short a time, thus depleting free subunits, which are thereby not in sufficient amount to form closed-up structures. Additional subunits must be released from small objects to allow the large ones to keep growing and become closed, a process reminiscent of the Ostwald ripening in emulsions.

At the microscopic level, we propose that the modulators have a twofold mechanism of action. Firstly, they enhance the binding energy between dimeric subunit owing to their hydrophobic nature. The higher subunit binding energy  $-\beta\Delta g$  translates into increased stretching energies at the capsid

scale, and subsequently, higher FvK numbers. A second effect may arise from their steric hindrance that can disturb the binding between adjacent subunits, since modulators sit at the surface of contact. As a result, the subunits in contact gain more flexibility to deviate around their preferred angle of curvature, thus lowering the bending rigidity of the capsid, which contributes to further increase the FvK number. It is noteworthy that, because of the slight differences between the four quasi-equivalent conformations 'A', 'B', 'C' and 'D' of Core in the  $T = 4$  capsid, there are four variations in the aforementioned dimer-dimer interface depending on whether it is A-to-A (pentameric contact), or B-to-C, C-to-D or D-to-B (three hexameric contacts). Different CAMs preferentially bind to different of these interface variations, giving a rationale as to how they may modify capsid curvature.<sup>39</sup>

We show here that JNJ-632, classified as a CAM-E, actually misdirects capsid assembly, albeit in a subtler way than the CAM-A BAY 41-4109. This has just been also reported by Lecoq and coworkers,<sup>44</sup> who showed that CAM-E, including JNJ-632, actually homogenize the four different dimer-dimer contacts, albeit to a lesser extent than CAM-A. Although negative-staining electron microscopy could not distinguish capsids reassembled with and without JNJ-632 at the mesoscopic level, solid-state nuclear magnetic resonance clearly showed that JNJ-632 modified dimer-dimer contacts at the molecular level, almost to the same extent as one CAM-A. Here we also demonstrate by TR-SAXS that JNJ-632 increases the average stoichiometry of assembly above 120 dimers, and by cryoTEM that the larger capsids are detectably ellipsoidal.

## Conclusions

The purpose of this study was to elucidate how two classes of modulators disturb the assembly pathway and alter the energetics of HBV capsids. We bring new light to several important features of CAMs: *e.g.*, some CAM-E actually can misdirect assembly; CAMs modulate, but do not prevent or grossly modify, the local arrangement of subunits; CAMs increase the subunit binding energy; and CAM binding also alters elastic energy. These modulators bind to the interface between dimers. By increasing the amount of buried hydrophobic surface area, thus raising subunit binding energy, these molecules promote capsid assembly. However, some CAMs overfill this hydrophobic pocket, causing the formation of aberrant, much longer, and distorted capsids. Here, we demonstrate that it is possible to simulate the formation of such objects by increasing the Föppl-von-Kármán number (FvK). This number is proportional to  $\gamma$ , which in turn is proportional to the ratio of stretching to bending energy. Thus, by lowering the bending energy while leaving the stretching energy unchanged (thereby increasing the FvK value), dimers could more easily deviate from their spontaneous curvature and form large, flat, non-curved structures such as those observed during assembly in the presence of type A modulators. Class E modulators, which do not lead to the

formation of this type of structure, would therefore have a less substantial impact on this bending energy. We believe that further quantitative and physical investigations are necessary to understand the mechanisms of action of modulators. Although this research alone cannot be used to design new antiviral molecules, it does enable us to carry out fingerprinting. The techniques used here provide both clarification and new insights on the mode of action of these molecules. These techniques will therefore enable us to characterize, from a structural and molecular point of view, the mode of action of newly synthesized CAMs with observable antiviral activity after screening on cellular systems. Moreover, our approach may initiate basic investigations on virus-derived, biocompatible nanocapsules exhibiting diverse morphologies tuned by yet-to-define assembly modulators, for controlled drug delivery applications for instance. The HBV Core protein can self-assemble *in vitro*. This makes it a good cargo protein engineering system for the transport and delivery of large and small molecules. The ability to chemically trigger the dissociation of virus-like particles derived from HBV Core proteins has already been demonstrated<sup>45</sup>. Therefore, the assembly of highly stable cargo capsids utilized to deliver therapeutic molecules could be regulated by these assembly modulators.

## Materials and Methods

### Sample preparation

*Expression of Cp149 capsids.* Hepatitis B virus capsids were expressed in *E. coli* as previously described.<sup>11,20</sup> Briefly, *E. coli* BL21\*Codon Plus cells were transfected with the pRSF-T7-HBc149 opt plasmid and grown at 37°C overnight on an LB-agar plate containing 50 mg.L<sup>-1</sup> of kanamycin and 34 mg.L<sup>-1</sup> of chloramphenicol. A single colony was then inoculated into 2x10 mL of LB medium and grown at 37°C overnight. After that, it was diluted into 2x500 mL of LB medium containing the same concentration of antibiotics and incubated at 37°C for 5 h. When OD<sub>600</sub> reached 0.8, the induction was done by adding 1 mM IPTG and the bacterial cells were grown at 25°C overnight.

*Capsid purification.* After the expression of Cp149 capsids, bacterial cells were pelleted by centrifugation and resuspended in lysis buffer containing 300 mM NaCl, 2 mM DTT, 50 mM Tris-HCl pH 7.5. 1 g.L<sup>-1</sup> of lysozyme, 0.5% Triton-X-100 and protease inhibitor cocktail were added to the cells and mixed on ice for 1 h. Then 4 µL of commercial Pierce nuclease were added to the solution and mixed at room temperature for 45 min. The cells were lysed by sonication and centrifuged for 45 min. The supernatant was deposited on a 10% to 60% sucrose gradient and centrifuged with a SW-32 Ti rotor (Beckman, France) at 28,800 rpm for 3 h at 4°C. The fractions containing Cp149 capsids were identified by 16% SDS-PAGE gel and precipitated by slowly adding ammonium sulfate until 40% saturation was reached. After incubation for 1 h at room temperature

and 2 h on ice, the solution was centrifuged. The pellet was resuspended in a purification buffer composed of 5% sucrose, 1 mM DTT, 50 mM Tris HCl, pH 7.5 and centrifuged to remove insoluble pellet. The supernatant containing soluble Cp149 capsids was stored at 4°C.

*Subunit purification.* The purification of Cp149 subunits was adapted from a protocol developed previously.<sup>20</sup> Cp149 capsids were dialyzed at 4°C in the disassembly buffer composed of 1 mM DTT, 50 mM CHES pH 9 and completely dissociated by adding solid urea up to 3 M. Following the dissociation, the subunits were then purified by size exclusion chromatography (SEC) using a Superdex S200 increase GL pre-equilibrated with the disassembly buffer. Fractions where dimers had been identified were dialyzed in the assembly buffer containing 500 mM NaCl, 1 mM DTT, 50 mM Tris-HCl pH 7.5. Assembled capsids were separated from dimers by size exclusion chromatography. Finally, an additional dissociation-purification step was added to remove inactive protein. Subunits could be stored at -80°C for several months without any activity loss.

### **Time-resolved and static small-angle X-ray scattering**

Prior to any measurement, Cp149 subunits were incubated for 1 h with 3 M of solid urea and dialyzed in a buffer containing 1 mM DTT, 50 mM Tris-HCl pH 7.5 for 2 h.

Time-resolved small-angle X-ray scattering experiments were carried out at the SWING beamline of the SOLEIL synchrotron facility. The sample-to-detector distance was set to 2.5 or 5.0 m which provided momentum transfers  $q$  ranging from  $2.92 \times 10^{-3}$  to  $0.463 \text{ \AA}^{-1}$  or  $1.46 \times 10^{-3}$  to  $0.232 \text{ \AA}^{-1}$ . Assembly was triggered with a stopped-flow mixer (BioLogic SFM-400) by mixing subunits, the assembly buffer composed of ammonium acetate and 50 mM Tris-HCl pH 7.5, and the modulators JNJ-632 and Bay 41-4109 at different modulators-to-subunit ratio, into a 1 mm diameter quartz capillary, at 37°C. Beam exposure time was set to 20 ms.

The samples for static small-angle X-ray scattering measurements were prepared by manually mixing subunits with 150 mM ammonium acetate in the presence of modulators, incubated at 37°C for 1 hour and stored for 2 to 3 days at 4°C. During measurements, the samples were injected into a 1.5 mm diameter quartz capillary cell using an autosampler and the sample-to-detector distance was set to 2.5 m. The temperature was maintained at 37°C with a heat bath.

The two-dimensional images were radially averaged to obtain the one-dimensional scattering profiles using Foxtrot,<sup>46</sup> and intensities were converted into absolute units after subtraction of the contribution of buffer solutions. The forward scattering intensity  $I_0 = I(q \rightarrow 0)$  and the radius of gyration  $R_g$  – or the mean (weight-averaged) radius of gyration  $\langle R_g \rangle_w$  for mixtures – were determined with AUTORG and PRIMUS from the ATSAS suite<sup>47</sup> with the condition  $q \langle R_g \rangle_w < 1.3$  defining the

Guinier region.  $I_0$  was used to estimate the mean aggregation number  $\langle N \rangle_w$  corresponding to the weight-averaged number of subunits per object:

$$\langle N \rangle_w(t) = \frac{c^S I_0(t)}{c I_0^S} \quad (4)$$

where  $I_0^S$  stands for the forward scattering intensity of subunits measured separately at a molar concentration  $c^S$ , while  $c$  is the initial molar concentration of subunits in an assembly experiment.

### Kinetic modelling

The form factors of subunits and  $T = 4$  capsids were computed with CRY SOL<sup>48</sup> from the crystal structures (Protein Data Bank code 1QGT). The subunit concentration before quench  $c$  in Eqs. 2 and 3 was taken by normalizing the subunit molar concentration to the reference state  $c_{\text{ref}} = 21.7$  mM, this value being estimated from the volume occupied by a single subunit. The semi-experimental intensities of Figure 4A were obtained by fitting Eq. 1 to experimental scattering intensities with the fraction of subunits in capsids  $\phi_{\text{se}}(t)$  as a fitting parameter for each time point, with bound constraints  $0 \leq \phi_{\text{se}}(t) \leq 1$ . The distributions of  $\Gamma$  and  $-\beta\Delta g$  in Figures 4B,C were computed with a Hamiltonian Markov chain Monte Carlo algorithm: prior distributions were set to be uniform and the likelihood was chosen as gaussian using the experimental uncertainties of the scattering intensities. The form factors were fixed to their calculated values, and at each Monte Carlo step,  $\phi(t)$  was calculated by solving Eqs. 2 and 3. Afterwards, the modelled intensities were computed from Eq. 1 and compared with the experimental intensities through the likelihood. Sampling was carried out using Matlab® with 10,000 uncorrelated Monte Carlo steps.

Throughout the text, the goodness-of-fit parameter is defined as

$$\chi^2 = \frac{1}{(N_{\text{data}} - N_{\text{param}})} \sum_{i,j} \left[ \frac{I(q_i t_j) - I_m(q_i t_j)}{\sigma_{i,j}} \right]^2 \quad (5)$$

where  $N_{\text{data}}$  is the number of data points,  $N_{\text{param}}$  the number of parameters including the data points entering the form factors,  $I_m$  the modelled scattering intensities, and  $\sigma$  the experimental uncertainties.

The subunit binding energies in the presence of modulators were estimated as follows: firstly, a value of  $-\beta\Delta g$ , and subsequently  $\phi_\infty$  through Eq. 3, was chosen. The evolution of  $\phi_m(t)$  – the index ‘m’ stands for ‘modelled’ – was calculated by solving Eq. 2,  $\Gamma$  being set to  $440 \text{ s}^{-1}$ . The traces are

plotted as dashed lines on [Figures 5A,D](#) and are referred to as ‘model A’. With the assumption that the ten last experimental scattering curves were measured near the equilibrium, their average was used to estimate the form factor of the final capsids  $P_N(q)$  via [Eq. 1](#) after plugging  $\phi_\infty$  into  $\phi(t)$ , while  $P_1(q)$  was calculated from the crystal structure of subunits using CRY SOL.<sup>48</sup>  $P_N(q)$  are plotted on [Figures 5B,E](#). The semi-experimental fractions of subunits in capsids  $\phi_{se}(t)$  were inferred from [Eq. 1](#) applied on experimental scattering intensities through a least-squares minimization with bound constraints ( $0 \leq \phi_{se}(t) \leq 1$ ) at each time point. These fractions are displayed on [Figures 5A,D](#) as discs.  $-\beta\Delta g$  was adjusted manually in such a way that  $\phi_m(t)$  and  $\phi_{se}(t)$  coincided over the short time scales, *i.e.*, typically for time points verifying  $\phi_m(t) \sim \phi_{se}(t) < 0.5$ . The number of subunits in capsid  $N$  was determined by extrapolation of  $NP_N(q)$  to the low  $q$ -values and reinjected into [Eqs. 2](#) and [3](#) in such a way that ‘model A’ was self-consistent.

The method for the determination of the radius of gyration and of the axial ratio on [Figure 5](#) was proposed by Roig-Solvas and coworkers<sup>49</sup> as an extension of the Guinier approximation. A Taylor expansion of the scattering intensity of an ellipsoid with semi-axes  $R, R, \epsilon R - \epsilon$  being the axial ratio  $\epsilon$ , about  $q = 0$ , yields

$$I(q) = I_0 \left[ 1 - \frac{1}{3} R_g^2 q^2 + \frac{1}{21} \left( R_g^4 + \frac{1}{5} A_F^2 \right) q^4 + \mathcal{O}(q^6) \right] \quad (6)$$

with  $R_g$  the radius of gyration and  $A_F$  the anisotropy factor. For a prolate ellipsoid ( $\epsilon \geq 1$ ), the axial ratio is related to the anisotropy factor through  $\epsilon = \sqrt{(1 + A_F/R_g)/(1 - A_F/(2R_g))}$ . The axial ratio is 1.0 for the spherical case and increases as the structure becomes elongated. Accordingly,  $P_N(q)$  was fitted with an 8<sup>th</sup>-order polynomial over the range  $1 \leq qR_g \leq 3$ , and both  $R_g$  and  $\epsilon$  were determined by identification of the three first polynomial coefficients.

### Cryotransmission electron microscopy and image processing

For the cryoTEM experiments, the assembly was triggered by mixing manually 30  $\mu\text{M}$  of subunits in a buffer containing 1 mM DTT, 20 mM Tris-HCl pH 7.5, in the presence or not of modulators (20 or 200  $\mu\text{M}$  – *i.e.*,  $\rho = 0.7$  or 7.0 – for CAM-E, and 100  $\mu\text{M}$  – *i.e.*,  $\rho = 3.4$  – for CAM-A) and 200 mM ammonium acetate. The latter salt concentration was found to be essential for the viability of capsids embedded in thin ice after cryofixation. Capsids assembled without modulator or with CAM-E were incubated at 37°C for 1 h and concentrated ten times to allow a faster collection of a large number of single particles during image processing. Capsids assembled with CAM-A were incubated at 37°C for 24 h with no further concentration. 4  $\mu\text{L}$  of solutions were deposited onto glow-discharged Quantifoil holey-carbon grids (R2/2 or R2/1). The grids were blotted with filter paper for 2 s before

automated plunging into liquid ethane cooled down by liquid nitrogen using FEI Vitroblot. The grids were stored in liquid nitrogen until use. The frozen samples were transferred into a Gatan 626 cryo-holder and imaging was carried out at  $-180^{\circ}\text{C}$  on a JEOL JEM-2010 microscope equipped with a 200-kV field emission gun. The samples were imaged with a magnification of  $\times 50,000$  using a minimal dose system. Images were recorded with a Gatan Ultrascan 4K CCD camera at  $2\ \mu\text{m}$  of nominal defocus.

CryoTEM particle picking, classification and reconstruction were carried out using standard protocols with the cryoSPARC software.<sup>50</sup> After CTF correction, a first 2D classification was performed using manually picked and extracted particles. Then, from the selected 2D classes, the particles were verified manually and extracted from the micrographs, and a second round of 2D classification was performed. The 3D maps were obtained after several steps of *ab initio* reconstruction and homogeneous or heterogeneous refinement jobs.

Sample	Number of micrographs	Number of particles after 2D classification	Number of particles for 3D maps
No modulator	39	3,478	3,478
CAM-E ( $\rho = 0.7$ )	42	3,912	3,912
CAM-E ( $\rho = 7.0$ )	54	3,010	3,010
CAM-A ( $\rho = 3.4$ )	95	—	—

**Table 1. Number of micrographs and single particles used for image processing.**

### Coarse-grained simulations

The assembly simulations were performed through triangulation growth model,<sup>51–53</sup> where we coarse-grained the subunits as the edges of each triangle, as shown in [Figure 7](#). Since the subunit concentration was low, at each step of the shell growth, we assumed that the subunits had enough time to find an optimal position to maximize its neighbors due to the gain of hydrophobic interaction. Meanwhile, the energy of the shell was minimized at each step, which can be written as a sum of stretching energy and bending energy

$$E_{\text{shell}} = E_s + E_b = \frac{1}{2}k_s \sum_i (l_i - l_0)^2 + k_b \sum_i [1 - \cos(\theta_i - \theta_0)] \quad (7)$$

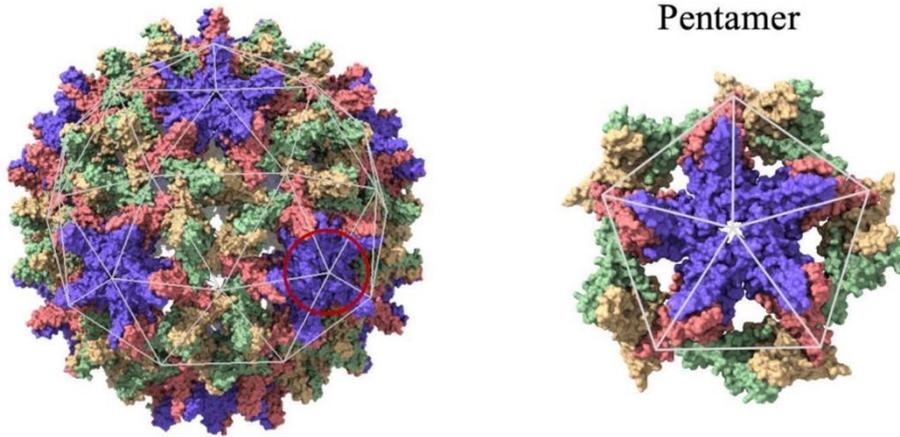
where  $k_s$  and  $k_b$  are the stretching and bending moduli,  $l_i$  and  $\theta_i$  are the length and dihedral angles of bond  $i$ , and  $l_0$ ,  $\theta_0$  are the equilibrium bond length and preferred angle. The shell energy is a discretization of continuum energy<sup>54</sup>

$$E = \frac{1}{2} \int dA \frac{Y}{1+\nu} \left( u_{ij}^2 + \frac{\nu}{1-\nu} u_{kk}^2 \right) + \int dA \left[ \frac{1}{2} \kappa (H - H_0)^2 + \kappa_g K \right] \quad (8)$$

where  $u_{ij}$ ,  $Y$  and  $\nu$  are the strain tensor, 2D Young's modulus, and Poisson's ratio, respectively.  $H$ ,  $H_0$ ,  $K$ ,  $\kappa$  and  $\kappa_g$  are mean curvature, spontaneous curvature, Gaussian curvature, bending rigidity, and Gaussian rigidity, while  $dA$  is the area element. The asphericity is defined as

$$\mathcal{A} = \frac{1}{N_v} \sum_{i=1}^{N_v} \frac{(R_i - \langle R \rangle)^2}{\langle R \rangle^2} \quad (9)$$

with  $N_v$  the vertex number,  $R_i$  the radial distance of vertex  $i$ , and  $\langle R \rangle$  the mean radius.



**Figure 7. Coarse-grained model.** Illustration of a  $T = 4$  HBV capsid along with triangular subunits used in the simulations.

## ACKNOWLEDGMENTS

We thank S. Fiulaine and R. Ruedas for their help during sample preparation and TR-SAXS experiments. K.K., S.B. and G.T. acknowledge financial support from the Agence Nationale de Recherche sur le Sida et les hépatites virales (contracts ANRS ECTZ117006 and ECTZ123887). K.K, S.B. and G.T. also acknowledge the synchrotron SOLEIL for allocating beamtime and A. Thureau for his technical assistance at the SWING beamline. The electron microscopy imaging is supported by "Investissements d'Avenir" LabEx PALM (ANR-10-LABX-0039-PALM). S.L. and R.Z. acknowledge support from NSF DMR-2131963 and the University of California Multicampus Research Programs and Initiatives (grant No. M21PR3267).

## Supporting Information

The following files are available free of charge.

Additional supporting figures, experimental details, and method (PDF)

## Author contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

**Notes**

The authors declare no competing financial interest.

## REFERENCES

1. Beasley, R. P. Hepatitis B virus. The major etiology of hepatocellular carcinoma. *Cancer* **61**, 1942–1956 (1988).
2. Taverniti, V. *et al.* Capsid Assembly Modulators as Antiviral Agents against HBV: Molecular Mechanisms and Clinical Perspectives. *J Clin Med* **11**, 1349 (2022).
3. Zlotnick, A. *et al.* Core protein: A pleiotropic keystone in the HBV lifecycle. *Antiviral Research* **121**, 82–93 (2015).
4. Birnbaum, F. & Nassal, M. Hepatitis B virus nucleocapsid assembly: primary structure requirements in the core protein. *J Virol* **64**, 3319–3330 (1990).
5. Venkatakrishnan, B. & Zlotnick, A. The Structural Biology of Hepatitis B Virus: Form and Function. *Annual Review of Virology* **3**, 429–451 (2016).
6. Nassal, M. The arginine-rich domain of the hepatitis B virus core protein is required for pregenome encapsidation and productive viral positive-strand DNA synthesis but not for virus assembly. *J Virol* **66**, 4107–4116 (1992).
7. Perlmutter, J. D. & Hagan, M. F. Mechanisms of Virus Assembly. *Annu Rev Phys Chem* **66**, 217–239 (2015).
8. Zlotnick, A., Johnson, J. M., Wingfield, P. W., Stahl, S. J. & Endres, D. A Theoretical Model Successfully Identifies Features of Hepatitis B Virus Capsid Assembly. *Biochemistry* **38**, 14644–14652 (1999).
9. Endres, D. & Zlotnick, A. Model-Based Analysis of Assembly Kinetics for Virus Capsids or Other Spherical Polymers. *Biophysical Journal* **83**, 1217–1230 (2002).
10. Oliver, R. C. *et al.* Assembly of Capsids from Hepatitis B Virus Core Protein Progresses through Highly Populated Intermediates in the Presence and Absence of RNA. *ACS Nano* **14**, 10226–10238 (2020).
11. Ceres, P. & Zlotnick, A. Weak Protein–Protein Interactions Are Sufficient To Drive Assembly of Hepatitis B Virus Capsids. *Biochemistry* **41**, 11525–11531 (2002).
12. Patterson, A., Zhao, Z., Waymire, E., Zlotnick, A. & Bothner, B. Dynamics of Hepatitis B Virus Capsid Protein Dimer Regulate Assembly through an Allosteric Network. *ACS Chem. Biol.* **15**, 2273–2280 (2020).

13. Garmann, R. F., Comas-Garcia, M., Gopal, A., Knobler, C. M. & Gelbart, W. M. The Assembly Pathway of an Icosahedral Single-Stranded RNA Virus Depends on the Strength of Inter-Subunit Attractions. *Journal of Molecular Biology* **426**, 1050–1060 (2014).
14. Narayanan, T., Wacklin, H., Konovalov, O. & Lund, R. Recent applications of synchrotron radiation and neutrons in the study of soft matter. *Crystallogr. Rev.* vol. 23 160–226 (2017).
15. Kler, S. *et al.* RNA Encapsidation by SV40-Derived Nanoparticles Follows a Rapid Two-State Mechanism. *J. Am. Chem. Soc.* **134**, 8823–8830 (2012).
16. Tresset, G. *et al.* Norovirus Capsid Proteins Self-Assemble through Biphasic Kinetics via Long-Lived State-like Intermediates. *J. Am. Chem. Soc.* **135**, 15373–15381 (2013).
17. Chevreuil, M. *et al.* Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte. *Nat Commun* **9**, 3071 (2018).
18. Asor, R., Schlicksup, C. J., Zhao, Z., Zlotnick, A. & Raviv, U. Rapidly Forming Early Intermediate Structures Dictate the Pathway of Capsid Assembly. *J. Am. Chem. Soc.* **142**, 7868–7882 (2020).
19. Law-Hine, D., Zeghal, M., Bressanelli, S., Constantin, D. & Tresset, G. Identification of a major intermediate along the self-assembly pathway of an icosahedral viral capsid by using an analytical model of a spherical patch. *Soft Matter* **12**, 6728–6736 (2016).
20. Chevreuil, M. *et al.* Nonsymmetrical dynamics of the HBV capsid assembly and disassembly evidenced by their transient species. *J. Phys. Chem. B* **124**, 9987–9995 (2020).
21. Law-Hine, D. *et al.* Reconstruction of the disassembly pathway of an icosahedral viral capsid and shape determination of two successive intermediates. *J. Phys. Chem. Lett.* **6**, 3471–3476 (2015).
22. Burke, A. *et al.* Nanoparticle-Templated Self-Assembly of Viral Capsids Probed by Time-Resolved Absorbance Spectroscopy and X-Ray Scattering. *Phys. Rev. Applied* **10**, 054065 (2018).
23. Zoulim, F. New insight on hepatitis B virus persistence from the study of intrahepatic viral cccDNA. *Journal of Hepatology* **42**, 302–308 (2005).
24. Stray, S. J. *et al.* A heteroaryldihydropyrimidine activates and can misdirect hepatitis B virus capsid assembly. *Proceedings of the National Academy of Sciences* **102**, 8138–8143 (2005).

25. Berke, J. M. *et al.* Capsid Assembly Modulators Have a Dual Mechanism of Action in Primary Human Hepatocytes Infected with Hepatitis B Virus. *Antimicrob Agents Chemother* **61**, (2017).
26. Corcuera, A. *et al.* Novel non-heteroarylpyrimidine (HAP) capsid assembly modifiers have a different mode of action from HAPs in vitro. *Antiviral Research* **158**, 135–142 (2018).
27. Zoulim, F. *et al.* Nomenclature of HBV core protein-targeting antivirals. *Nat. Rev. Gastroenterol. Hepatol.* **19**, 748–750 (2022).
28. Zhou, Z. *et al.* Heteroaryldihydropyrimidine (HAP) and Sulfamoylbenzamide (SBA) Inhibit Hepatitis B Virus Replication by Different Molecular Mechanisms. *Sci Rep* **7**, 42374 (2017).
29. Campagna, M. R. *et al.* Sulfamoylbenzamide Derivatives Inhibit the Assembly of Hepatitis B Virus Nucleocapsids. *J Virol* **87**, 6931–6942 (2013).
30. Cole, A. G. Modulators of HBV capsid assembly as an approach to treating hepatitis B virus infection. *Current Opinion in Pharmacology* **30**, 131–137 (2016).
31. Vandyck, K. *et al.* Synthesis and Evaluation of N-Phenyl-3-sulfamoyl-benzamide Derivatives as Capsid Assembly Modulators Inhibiting Hepatitis B Virus (HBV). *J. Med. Chem.* **61**, 6247–6260 (2018).
32. Weber, O. *et al.* Inhibition of human hepatitis B virus (HBV) by a novel non-nucleosidic compound in a transgenic mouse model. *Antiviral Research* **54**, 69–78 (2002).
33. Stray, S. J. & Zlotnick, A. BAY 41-4109 has multiple effects on Hepatitis B virus capsid assembly. *Journal of Molecular Recognition* **19**, 542–548 (2006).
34. van der Schoot, P. & Zandi, R. Kinetic theory of virus capsid assembly. *Phys. Biol.* **4**, 296–304 (2007).
35. Roos, W. H., Bruinsma, R. & Wuite, G. J. L. Physical virology. *Nature Phys.* **6**, 733–743 (2010).
36. Hadden, J. A. *et al.* All-atom molecular dynamics of the HBV capsid reveals insights into biological function and cryo-EM resolution limits. *eLife* **7**, e32478 (2018).
37. Lahlali, T. *et al.* Novel Potent Capsid Assembly Modulators Regulate Multiple Steps of the Hepatitis B Virus Life Cycle. *Antimicrob Agents Chemother* **62**, e00835-18 (2018).
38. Zlotnick, A. & Mukhopadhyay, S. Virus assembly, allostery and antivirals. *Trends Microbiol.* **19**, 14–23 (2011).

39. Katen, S. P., Tan, Z., Chirapu, S. R., Finn, M. G. & Zlotnick, A. Assembly-Directed Antivirals Differentially Bind Quasiequivalent Pockets to Modify Hepatitis B Virus Capsid Tertiary and Quaternary Structure. *Structure* **21**, 1406–1416 (2013).
40. Klumpp, K. *et al.* High-resolution crystal structure of a hepatitis B virus replication inhibitor bound to the viral core protein. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15196–15201 (2015).
41. Zandi, R., van der Schoot, P., Reguera, D., Kegel, W. & Reiss, H. Classical nucleation theory of virus capsids. *Biophys. J.* **90**, 1939–1948 (2006).
42. Timmermans, S. B. P. E. *et al.* The Dynamics of Viruslike Capsid Assembly and Disassembly. *J. Am. Chem. Soc.* **144**, 12608–12612 (2022).
43. Hagan, M. F. & Grason, G. M. Equilibrium mechanisms of self-limiting assembly. *Rev. Mod. Phys.* **93**, 025008 (2021).
44. Lecoq, L. *et al.* Molecular elucidation of drug-induced abnormal assemblies of the hepatitis B virus capsid protein by solid-state NMR. *Nat. Commun.* **14**, 471 (2023).
45. Starr, C. A. *et al.* Engineering Metastability into a Virus-like Particle to Enable Triggered Dissociation. *J. Am. Chem. Soc.* **145**, 2322–2331 (2023).
46. Thureau, A., Roblin, P. & Pérez, J. BioSAXS on the SWING beamline at Synchrotron SOLEIL. *J Appl Cryst* **54**, 1698–1710 (2021).
47. Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* **50**, 1212–1225 (2017).
48. Svergun, D. I., Barberato, C. & Koch, M. H. J. CRY SOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
49. Roig-Solvas, B., Brooks, D. & Makowski, L. A direct approach to estimate the anisotropy of protein structures from small-angle X-ray scattering. *J. Appl. Crystallogr.* **52**, 274–283 (2019).
50. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
51. Wagner, J. & Zandi, R. The Robust Assembly of Small Symmetric Nanoshells. *Biophys. J.* **109**, 956–965 (2015).
52. Panahandeh, S., Li, S. & Zandi, R. The equilibrium structure of self-assembled protein nano-cages. *Nanoscale* **10**, 22802–22809 (2018).

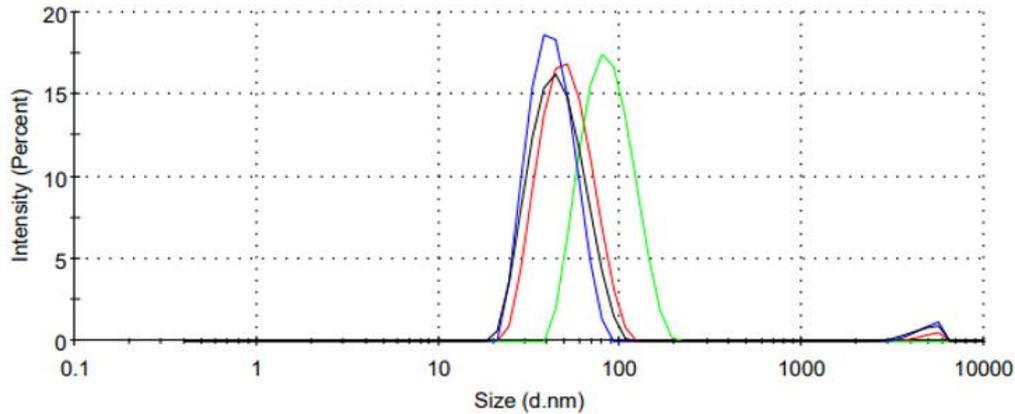
53. Panahandeh, S. *et al.* How a Virus Circumvents Energy Barriers to Form Symmetric Shells. *ACS Nano* **14**, 3170–3180 (2020).
54. Seung, H. S. & Nelson, D. R. Defects in flexible membranes with crystalline order. *Phys. Rev. A* **38**, 1005–1018 (1988).

### 3.2 Résumé de l'article

Les conclusions de cette étude fournissent de nouvelles informations sur les mécanismes d'action des modulateurs d'assemblage lors de l'assemblage de la capsid du virus de l'Hépatite B. Les analyses de diffusion des rayons X aux petits angles résolue en temps, ont permis de mettre en évidence l'accélération de la cinétique d'assemblage des capsides en présence des modulateurs d'assemblage. Ces molécules provoquent également une augmentation de l'énergie de liaison des sous-unités. De plus, ces analyses mais également l'analyse de la morphologie des capsides observées par cryomicroscopie électronique ont mis en évidence les différents changements morphologiques et structuraux induits par la CAM-A mais également par la CAM-E. Ces changements vont d'une légère élongation (dans le cas de la CAM-E) à une déformation significative accompagnée d'une forte augmentation de la taille des objets (dans le cas de la CAM-A). Enfin, ces morphologies ont été reproduites lors de simulation gros grains, mettant en évidence le rôle des modulateurs sur l'énergie élastique de la capsid. La découverte majeure de cette étude est la capacité de la CAM-E, JNJ-632 à induire des modifications structurales des capsides lorsqu'elle est présente lors de l'assemblage. Et cela, malgré le fait que les CAM-E soient classées comme des modulateurs promouvant l'assemblage de capsides normales

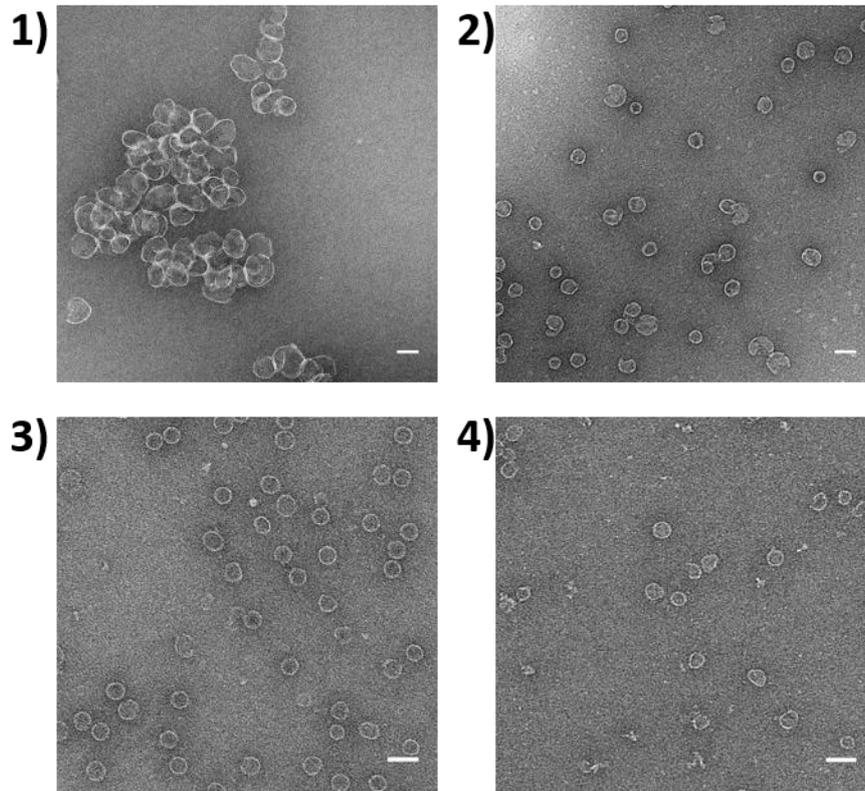
### 3.3 Complément de l'article : Caractérisation par DLS et microscopie électronique des capsides formées en présence des modulateurs de type A et de type E

En amont des expériences au synchrotron et des expériences de cryomicroscopie électronique, nous avons voulu nous assurer qu'il était bien possible d'assembler *in vitro*, des objets en présence des différents modulateurs. Dans cette optique, les objets formés après assemblage en présence des modulateurs ont donc été préalablement analysés par DLS et microscopie électronique.



**Figure 4.10 : Mesures par DLS des objets formés après l'assemblage de la Cp149 en présence des modulateurs à différentes concentrations.** 30  $\mu\text{M}$  de dimères de Cp149 ont été assemblés en présence de 10  $\mu\text{M}$  (ligne rouge) ou 100  $\mu\text{M}$  (ligne verte) de CAM-A, de 20  $\mu\text{M}$  (ligne bleue) ou 200  $\mu\text{M}$  (ligne noire) de CAM-E.

30  $\mu\text{M}$  de dimères de Cp149 ont été assemblés sans modulateurs ou en présence de 10/100  $\mu\text{M}$  de CAM-A ou 20/200  $\mu\text{M}$  de CAM-E à 37°C puis analysés par DLS. La figure 4.10 ci-dessus, indique le diamètre hydrodynamique des objets mesurés en fonction des conditions d'assemblage. Nous observons que la présence de la CAM-A (courbes rouge et verte) entraîne une augmentation du diamètre hydrodynamique des objets formés. En absence de modulateurs, le diamètre hydrodynamique est égal à 37 nm, tandis qu'en présence de 10  $\mu\text{M}$  de CAM-A, celui-ci est de 48 nm, et en présence de 100  $\mu\text{M}$  de CAM-A, il atteint 80 nm. En revanche, en présence de la CAM-E (représentée par les courbes bleue et noire), le diamètre hydrodynamique des objets formés augmente légèrement mais reste similaire à celui des capsides normales du VHB. Le diamètre hydrodynamique est égal à 42 nm et 44 nm respectivement lors d'un assemblage en présence de 20  $\mu\text{M}$  et 200  $\mu\text{M}$  de CAM-E.

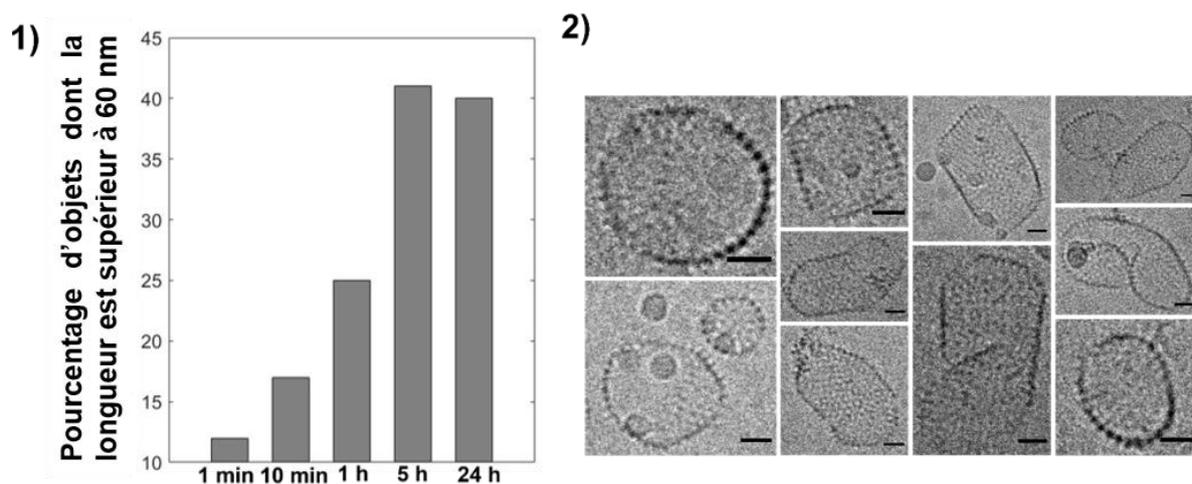


**Figure 4.11 : Micrographies des objets formés lors de l'assemblage de la Cp149 à partir des dimères de la Cp149 et en présence de modulateurs à différentes concentrations.** Objets formés lors de l'assemblage en présence de 100  $\mu\text{M}$  de CAM-A (1), 10  $\mu\text{M}$  de CAM-A (2), 200  $\mu\text{M}$  de CAM-E (3), et 20  $\mu\text{M}$  de CAM-E (4). Les barres d'échelle sont égales à 50 nm. Les échantillons sont observés sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN Ultrascan 4K CCD.

Ensuite, 22,5  $\mu\text{M}$  (0,75 mg/ml) de dimères de Cp149 ont été assemblés en présence de 10/100  $\mu\text{M}$  de CAM-A ou 20/200  $\mu\text{M}$  de CAM-E à 37°C. La concentration en protéines a été ajustée à 0,2 mg/ml avant d'observer les objets formés en microscopie électronique (Figure 4.11). L'image de microscopie électronique des capsides formées en présence de 100  $\mu\text{M}$  de CAM-A (1), montre la présence de gros objets, ouverts, qui ne présentent pas la morphologie typique d'une capside du VHB. De plus, ces objets sont agrégés. Cette agrégation peut être due à l'utilisation du colorant qui déshydrate l'échantillon ou peut être un réel effet du modulateur. L'image de microscopie électronique des capsides formées en présence de 10  $\mu\text{M}$  de CAM-A (2) montre la présence de deux types de populations. Nous pouvons observer des objets ressemblant aux capsides du VHB mais également un autre type d'objet beaucoup plus gros, ayant une morphologie complètement différente et dont certains sont ouverts. En présence de la CAM-E (3,4), les images de microscopie électronique

montrent la présence de capsides qui ressemblent aux capsides classiques du VHB mais également de capsides légèrement plus grandes. Enfin, en présence de 200  $\mu\text{M}$  de CAM-E (4), des d'objets ouverts ou cassés sont observés.

Ces premières expériences ont donc montré qu'il est possible d'assembler des capsides *in vitro* en présence de ces modulateurs. Ces résultats nous ont permis de réaliser par la suite des études structurales et cinétiques de l'assemblage des capsides du VHB en présence des modulateurs.



**Figure 4.12 : Assemblage de la capside Cp149 en présence de la CAM-A pendant 24 h.** 1) Evolution du pourcentage d'objets de taille supérieure à 60 nm en fonction de la durée d'assemblage en présence de 100  $\mu\text{M}$  de CAM-A. 2) Micrographies de cryoEM des objets formés en présence de 100  $\mu\text{M}$  de CAM-A après 24 h d'assemblage à 37°C. Les barres d'échelle sont égales à 15 nm. Les échantillons sont observés sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN Ultrascan 4K CCD.

Les données obtenues à la suite des expériences de cinétique d'assemblage des objets formés à partir de 30  $\mu\text{M}$  de dimères de Cp149 en présence de 100  $\mu\text{M}$  de CAM-A montrent que la taille des objets ne se stabilisent pas après 15 minutes d'assemblage à 37°C et continuent d'augmenter. Une cinétique de cet l'assemblage dans les mêmes conditions a donc été réalisée pendant 24 h et des grilles de microscopie électronique sur lesquelles les objets sont déposés ont été congelées à différents temps d'incubation. L'histogramme de la figure 4.12.1 représente l'évolution du pourcentage d'objets dont la taille est supérieure à 60 nm au cours du temps. Il montre que les objets continuent d'évoluer pendant 5 h avant de se stabiliser. La figure 4.12.2 montre la polydispersité des objets après 24h d'assemblage. Ces résultats

montrent que l'assemblage des objets formés en présence de la CAM-A est un processus très long et qu'il conduit à la formation d'objets très divers.

## 4. Assemblage de la protéine cp149 en présence de deux nouvelles molécules GLP-26 et RS-3427

### 4.1 Introduction

Dans le cadre d'une collaboration avec l'équipe du Professeur Raymond Schinazi, nous avons entrepris une étude sur la caractérisation de deux nouvelles molécules GLP-26 et RS-3427.

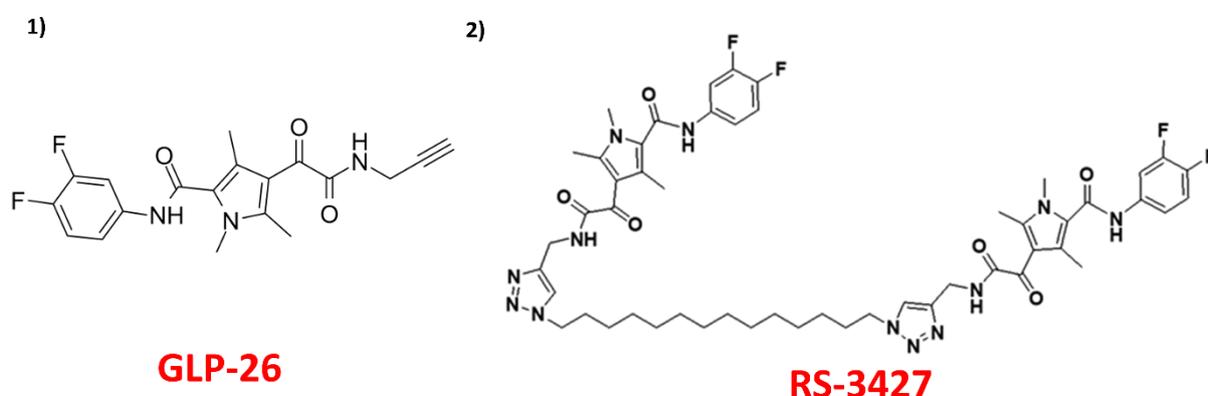


Figure 4.13 : Structures chimiques de GLP-26 (1) et de RS-3427 (2).

Ces deux nouveaux composés, dont les structures chimiques sont représentées sur la figure 4.13 ont été identifiés comme candidats potentiels pour interagir avec la Cp149 et influencer son assemblage. Ils appartiennent à une nouvelle classe de CAMs, celle des dérivés de glyoxamoylpyrroloxamide (GLP). Cette classe se caractérise par la substitution du groupement sulfonamide présent dans les SPA par un groupement glyoxamide.

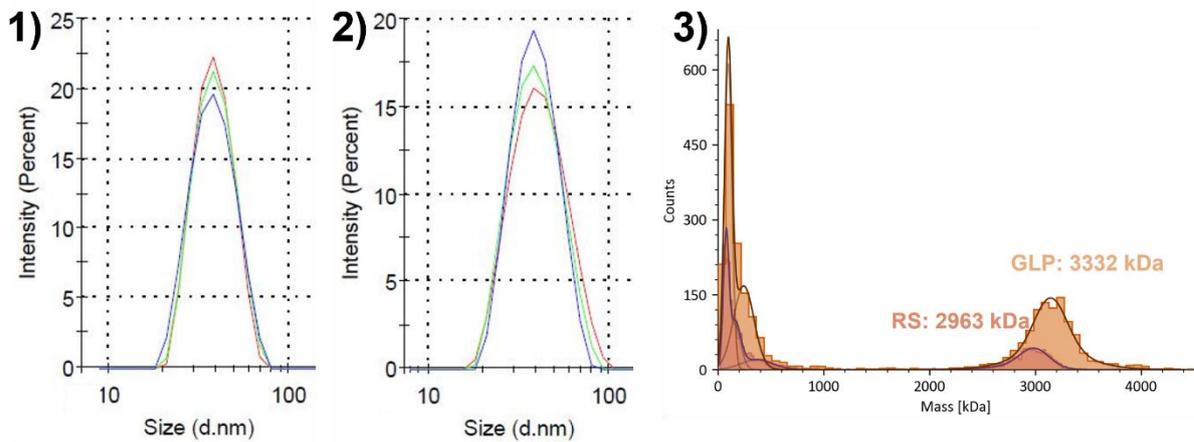
Des études *in vivo* ont révélé que GLP-26 réduit de manière significative la quantité d'ADNccc du VHB, jusqu'à 90% lorsqu'elle est utilisée à une concentration de 1  $\mu\text{M}$  sur les cellules HepAD38, tout en déformant les capsides déjà formées. De plus, des expériences menées sur des souris ont montré que cette molécule associée à un autre médicament antiviral, l'entécavir, présente une bonne activité antivirale, sans effets nocifs sur les fonctions cardiaques, même à des doses élevées. D'autres études, réalisées sur des souris humanisées et infectées par le VHB ont montré que GLP-26

réduit les niveaux d'ADN viral, d'antigènes HBs et d'antigènes HBe. Enfin, cette molécule, qui est spécifique du VHB, peut conduire à la formation de capsides sphériques, mais malformées lors de simulations de dynamique moléculaire.

La revue de littérature montre que la molécule GLP-26 a principalement été caractérisée sur le plan cellulaire jusqu'à présent. Peu d'études structurales ou *in vitro* ont à notre connaissance été réalisées. De plus, la molécule RS-3427, dérivée de GLP-26 n'a pas encore fait l'objet d'études approfondies. Ainsi, l'objectif de notre étude était de déterminer comment ces molécules modulent la structure des capsides du virus de l'Hépatite B, en apportant si possible des informations sur leur mode de fixation à la protéine Core ainsi que de caractériser leurs effets sur la cinétique d'assemblage.

Dans cette optique, nous avons réalisé des études cinétiques par TR-SAXS et des études structurales par cryomicroscopie électronique, des capsides du VHB formées à partir de dimères de la protéine Cp149 en présence de ces nouvelles molécules.

## 4.2 *In vitro*, l'assemblage en présence des deux molécules conduit à la formation d'une grande majorité de capsides T=4

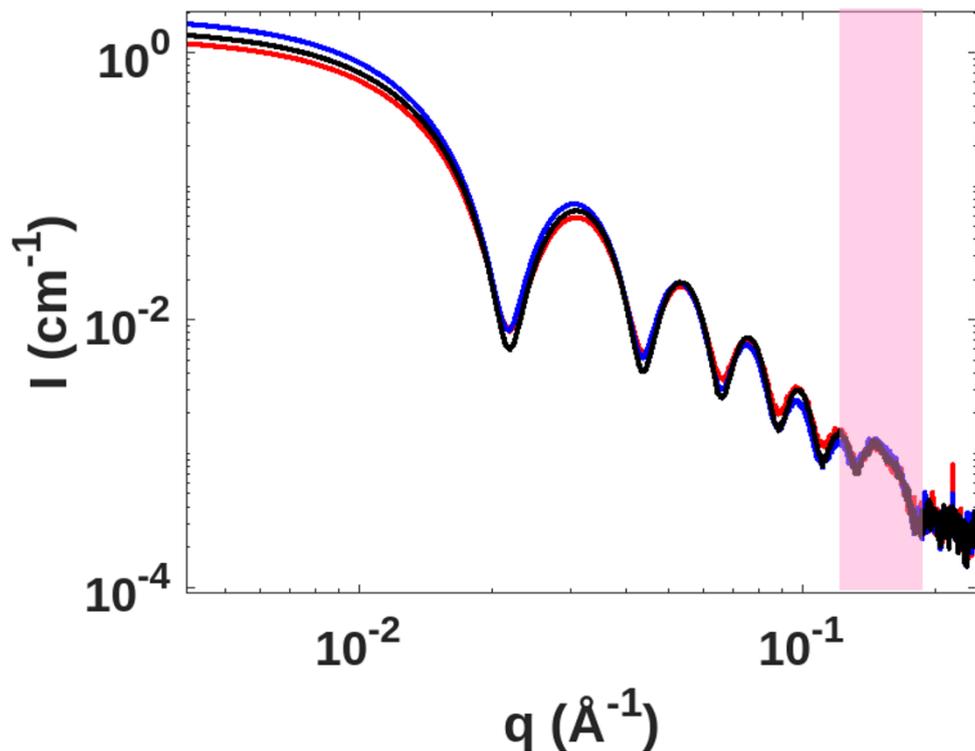


**Figure 4.14 : Caractérisation optique des objets formés en présence de GLP-26 et RS-3427.** (1) et (2) Mesures par DLS des objets formés lors de l'assemblage de 30  $\mu\text{M}$  en présence de 90  $\mu\text{M}$  de GLP-26 (1) et 90  $\mu\text{M}$  de RS-3527 (2). Chaque échantillon a été mesuré 3 fois. (3) Mesure par photométrie de la distribution de masse des objets formés dans les mêmes conditions que pour (1) et (2).

30  $\mu\text{M}$  de sous-unités de Cp149 ont été assemblés en présence de 90  $\mu\text{M}$  de RS-3427 ou GLP-26 et 150 mM d'acétate d'ammonium pendant 1 h à 4°C. L'objectif initial de l'étude était d'analyser la structure des objets formés. Pour ce faire, nous avons décidé de réaliser les assemblages à 4°C au lieu de 37°C. La figure 4.14 ci-dessus indique les résultats obtenus par DLS des mesures des objets formés dans ces conditions. L'analyse par DLS (1,2), indique que les objets formés en présence de ces deux modulateurs sont similaires aux capsides T=4 du VHB. Le diamètre hydrodynamique des objets formés en présence de ces modulateurs est similaire à celui des capsides assemblées dans des conditions classiques c'est-à-dire sans modulateurs. Les diamètres hydrodynamiques sont égaux à 37 et 38 nm respectivement pour l'assemblage avec 90  $\mu\text{M}$  de GLP-26 ou 90  $\mu\text{M}$  de RS-3427. De plus, un seul pic est observé, indiquant que l'échantillon est homogène dans les deux conditions. Cette homogénéité est confirmée par la très faible valeur de l'indice de polydispersité qui est d'environ 0,1 pour les deux échantillons.

Ces objets ont ensuite été analysés par photométrie de masse (3), afin d'obtenir des informations sur leur masse exacte. Pour cela, un peu moins de 2 minutes avant la mesure, une dilution des échantillons a été réalisée. Les échantillons ont ainsi été dilués de façon à obtenir une concentration finale de 6 nM de capsides. Nous supposons que l'ensemble des dimères se sont assemblés en capsides, ce qui signifie donc que la concentration initiale en capsidite est égale à 247 nM. En conséquence, les échantillons ont été dilués environ 41 fois à l'aide d'un tampon contenant 50 mM Tris HCl pH 7,5 et 150 mM AmAc. La figure ci-dessus indique la masse des différents objets mesurés dans chaque condition d'assemblage. La masse mesurée pour des capsides de Cp149 formées en condition classique (150 mM AmAc) était égale à 3606 kDa. En présence de 90  $\mu$ M de GLP-26, la masse moyenne des capsides est mesurée à 3332 kDa et en présence de 90  $\mu$ M de RS-3427, elle est égale à 2963 kDa. Nous remarquons que la masse moyenne des objets est légèrement différente en fonction des conditions d'assemblage, mais reste globalement similaire à celle des capsides mesurées après un assemblage classique.

Ces premiers résultats indiquent que les molécules RS-3427 et GLP-26 seraient des modulateurs de type CAM-E puisqu'ils conduisent à la formation d'objets qui comme JNJ-632, d'un point de vue caractérisation optique, sont similaires aux capsides du VHB.

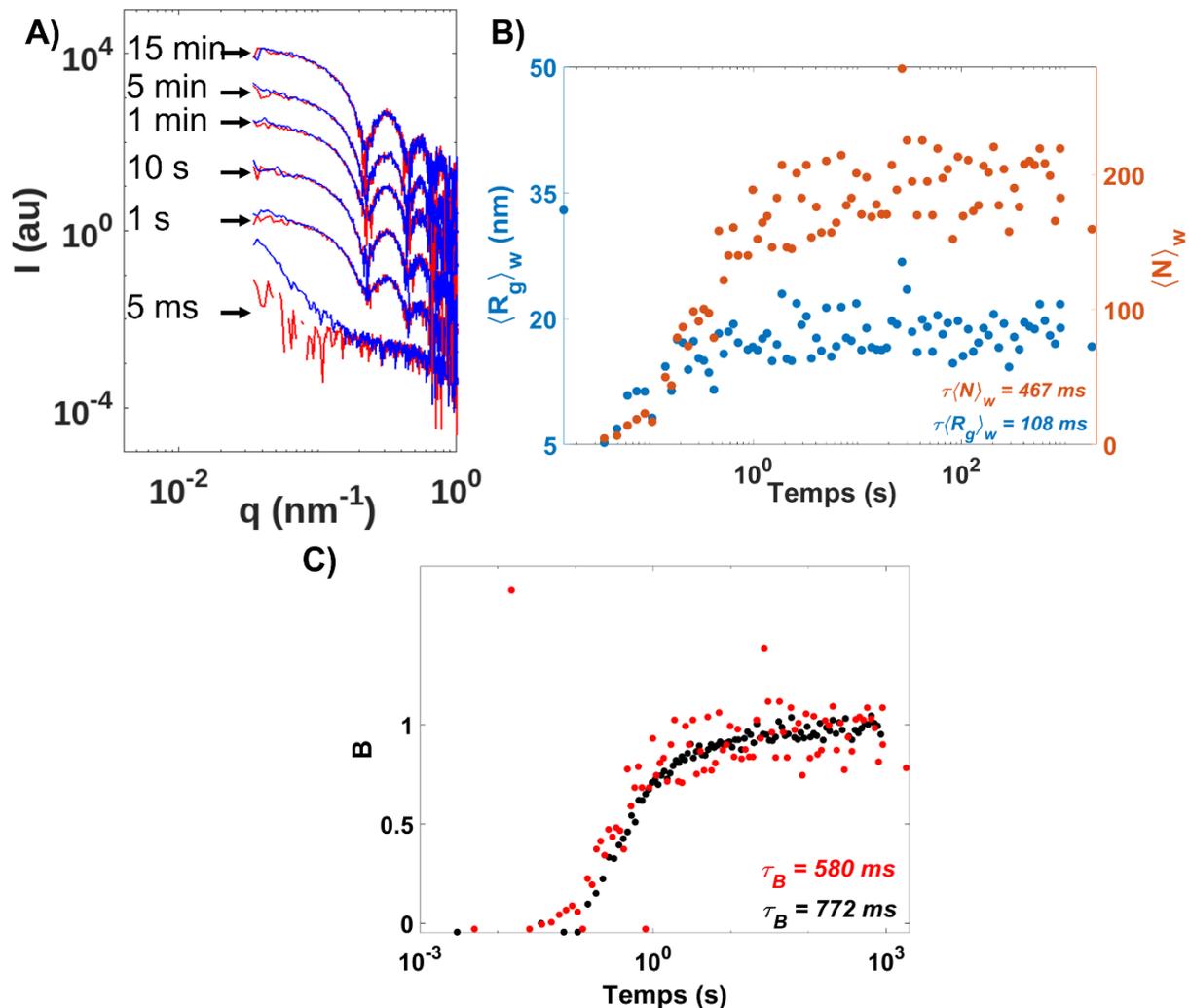


**Figure 4.15 : Courbes SAXS des capsides assemblées en présence des deux molécules.** Courbes SAXS expérimentales des capsides formées avec 30  $\mu\text{M}$  de sous-unités sans modulateurs (courbe noire), avec 90  $\mu\text{M}$  de GLP-26 (courbe bleue) et avec 90  $\mu\text{M}$  de RS-3427 (courbe rouge). L'assemblage a été déclenché avec 500 mM d'acétate d'ammonium à 4°C. La zone marquée en rose, met en évidence la région où toutes les courbes SAXS sont superposées.

Les courbes SAXS des capsides formées en présence des modulateurs GLP-3427 et RS-26, et mesurées à l'équilibre (Figure 4.15), indiquent également que ces deux molécules seraient bien des modulateurs de type CAM-E. Dans les deux conditions d'assemblages, la première oscillation apparaît à un nombre d'onde  $q = 0,022 \text{ \AA}^{-1}$ , ce qui signifie que la majorité des objets formés en présence de ces deux modulateurs sont des capsides T=4. En présence de 90  $\mu\text{M}$  GLP-26 lors de l'assemblage, la courbe SAXS obtenue (courbe bleue) est similaire à la courbe SAXS des capsides formées en conditions classiques (courbe noire). Cependant, l'intensité  $I(0)$  est légèrement plus grande et les oscillations sont moins profondes. Certains objets formés dans cette condition sont donc légèrement plus gros que des capsides T=4 ou présenteraient des variations structurales par rapport aux objets formés dans des conditions d'assemblage classique. En présence de 90  $\mu\text{M}$  RS-3427 lors de l'assemblage, la courbe SAXS obtenue (courbe rouge), est également similaire à la courbe SAXS des capsides formées en conditions normales. Mais l'intensité  $I(0)$  est légèrement plus basse et les oscillations sont également moins profondes. Dans cette condition, il

semblerait que moins de capsides soient formées. Enfin, toutes les courbes SAXS se superposent aux grandes valeurs de  $q$  (zone rose), ce qui indique que, tout comme avec les modulateurs Bay 41-4109 et JNJ-632, la présence des molécules GLP-26 et RS-3427 lors de l'assemblage n'affecte pas l'organisation locale des objets formés.

### 4.3 GLP-26 accélère la cinétique d'assemblage et induit des changements structuraux



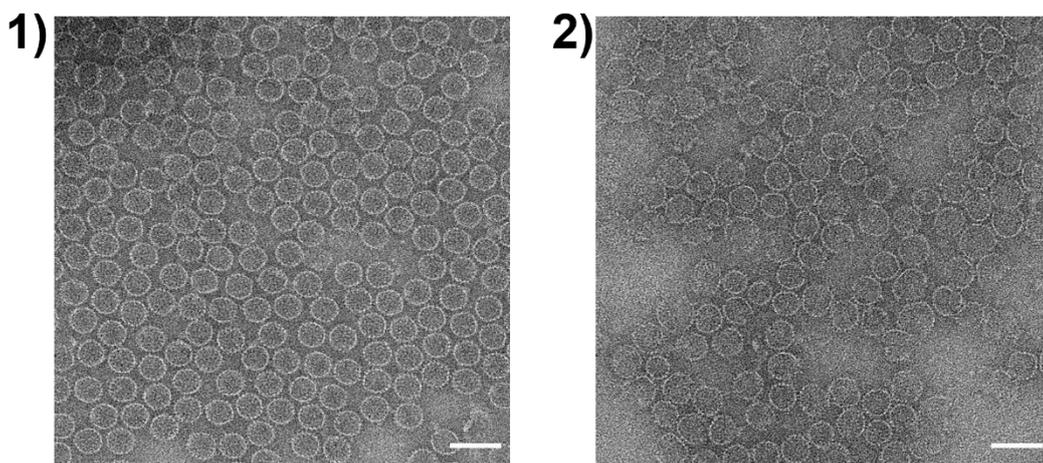
**Figure 4.16 : Mesures TR-SAXS de l'assemblage de capsides Cp149 en présence de la molécule GLP-26.** (A) Courbes TR-SAXS de deux expériences identiques d'assemblage (courbes bleue et rouge) de  $30 \mu\text{M}$  de sous-unités en présence de  $90 \mu\text{M}$  de GLP-26. (B) Evolution du nombre moyen d'agrégation  $\langle N \rangle_w$  (disques orange) et du rayon de giration moyen  $\langle R_g \rangle_w$  (disques bleus) en fonction du temps pour l'assemblage de capsides à partir de  $30 \mu\text{M}$  de sous-unités en présence de  $90 \mu\text{M}$  de GLP-26. (C) Evolution du facteur  $B$  défini comme  $B = \frac{I(0)_t - I(0)_i}{I(0)_f - I(0)_i}$  en fonction du temps lors d'un assemblage sans modulateurs (disques noirs) et en présence de  $90 \mu\text{M}$  de GLP-26 (disques rouges).

$I(0)_t$ ,  $\langle I(0)_i \rangle$ , et  $\langle I(0)_f \rangle$ , correspondent respectivement à l'intensité  $I(0)$  au temps  $t$ , à l'intensité moyenne  $I(0)$  des deux premières courbes de la cinétique et à l'intensité moyenne  $I(0)$  des cinq dernières courbes de la cinétique. Pour (B) et (C), les échelles de temps caractéristiques ont été obtenues en ajustant une fonction de décroissance exponentielle.

30  $\mu\text{M}$  de dimères de Cp149 ont été assemblés avec 150 mM d'acétate d'ammonium et 90  $\mu\text{M}$  de GLP-26 à 37°C. Les courbes SAXS des cinétiques de formation des capsides dans cette condition sont représentées ci-dessus (Figure 4.16.a). En présence de GLP-26, des objets se forment et une première oscillation à  $q = 0,023 \text{ \AA}^{-1}$  est observée. La valeur du nombre d'onde  $q$  de cette première oscillation est identique à celle observée lors de la formation de capsides T=4 durant un assemblage sans modulateurs. Cela indique donc que la majorité des objets formés dans cette condition sont des capsides T=4.

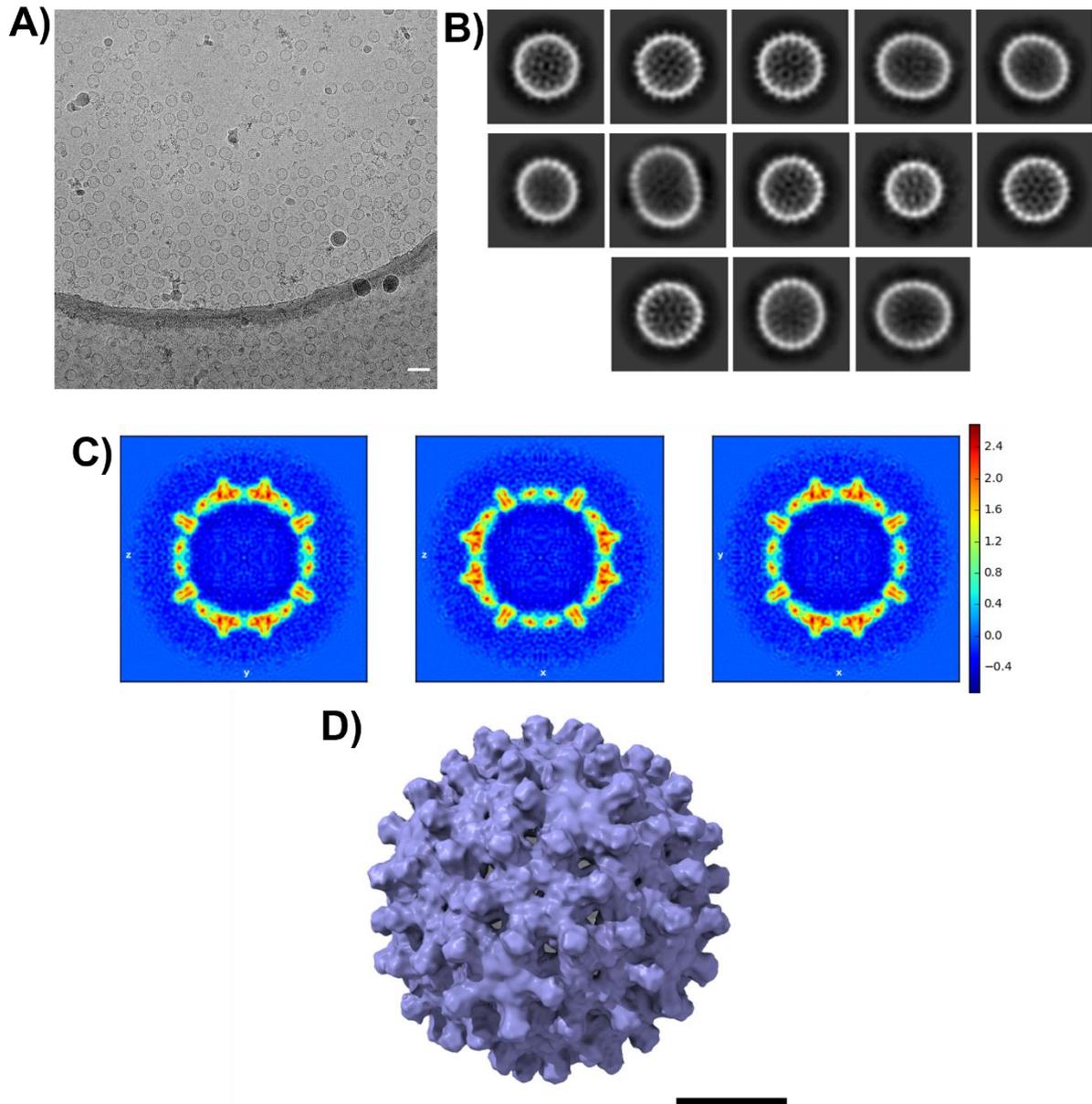
La figure 4.16.b ci-dessus représente l'évolution des  $\langle R_g \rangle_w$  et des  $\langle N \rangle_w$  lors de l'assemblage de capsides formées à partir de 30  $\mu\text{M}$  de sous-unités et 90  $\mu\text{M}$  de GLP-26. Les données fournies ici sont comparées à la cinétique d'assemblage de 30  $\mu\text{M}$  de dimères de Cp149 réalisée au synchrotron SOLEIL, et présentée dans l'article ci-dessus. En effet, les données obtenues pour cette même cinétique mais réalisée à l'ESRF en parallèle de la cinétique avec 90  $\mu\text{M}$  de GLP-26 ne sont pas concluantes. Le  $\langle R_g \rangle_w$  est surestimé (176  $\text{\AA}$  pour la cinétique réalisée à l'ESRF au lieu de 150  $\text{\AA}$  pour la cinétique réalisée à SOLEIL) et l'échelle de temps associée également (237 ms pour la cinétique à l'ESRF au lieu de  $130 \pm 22$  ms pour la cinétique à SOLEIL). En présence de 90  $\mu\text{M}$  de GLP-26 lors de l'assemblage, le  $\langle R_g \rangle_w$  est égal à 180  $\text{\AA}$  à la fin de la cinétique, et l'échelle de temps caractéristique extraite pour ce paramètre est égale à  $108 \pm 66$  ms. Les valeurs du  $\langle N \rangle_w$  calculées pour cette condition d'assemblage sont supérieures à celles obtenues lors d'un assemblage classique, et l'échelle de temps associée à ce paramètre est égale à  $467 \pm 178$  ms (En absence de modulateurs, cette échelle de temps est égale à  $717 \pm 54$  ms). Enfin, la valeur du nombre moyen d'agrégation à l'équilibre  $\langle N \rangle_w^\infty$  est supérieure à 120. Elle est égale à 190 à la fin de la cinétique. Ces résultats indiquent que les objets formés dans cette condition sont plus gros que des capsides T=4 et que des différences structurales peuvent être observés lors d'un assemblage en présence de cette molécule. Par ailleurs, il est intéressant de noter que la valeur de  $\langle N \rangle_w$  évolue plus lentement que la valeur du  $\langle R_g \rangle_w$  lors de la cinétique d'assemblage.

Pour comparer les vitesses d'assemblage, la figure 4.16.c illustre l'évolution du facteur B au cours du temps lors de la cinétique d'assemblage, en présence ou non de la molécule GLP-26. Ce facteur B représente l'intensité  $I(0)$  calculée et normalisée pour chaque courbe de la cinétique. L'échelle de temps caractéristique associée à ce paramètre est égale à  $580 \pm 229$  ms en présence de  $90 \mu\text{M}$  de GLP-26 et à  $772 \pm 64$  ms en absence de modulateurs. Les échelles de temps extraites pour le rayon de giration et le facteur B lors de l'assemblage en présence de  $90 \mu\text{M}$  de GLP-26 sont similaires à celles obtenues pour ces mêmes paramètres lors d'un assemblage en condition classique. Cependant, l'échelle de temps calculée pour le nombre d'agrégation moyen est plus courte en présence de  $90 \mu\text{M}$  de GLP-26. Par conséquent, nous pouvons en conclure que cette molécule accélère la cinétique d'assemblage.



**Figure 4.17 : Micrographies des capsides assemblées en présence ou non de GLP-26 et observées en microscopie électronique en coloration négative.**  $30 \mu\text{M}$  de capsides ont été assemblées en absence de modulateurs (1) ou en présence de  $90 \mu\text{M}$  de GLP-26 (2). Les barres d'échelle sont égales à  $50 \text{ nm}$ . Les échantillons sont observés sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN Ultrascan 4K CCD.

La figure 4.17 ci-dessus montre les images de microscopie électronique des capsides de Cp149 assemblées à partir de  $30 \mu\text{M}$  de sous-unités en présence ou non de  $90 \mu\text{M}$  de GLP-26. Dans les deux conditions, les objets formés présentent des formes et des tailles similaires à celle de la capside T=4 du VHB. Cependant, en présence de  $90 \mu\text{M}$  de GLP-26, nous pouvons observer des capsides légèrement plus grandes mais également la présence de capsides ouvertes.

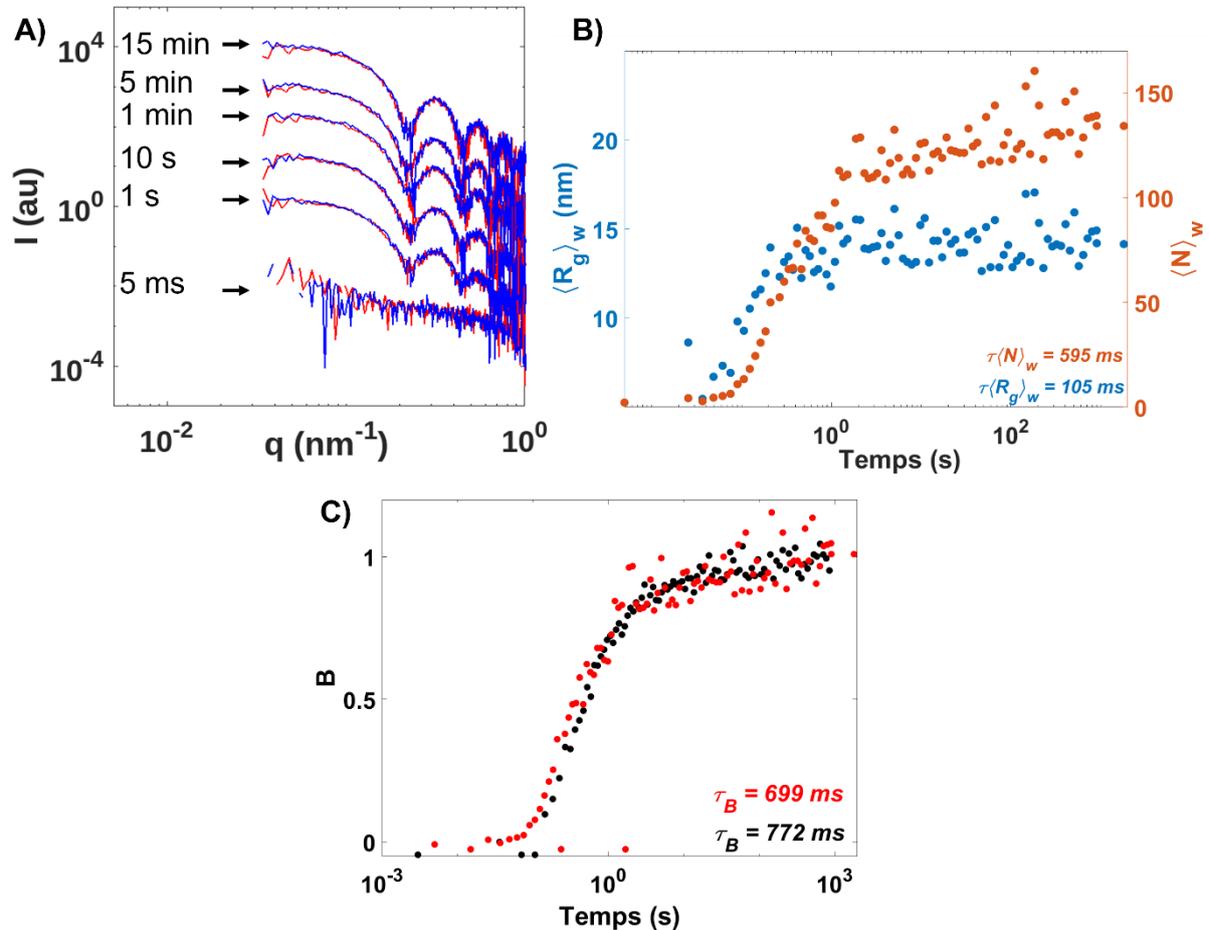


**Figure 4.18 : Données CryoEM obtenues pour l'assemblage des capsides en présence de GLP-26.** (A) Micrographie de cryoEM des capsides assemblées puis concentrées 11 fois. La concentration initiale en sous-unités pour l'assemblage était de  $30 \mu\text{M}$ . La barre d'échelle est de 50 nm. L'échantillon est observé sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN Ultrascan 4K CCD. (B) Classification 2D des particules identifiées par cryoSPARC. La taille de la boîte est de 50 nm. (C) Tranches d'espace réel de la reconstruction 3D de cryoSPARC pour les particules de la taille de la capside T=4. (D) Reconstruction 3D de la capside T=4. La barre d'échelle est de 10 nm.

Les capsides formées à partir de  $30 \mu\text{M}$  de sous-unités et  $200 \text{ mM}$  d'AmAc, en présence de  $90 \mu\text{M}$  de GLP-26 ont ensuite été analysées par cryomicroscopie électronique. La figure 4.18 ci-dessus illustre les différentes données obtenues par

cryomicroscopie électronique pour cette condition. La classification 2D (B) a été réalisée en utilisant les objets extraits de ces micrographies, permettant ainsi de sélectionner 4943 particules issues de 43 micrographies. Cette classification 2D montre qu'en présence de 90  $\mu\text{M}$  de GLP-26, l'assemblage conduit à la formation de trois populations apparentes d'objets : une minorité semblant correspondre à des capsides T=3 (1,8 %), une grande majorité d'objets très similaires aux capsides T=4 formées en absence de modulateurs (81,3 %), et 17,5 % d'objets plus grands et d'apparence ellipsoïdale. A partir des particules similaires aux capsides T=4, nous avons effectué une reconstruction 3D en imposant la symétrie icosaédrique. Cette démarche nous a permis d'obtenir une carte de densité électronique de la capsid de Cp149 à 8,9 Å de résolution, ce qui correspond à la limite de résolution atteignable avec le microscope et la caméra utilisés, en utilisant 4284 particules. Cependant, étant donné que le nombre de particules formant des objets allongés et ellipsoïdaux est très faible, et que ces particules n'ont pas de symétrie icosaédrique qui aurait pu permettre d'augmenter d'un facteur 60 le nombre de particules utilisées lors de la reconstruction, nous n'avons malheureusement pas pu reconstruire ces objets avec une résolution satisfaisante.

#### 4.4 Contrairement à GLP-26, RS-3427 n'accélère pas la formation des capsides et entraîne de légères modifications structurales



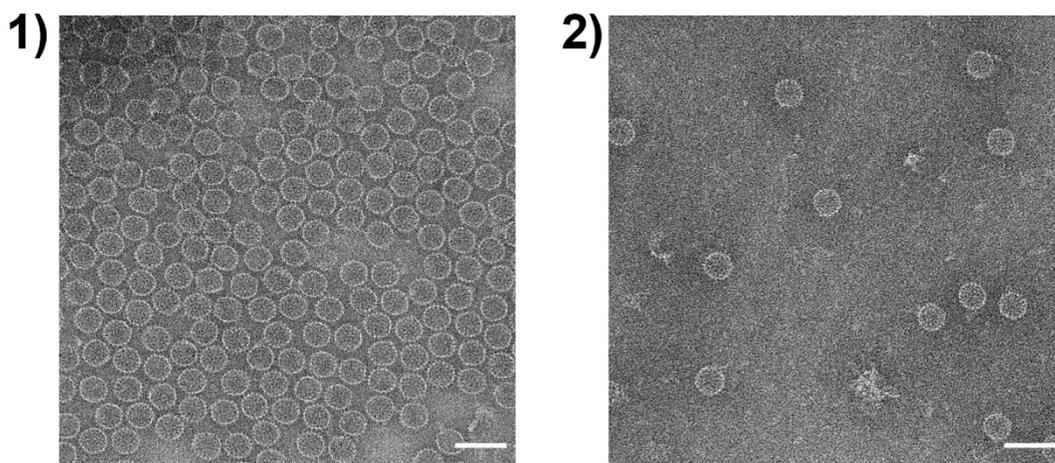
**Figure 4.19 : Mesures TR-SAXS de l'assemblage de capsides Cp149 en présence de la molécule RS-3427.** (A) Courbes TR-SAXS de deux expériences d'assemblages (courbes bleue et rouge) de 30  $\mu\text{M}$  de sous-unités en présence de 90  $\mu\text{M}$  de RS-3427. (B) Evolution du nombre moyen d'agrégation  $\langle N \rangle_w$  (disques orange) et du rayon de giration moyen  $\langle R_g \rangle_w$  (disques bleus) en fonction du temps pour l'assemblage de capsides avec 30  $\mu\text{M}$  de sous-unités en présence de 90  $\mu\text{M}$  de RS-3427. (C) Evolution du facteur B défini comme  $B = \frac{I(0)_t - \langle I(0)_i \rangle}{\langle I(0)_f \rangle - \langle I(0)_i \rangle}$  en fonction du temps lors d'un assemblage sans modulateurs (disques noirs) et en présence de 90  $\mu\text{M}$  de RS-3427 (disques rouges).  $I(0)_t$ ,  $\langle I(0)_i \rangle$ , et  $\langle I(0)_f \rangle$ , correspondent respectivement à l'intensité  $I(0)$  au temps  $t$ , à l'intensité moyenne  $I(0)$  des deux premières courbes de la cinétique et à l'intensité moyenne  $I(0)$  des cinq dernières courbes de la cinétique. Pour (B) et (C), les échelles de temps caractéristiques ont été obtenues en ajustant une fonction de décroissance exponentielle.

30  $\mu\text{M}$  de dimères de Cp149 ont été assemblés avec 200 mM d'acétate d'ammonium en présence de 90  $\mu\text{M}$  de RS-3427. Les courbes SAXS (Figure 4.19.a) des cinétiques

de formation des capsides dans cette condition sont représentées ci-dessus. En présence de 90  $\mu\text{M}$  de RS-3427, des objets sont formés et une première oscillation à  $q = 0,023 \text{ \AA}^{-1}$  est observée. La valeur du nombre  $q$  de cette première oscillation indique qu'une majorité des objets formés dans cette condition sont des capsides  $T=4$ .

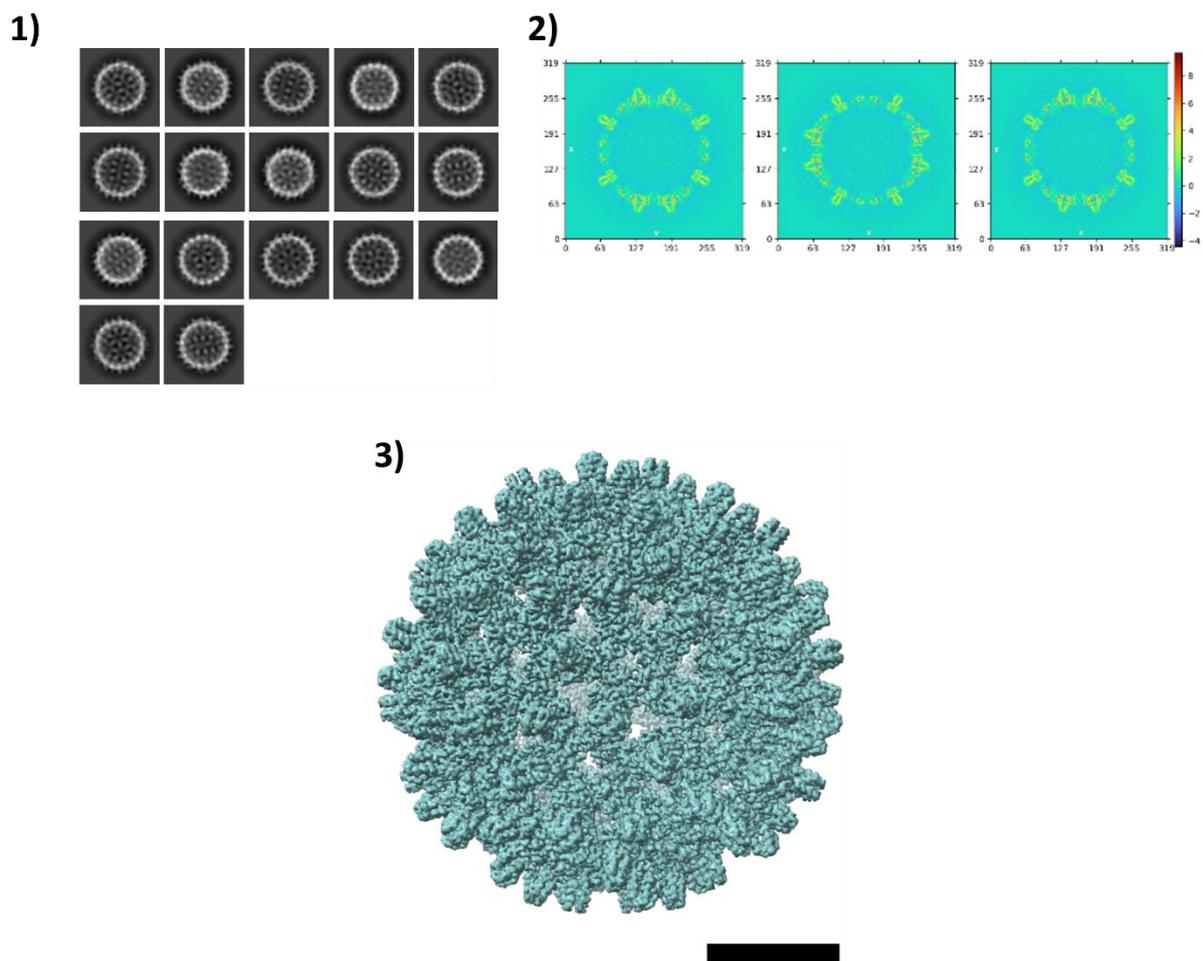
La figure 4.19.b ci-dessus représente l'évolution des  $\langle R_g \rangle_w$  et des  $\langle N \rangle_w$  lors de l'assemblage de capsides formées à partir de 30  $\mu\text{M}$  de sous-unités et 90  $\mu\text{M}$  de RS-3427. Comme pour les cinétiques d'assemblage en présence de GLP-26, les données présentées ici sont comparées à la cinétique d'assemblage de 30  $\mu\text{M}$  de dimères de Cp149 réalisée au synchrotron SOLEIL. En présence de 90  $\mu\text{M}$  de RS-3427, la formation de la capside n'est pas accélérée mais les objets formés semblent légèrement plus petits. Le  $R_g$  est égal à 141  $\text{\AA}$  et l'échelle de temps associée est égale à  $105 \pm 19 \text{ ms}$ . A la fin de la cinétique, la valeur du nombre moyen d'agrégation à l'équilibre  $\langle N \rangle_w^\infty$  est égale à 126. La valeur du  $\langle N \rangle_w^\infty$  estimée pour cette condition indique également que les objets formés sont des capsides  $T=4$ . Contrairement à la valeur du  $\langle R_g \rangle_w$  qui se stabilise à la fin des cinétiques d'assemblage, le  $\langle N \rangle_w$  ne se stabilise pas et continue d'augmenter.

Enfin, l'échelle de temps caractéristique extraite pour le paramètre B (Figure 4.19.c), qui est égale à  $699 \pm 142 \text{ ms}$  est similaire à celle obtenue lors de la cinétique d'assemblage des capsides en absence de modulateur.



**Figure 4.20 : Micrographies des capsides assemblées en présence ou non de RS-3427 et observées en microscopie électronique en coloration négative.** 30  $\mu\text{M}$  de capsides ont été assemblées en absence de modulateurs (1) ou en présence de 90  $\mu\text{M}$  de RS-3427 (2). Les barres d'échelle sont égales à 50 nm. Les échantillons sont observés sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN Ultrascan 4K CCD.

La figure 4.20 ci-dessus montre les images de microscopie électronique des capsides de Cp149 assemblées à partir de 30  $\mu\text{M}$  de sous-unités en présence ou non de 90  $\mu\text{M}$  de RS-3427. Les objets formés en présence de RS-3427 ont des formes et des tailles similaires à celle de la capside T=4 du VHB. Cependant, nous remarquons une moindre quantité de capsides formées sur la grille par rapport à un assemblage classique. Cette réduction du nombre de capsides lors d'un assemblage en présence de ce modulateur pourrait donc expliquer la baisse de la valeur du  $I(0)$  mesurée à partir de la courbe SAXS de ces objets à l'équilibre.



**Figure 4.21 : Données CryoEM obtenues pour l'assemblage des capsides en présence de 90  $\mu\text{M}$  de RS-3427.** (1) Classification 2D des particules extraites dans cryoSPARC (2) Tranches d'espace réel de la reconstruction 3D (3) Carte de densité électronique de la capside Cp149 à 3,33 Å. La taille de la boîte est de 50 nm et les barres d'échelle sont de 10 nm. L'échantillon est observé sur un microscope Glacios cryo-TEM et les images enregistrées avec une caméra Falcon 3EC.

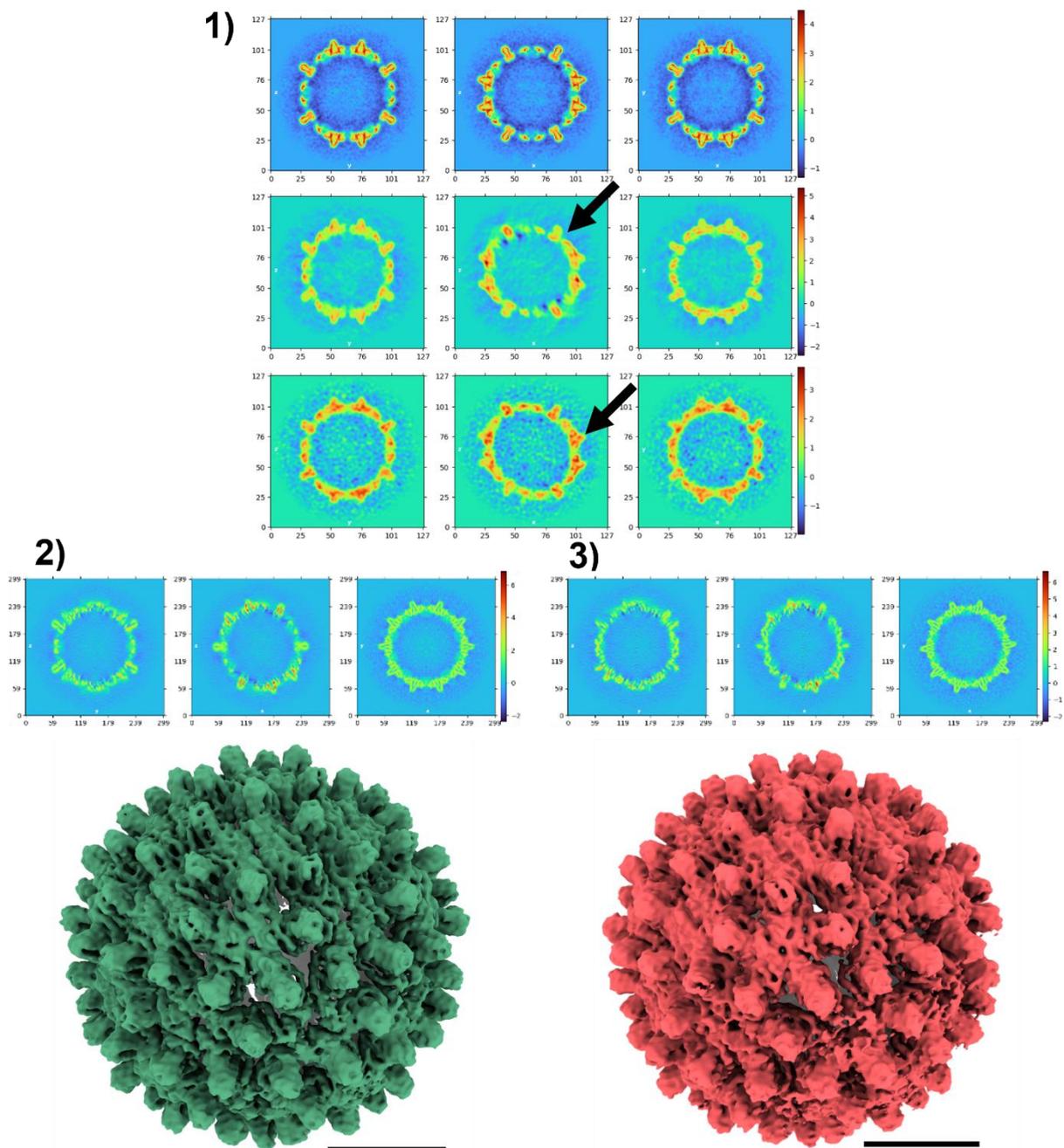
Les capsides formées à partir de 30  $\mu\text{M}$  de sous-unités en présence de 90  $\mu\text{M}$  de RS-3427 ont ensuite été analysées par cryomicroscopie électronique. La figure 4.21 ci-dessus illustre les différentes données obtenues dans cette condition expérimentale.

Une première étape de « *pre-processing* » des données a été réalisée (Figure 4.21.1). Cette étape a permis d'extraire 68 789 particules à partir de 1 349 micrographies. Ensuite, une classification 2D de ces particules a été effectuée, révélant une homogénéité marquée et mettant en évidence la présence de capsides fortement similaires à la capsid T=4 du VHB. Contrairement aux assemblages en présence de GLP-26, aucune capsid ellipsoïdale, allongée ou cassée n'a été observée dans cette condition. À la suite de cette classification, une première reconstruction 3D a été réalisée en imposant la symétrie icosaédrique. Cette reconstruction a permis d'obtenir une carte de densité électronique de la capsid de Cp149 avec une résolution de 3,86 Å, en utilisant un total de 32 606 particules.

Ensuite, une deuxième étape dite de « *polishing* », a été entreprise. Dans cette étape, plusieurs affinements de la carte 3D ont été effectués. Nous avons ainsi réussi à obtenir une seconde carte de densité électronique de la capsid avec une résolution de 3,73 Å. Bien que cette carte soit légèrement plus précise que la précédente, malheureusement, elle ne nous permet toujours pas de déterminer clairement la présence de la molécule RS-3427.

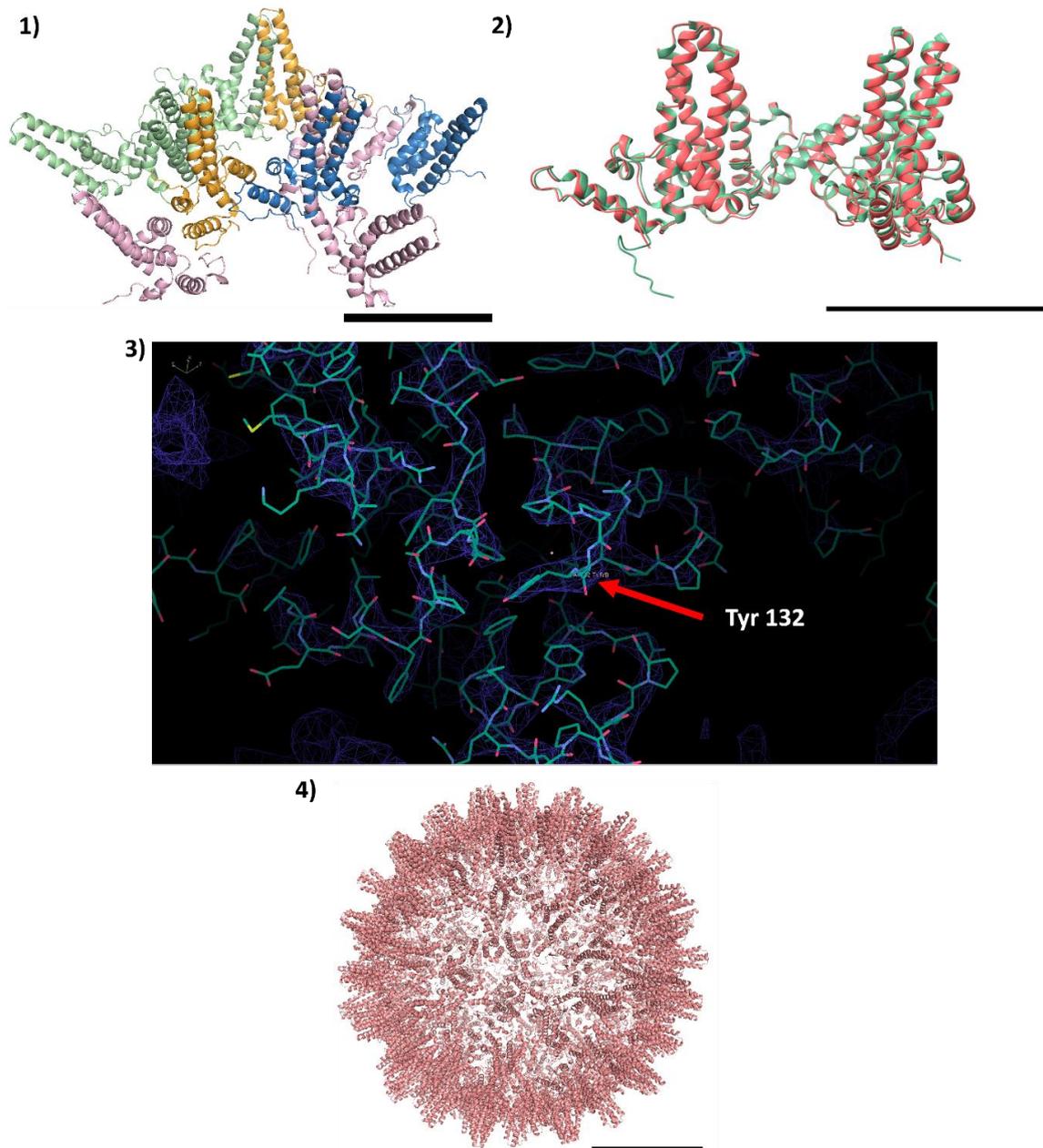
Afin d'observer la densité électronique de RS-3427, nous avons entrepris deux types d'analyses. Tout d'abord, nous avons effectué des reconstructions localisées de la capsid en utilisant les logiciels Scipion et cryoSPARC. Malheureusement, ces tentatives n'ont pas abouti, et nous n'avons pas pu observer la densité électronique souhaitée pour la molécule RS-3427. Par la suite, nous avons essayé d'améliorer la résolution de la carte de densité de la capsid Cp149. Pour cela, nous avons estimé les paramètres de transfert de contraste de chaque particule et corrigé les mouvements locaux dans les micrographies afin d'améliorer la qualité des données. Ces étapes ont permis une reconstruction tridimensionnelle plus précise de la capsid. Enfin, après avoir obtenu un modèle initial, celui-ci a été affiné de manière non-uniforme pour obtenir la meilleure résolution possible. Finalement, nous avons réussi à obtenir une carte de la capsid Cp149 à une résolution de 3,33 Å en utilisant 33 388 particules (Figure 4.21.3). Malgré ces améliorations significatives, la nouvelle carte ne

permet toujours pas d'observer de densité électronique correspondant à la molécule RS-3427.



**Figure 4.22 : Données cryoEM obtenues pour l'analyse des capsides en présence de RS-3427 et présentant des défauts.** (1) Tranches d'espace réel de la reconstruction 3D des 3 types de capsides présents dans l'échantillon. Les flèches noires indiquent la présence de défauts dans la capside. (2) Tranches d'espaces réel de la reconstruction 3D (haut) et carte 3D de la capside Cp149 (bas) reconstruite en imposant la symétrie D5. (3) Comme la (2) mais la symétrie C5 a été imposée. Les barres d'échelle sont de 10 nm.

Enfin, lors de la deuxième étape, appelée étape de « *polishing* », la première reconstruction 3D a mis en évidence la présence d'un nombre significatif de capsides qui ne semblaient pas suivre la symétrie icosaédrique. Environ 45% des capsides présentaient des défauts (Figure 4.22.1). Nous avons donc procédé à de nouvelles reconstructions tridimensionnelles en imposant soit la symétrie d'ordre 5 (C5) ou la symétrie D5, qui comprend des axes C5 et des axes C2 perpendiculaires à ces axes C5. Ces approches ont permis d'obtenir des cartes de densités électroniques à une résolution d'environ 6 Å à partir d'environ 27 500 particules (Figure 4.22.2 et 4.22.3).

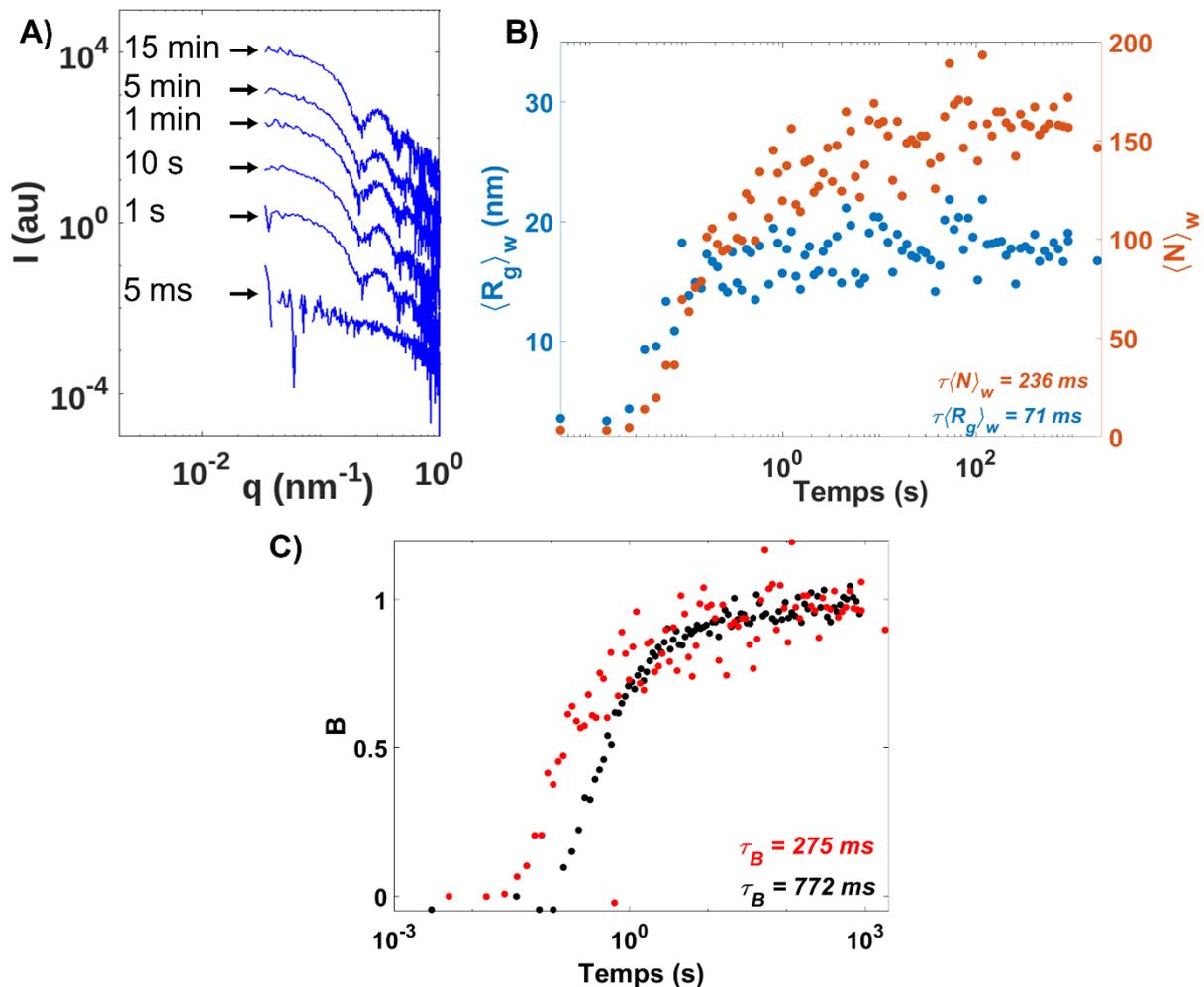


**Figure 4.23 : Reconstruction de la structure atomique de la capside Cp149.** (1) Unité asymétrique de la capside Cp149 reconstruite. Les chaînes A, B, C, et D sont colorées en vert, orange, bleu et rose respectivement. La barre d'échelle est égale à 5 nm. (2) Superposition du dimère de Cp149 reconstruit (rose) sur le dimère de Cp149 issu de la structure cryo-EM de la capside Cp183 (vert). Code PDB : 6HTX. La barre d'échelle est égale à 5 nm. (3) Superposition de l'unité asymétrique dans la carte de densité électronique de la capside de Cp149 focalisé sur la tyrosine 132. (4) Structure atomique de la capside Cp149 formée en présence de 90  $\mu\text{M}$  de RS-3427. La barre d'échelle est égale à 10 nm.

La structure de la capside de Cp149 a été reconstruite (Figure 4.23) en utilisant la carte cryo-EM intermédiaire de la capside à 3,73 Å obtenue grâce aux analyses sur cryoSPARC et la structure 6HTX de la capside de Cp183 comme modèle initial. Cette reconstruction a été réalisée en utilisant les logiciels Phenix et Coot. La superposition

du modèle obtenu avec le dimère issu du modèle initial révèle une grande similitude entre les deux (Figure 4.23.2). Néanmoins, quelques légères différences ont pu être observées. Dans notre modèle, l'hélice alpha 5 qui est impliqué dans l'interface dimérique, est légèrement décalée par rapport à l'hélice alpha 5 du modèle cristallographique. De plus, le sommet des spicules est moins bien défini dans notre construction. Enfin, la superposition de notre modèle avec la carte expérimentale (Figure 4.23.3) confirme l'absence de densité pouvant correspondre à la molécule RS-3427 au niveau des interfaces de contact entre deux dimères.

## 4.5 Conclusion : Les modulateurs de type CAM-E ; une classe diverse et variée



**Figure 4.24 : Mesures TR-SAXS de l'assemblage de capsides Cp149 en présence de 90  $\mu\text{M}$  de JNJ-632.** (A) Courbes TR-SAXS de deux expériences d'assemblages (courbes bleue et rouge) de 30  $\mu\text{M}$  de sous-unités en présence de 90  $\mu\text{M}$  de JNJ-632. (B) Evolution du nombre moyen d'agrégation  $\langle N \rangle_w$  (disques orange) et du rayon de giration moyen  $\langle R_g \rangle_w$  (disques bleus) en fonction du temps pour l'assemblage de capsides formées à partir de 30  $\mu\text{M}$  de sous-unités en présence de 90  $\mu\text{M}$  de JNJ-632. (C) Evolution du facteur  $B$  défini comme  $B = \frac{I(0)_t - \langle I(0)_i \rangle}{\langle I(0)_f \rangle - \langle I(0)_i \rangle}$  en fonction du temps lors d'un assemblage sans modulateurs (disques noires) et en présence de 90  $\mu\text{M}$  de JNJ-632 (disques rouges).  $I(0)_t$ ,  $\langle I(0)_i \rangle$ , et  $\langle I(0)_f \rangle$ , correspondent respectivement à l'intensité  $I(0)$  au temps  $t$ , à l'intensité moyenne  $I(0)$  des deux premières courbes de la cinétique et à l'intensité moyenne  $I(0)$  des cinq dernières courbes de la cinétique. Pour (B) et (C), les échelles de temps caractéristiques ont été obtenues en ajustant une fonction de décroissance exponentielle.

Afin de comparer ces résultats avec un autre modulateur de type CAM-E, la cinétique de formation de 30  $\mu\text{M}$  de sous-unités en présence de 90  $\mu\text{M}$  de JNJ-632 a été

réalisée (Figure 4.24). Les valeurs obtenues pour  $\langle R_g \rangle_w$ , l'échelle caractéristique associée, ainsi que la valeur du  $\langle N \rangle_w^\infty$ , sont similaires à celles calculées lors de l'assemblage de 30  $\mu\text{M}$  de sous-unités en présence de 90  $\mu\text{M}$  de GLP-26. A la fin de la cinétique, la valeur du  $\langle R_g \rangle_w$  est de 175 Å, l'échelle de temps extraite est de  $71 \pm 19$  ms et la valeur du nombre moyen d'agrégation à l'équilibre est égale à 149. Pour rappel, à la fin de la cinétique en présence de GLP-26, la valeur du  $\langle R_g \rangle_w$  est de 180 Å, l'échelle de temps extraite est de  $108 \pm 66$  ms et la valeur du  $\langle N \rangle_w^\infty$  est égale à 190.

Les données cinétiques et structurales indiquent clairement que la molécule GLP-26 est un modulateur de type CAM-E. Elle présente les mêmes effets cinétiques que le modulateur JNJ-632, et les courbes SAXS ainsi que la cryomicroscopie électronique montrent qu'elle a également les mêmes effets structuraux.

Cependant, les cinétiques réalisées par TR-SAXS suggèrent que la présence de la molécule RS-3427 lors de l'assemblage n'accélère pas la croissance des capsides. De plus, elle conduirait à la formation d'un nombre plus faible de capsides T=4. La diminution de la valeur  $I(0)$  sur la courbe SAXS des capsides formées dans ces conditions et mesurées à l'équilibre indique que le rayon de giration des objets formés est plus bas que celui des capsides classiques, ou bien que moins de capsides sont formées, ou encore les deux à la fois. Cependant, la première oscillation à  $q = 0,023 \text{ \AA}^{-1}$  indique que la majorité des objets formés dans cette condition sont des capsides T=4, suggérant ainsi qu'elles auraient le rayon de giration des capsides T=4 ce qui n'était pas le cas. Cette diminution de l'intensité  $I(0)$  est donc due à la diminution du nombre de capsides formées dans cette condition. Une hypothèse serait que la molécule RS-3427 empêcherait fortement la formation de capsides du VHB une fois liée à une sous-unité dimérique de Cp149.

## 5. Conclusion du chapitre

Dans ce chapitre, nous avons analysé les effets divers et variés des modulateurs d'assemblage de capsides sur l'auto-assemblage de la capsid Cp149 du VHB. Nos approches expérimentales et théoriques utilisées ici ont permis d'apporter plus d'éléments de réponses sur l'impact de ces modulateurs sur la structure mais également sur la cinétique d'assemblage des capsides du VHB.

Tout d'abord, nous avons retrouvé que l'auto-assemblage de la protéine dimérique Cp149 en capsides icosaédriques de triangulation  $T=4$  pouvait être réalisé en présence d'un autre sel que le chlorure de sodium, l'acétate d'ammonium (Asor *et al* 2020). Cela confirme que dans le cadre d'un déclenchement de l'assemblage par ajout de sel monovalent, seule la concentration ionique et non le type de molécules utilisées influence l'assemblage.

Nous avons ensuite étudié les effets des modulateurs Bay 41-4109 (un modulateur de type CAM-A) et JNJ-632 (un modulateur de type CAM-E) sur l'auto-assemblage de la capside Cp149 formée en présence de 150 mM d'acétate d'ammonium à pH 7,5. Nos résultats ont montré que les CAMs (de type A et de type E), accélèrent la cinétique d'assemblage des capsides et influencent leurs propriétés élastiques. De plus, les CAM-A entraînent des changements morphologiques drastiques lors de l'assemblage. Enfin, contrairement à ce qui est décrit dans la littérature, les CAM-E peuvent également modifier la structure des capsides du VHB et conduire à la formation de capsides légèrement allongées et ellipsoïdales.

Puis, en évaluant l'effet de deux nouvelles molécules GLP-26 et RS-3427, nous avons conclu que la molécule GLP-26 est un modulateur de type CAM-E puisqu'elle induit les mêmes effets cinétiques et structuraux que JNJ-632. En revanche, la molécule RS-3427 n'accélère pas l'auto-assemblage des capsides, n'induit pas de modification structurale conduisant à la formation d'objets allongés, et conduit à l'assemblage d'un nombre réduit de capsides. Dans la classe des modulateurs de type CAM-E, on retrouve des molécules qui accélèrent l'assemblage de la capside du VHB, et qui comme les CAM-A modifient la structure des objets formés mais de façon plus subtile. La molécule RS-3427 qui a des effets complètement nouveaux et différents des modulateurs connus jusqu'à présent, serait donc un nouveau type de modulateurs.

Ces résultats fournissent de précieuses informations pour la compréhension des mécanismes d'auto-assemblage et ouvrent la voie à de nouvelles perspectives de recherche pour le développement de nouvelles thérapies contre l'Hépatite B mais également pour le développement de protéines cargos servant au transport de molécules thérapeutiques.

## 6. Matériels et méthodes

### Expression de la protéine Cp149

Des précultures sont préparées en inoculant des bactéries transformées dans du milieu Lysogenic broth (LB). Pour cela, des bactéries transformées avec le plasmide codant pour la protéine Cp149 sont mélangées à 5 ml de LB contenant 100 µg/ml de carbénicilline et 34 µg/ml de chloramphénicol. La préculture est ensuite incubée toute la nuit à 37°C et 200 rpm.

Après l'incubation, si le milieu est trouble et que le témoin négatif ne l'est pas (Il s'agit d'un tube contenant uniquement 5 ml de LB. Il permet de s'assurer que le milieu utilisé pour la préculture n'est pas déjà contaminé), la préculture est diluée dans 2X500 ml de milieu LB contenant également 100 µg/ml de carbénicilline et 34 µg/ml de chloramphénicol. Ces cultures sont incubées pendant environ 5h à 37°C et 160 rpm. Un suivi de la densité optique à 600 nm ( $DO_{600nm}$ ) est effectué afin de suivre la croissance des bactéries. Lorsque cette DO atteint une valeur de 0,8, l'induction de l'expression de la protéine est réalisée en ajoutant 1 mM d'IPTG. Les bactéries sont ensuite incubées à 25°C et 160 rpm toute la nuit.

Après l'expression des capsides de Cp149, les cultures sont centrifugées pendant 20 minutes à 4°C et 6000 g (rotor JLA 10.500) afin de séparer les bactéries (qui seront dans le culot) du milieu de culture (surnageant). Les culots sont ensuite transférés dans des tubes Falcon de 50 ml et congelés à -80°C jusqu'à utilisation.

### Purification des capsides Cp149

La purification des capsides, dont l'essentiel des manipulations doit être réalisé à 4°C en chambre froide ou dans un bac à glace se déroule comme suit :

- 1- 2 jours avant la purification, préparer deux gradients de sucrose. Déposer dans des tubes « ultra-clear centrifuges » de chez Beckman 12 ml de sucrose 60 % puis 12 ml de sucrose 10 %. Les solutions de sucrose sont préparées à l'aide du tampon de lyse TN300. Ce tampon est composé de 50 mM Tris HCl pH 7,5, 2 mM DTT et 300 mM NaCl. Congeler les gradients en les laissant debout sur un support à -20°C jusqu'à utilisation.

- 2- Décongeler en chambre froide les deux gradients de sucrose ainsi que les culots bactérien la veille de la purification.
- 3- Re-suspendre les culots bactériens avec le tampon de lyse TN300. Le volume à ajouter dépend du poids du culot bactérien. 3 g de bactéries sont repris avec 7 ml de tampon TN300.
- 4- Ajouter 1 mg/ml de lysozyme, un cocktail inhibiteur de protéase et 0,5% de triton-X-100 dans la suspension bactérienne. Incuber sous agitation à 4°C pendant 1 h.
- 5- Ajouter 4 µl de nucléase commerciale Pierce et incuber la suspension à température ambiante sous agitation douce pendant 45 minutes.
- 6- Lyser les bactéries. Soniquer 15 secondes en mode pulse puis refroidir le tube dans un bac à glace pendant 50 secondes. Répéter 20 fois.
- 7- Récupérer la solution et la transférer dans un tube de centrifugation en polycarbonate de 26,3 ml. Centrifuger pendant 45 min à 35 000 g et 4°C (rotor 70 Ti, 19 000 rpm).
- 8- Récupérer le surnageant qui contient les capsides du VHB. Pour le gel SDS-page, prélever 10 µl et les réserver dans un tube Eppendorf noté S.
- 9- Déposer soigneusement le surnageant sur un gradient de sucrose linéaire de 10% à 60% et centrifuger 3 h à 100 000 g et 4°C (rotor SW-32 Ti, 28 800 rpm).
- 10-Récupérer le surnageant en 13 fractions de 2,7 ml en prenant toujours la partie la plus haute de la solution et conserver les fractions à 4°C. Pour le gel SDS-page, prélever 10 µl de chaque fraction et les réserver à 4°C dans des tubes Eppendorf noté 1 à 13.
- 11-Identifier les fractions où se trouve la protéine en utilisant le gel SDS-page.
- 12-Rassembler les fractions contenant les capsides du VHB et prélever 10 µl dans un tube noté SC.
- 13-Ajouter lentement du sulfate d'ammonium à la solution de protéines SC, tout en agitant, à température ambiante. La quantité de sulfate d'ammonium à ajouter pour une précipitation à 40% est calculée en fonction du volume de la solution de protéines. Elle est calculée sur le site suivant : <https://www.encorbio.com/protocols/AM-SO4.htm> en saisissant 0% de saturation initiale en sulfate d'ammonium. Après ajout du sel, laisser la solution sous agitation douce et à température ambiante pendant 1 h puis conserver à 4°C minimum 2 h ou toute la nuit.

- 14-Récupérer le mélange protéine/sulfate d'ammonium et centrifuger à 20 000 g et 4°C pendant 45 minutes (rotor 70-Ti, 14 000 rpm).
- 15-Récupérer le surnageant et prélever 10 µl noté S1. Re-suspendre le culot dans le tampon de purification contenant 5% de sucrose, 1 mM DTT et 50 mM Tris HCL pH 7,5. Centrifuger 45 minutes à 20 000 g et 4°C (rotor 70-Ti, 14 000 rpm).
- 16-Récupérer le surnageant, prélever 10 µl noté S2. Re-suspendre le culot avec le tampon de purification et prélever 10 µl noté P2.
- 17-Mesurer au NanoDrop la concentration en protéines des solutions S2 et P2.

La solution S2 contient les capsides de Cp149 solubles. Elle est conservée à 4°C.

## **Purification des dimères Cp149**

La purification des dimères à 4°C se déroule comme suit :

- 1- Dialyser la solution S2 à 4°C 1h puis toute la nuit dans un tampon de désassemblage contenant 1 mM DTT et 50 mM CHES pH 9. Utiliser pour cela une cassette de dialyse de volume adéquat et de cutoff 3,5 kDa.
- 2- Equilibrer la colonne Superdex S200 10/300 avec le tampon de désassemblage pendant la nuit.
- 3- Centrifuger la solution S2 dans un tube Amicon dont le seuil de coupure est de 3,5 kDa à 2000 g et 4°C pendant 5 minutes. Mesurer la concentration en protéines au NanoDrop et centrifuger autant de fois que nécessaire de façon à avoir une concentration de 4 mg/ml.
- 4- Ajouter 3 M d'urée solide dans la solution protéique pendant 1 h à 4°C sous agitation afin de complètement dissocier les capsides.
- 5- Filtrer la solution à l'aide d'un filtre 0,2 µm.
- 6- Purifier les dimères par chromatographie d'exclusion de taille (SEC).
- 7- Mesurer la concentration en protéines de chaque fraction contenant les dimères.
- 8- Rassembler les fractions contenant les dimères et mesurer la concentration finale en protéines.
- 9- Afin de reformer les capsides, dialyser les dimères à 4°C pendant 1 h puis toute la nuit dans le tampon d'assemblage composé de 50 mM Tris HCl pH 7,5, 1 mM DTT et 500 mM NaCl.
- 10-Purifier les capsides de Cp149 par SEC.

- 11-Récupérer les fractions contenant les capsides de Cp149, mesurer la concentration au NanoDrop.
- 12-Réaliser un deuxième cycle de dissociation des capsides et de purification des dimères afin d'obtenir une solution de dimères purs et actifs.
- 13-A la fin de la purification, mesurer la concentration finale en dimères de Cp149. Vérifier par DLS que la solution est bien monodisperse et que le rayon hydrodynamique moyen des objets correspond bien à celui de dimères de Cp149.
- 14-Lors des étapes 3, 5, 8, 11 et 13, prélever 10 µl d'échantillon. Réaliser un gel SDS-page avec ces échantillons afin de contrôler la pureté de dimères Cp149.
- 15-Aliquoter les dimères de Cp149 dans des tubes Eppendorf de 1 ml, congeler les tubes dans l'azote et les conserver à -80°C jusqu'à utilisation.

## Electrophorèse sur gel SDS-page

Le protocole de préparation et de réalisation de l'électrophorèse sur gel SDS-page est le suivant :

- 1- Préparer le gel SDS-page à l'aide des composés suivants :

	Gel de concentration (3 ml)	Gel de séparation 14% (5 ml)
H <sub>2</sub> O	2,25 ml	2,5 ml
Acrylamide 40%	270 µl	1,75 ml
3 M Tris HCl pH 8.8	0	625 µl
1 M Tris HCl pH 6.8	360 µl	0
SDS 10%	30 µl	50 µl
TEMED 100%	3 µl	5 µl
APS 10%	6 µl	70 µl

- 2- Préparer les échantillons : Dans un tube Eppendorf, mélanger 10 µl d'échantillons avec 2,5 µl de Laemmli 5X. Centrifuger quelques secondes et incubé 4 minutes à 95°C.
- 3- Déposer 12,5 µl de chaque échantillon dans les différents puits du gel ainsi que 6 µl du marqueur de poids moléculaires. Réaliser la migration du gel dans du tampon de migration composé de 2,5 mM de Tris HCl, 20 mM de glycine et

0,01% de SDS pendant 2 h environ à 125 V. Colorer le gel avec de l'Instant Blue afin de révéler les différentes bandes correspondant aux protéines séparées.

## **Assemblage en présence ou non des modulateurs d'assemblage**

Les dimères de Cp149 sont décongelés à 4°C, et incubés avec 3 M d'urée solide pendant 1 h à 4°C sous agitation. Les protéines sont ensuite dialysées pendant 2 h à 4°C dans le tampon 50 mM de Tris HCl pH 7,5 et 1 mM DTT. Le bain de dialyse est renouvelé après 1 h de dialyse.

Pendant ce temps, les solutions mères de sels et de modulateurs sont préparées dans un tampon Tris HCl à pH 7,5 ou dans un tampon CHES à pH 9 selon les conditions d'assemblage. Les sels testés sont le chlorure de sodium et l'acétate d'ammonium. Les modulateurs utilisés sont Bay 41-4109, un modulateur de la classe A, un modulateur de la classe E, JNJ-632, et deux nouvelles molécules RS-3427 et GLP-26.

Plusieurs conditions d'assemblage sont ensuite testées. L'assemblage est réalisé en mélangeant manuellement dans un tube Eppendorf une certaine concentration des protéines Cp149 dimériques avec une certaine concentration de sels en présence ou non des modulateurs. Les protéines sont ensuite incubées minimum 20 minutes à 37°C.

# Méthodes de caractérisation optique

## Diffusion dynamique de la lumière

Le diamètre hydrodynamique des dimères et des objets formés après des tests d'assemblage est vérifiée par DLS. Pour cela, 40  $\mu\text{l}$  de l'échantillon à mesurer est ajouté dans une cuve en quartz préalablement rincée à l'éthanol puis à l'eau. L'échantillon est ensuite mesuré dans l'appareil Malvern Zetasizer nano-S. Les mesures sont effectuées à 18°C et pendant environ 2 minutes. La durée d'une mesure varie selon la qualité de l'échantillon et est ajustée automatiquement par l'appareil.

## Photométrie de masse

L'appareil utilisé pour réaliser les expériences de photométrie de masse est le Refeyn Two<sup>MP</sup>. Le protocole d'une mesure est le suivant :

- 1- Nettoyer une lame de verre en la rinçant plusieurs fois à l'eau et à l'isopropanol pour rendre la surface du verre hydrophobe. Sécher la lame de verre.
- 2- Rincer plusieurs fois le support perforé en silicone (Silicon gasket) à l'isopropanol. Puis le rincer à l'eau et le sécher.
- 3- Déposer le support perforé sur la lame de verre et déposer la lame sur verre sur l'objectif de l'appareil. L'objectif aura été préalablement nettoyé et immergé dans de l'huile.
- 4- Calibrer l'appareil. Déposer 18  $\mu\text{l}$  de tampon dans un trou et faire le focus à l'aide du logiciel Refeyn. Déposer ensuite 2  $\mu\text{l}$  de protéines dans le même trou, homogénéiser et lancer la mesure. La calibration se fait avec deux protéines : la BSA (66,4 kDa, 132,9 kDa et 199,3 kDa) et l'uréase (272,3 kDa et 544,6 kDa). A la fin de la mesure, attribuer la masse moléculaire de chaque conformation (dimérique/tétramérique etc.) des protéines à chaque pic obtenu lors de la mesure.
- 5- Mesurer l'échantillon. Déposer entre 15 et 18  $\mu\text{l}$  du tampon d'intérêt et faire le focus. Puis déposer entre 2 et 5  $\mu\text{l}$  de protéines et lancer la mesure. On peut donc réaliser des dilutions de notre échantillon avec le tampon allant d'un facteur 4 à un facteur 10.

- 6- Après la mesure, appliquer la calibration sur les données brutes afin d'obtenir les informations concernant la masse des échantillons présents dans la solution.

Il est possible de modifier la durée d'exposition et la durée d'acquisition mais généralement, celle-ci est de 60 secondes et après 2 minutes d'acquisition, la surface de verre est saturée par les protéines.

## **Méthodes structurales**

### **SAXS et TR-SAXS**

Les mesures SAXS et TR-SAXS ont été réalisées sur la ligne SWING au synchrotron Soleil et sur la ligne ID02 à l'ESRF.

50 µl des capsides formées généralement 2 à 3 jours avant l'expérience sont mis dans des inserts puis placées dans le robot Autosampler. Lors de la mesure, le détecteur est placé à 2,5 mètres et chaque échantillon est injecté dans un capillaire en quartz de 1,5 mm de diamètre, sous vide, afin d'être mesuré. La température du système de mesure est maintenue à 37°C. Le tampon associé à l'échantillon analysé est également mesuré. Ce tampon contient la même concentration en sel, en DMSO et en solution tampon que l'échantillon.

Les images 2D obtenues sont ensuite moyennées radialement afin d'obtenir une courbe de diffusion en utilisant le logiciel de la ligne de lumière (Foxtrot sur la ligne SWING au synchrotron SOLEIL et SAXSUtilities 2 sur la ligne ID02 à l'ESRF). Après soustraction du tampon, les intensités sont converties en unités absolues. Les courbes de diffusion sont ensuite analysées avec Primus où le rayon de giration  $R_g$  et l'intensité  $I(0)$  des objets sont calculés.

Pour les expériences de TR-SAXS, la veille de la manipulation, les dimères de Cp149 sont incubés avec 3 M d'urée pendant 1 h à 4°C sous agitation, puis dialysés avec le tampon contenant 50 mM de Tris HCl pH 7,5 et 1 mM DTT pendant 2 h, à 4°C. Les dimères sont ensuite concentrés à 1,25 mg/ml ou à 2,5 mg/ml. Lors de l'expérience, le détecteur est placé à une distance de 2, 2,5 ou 5 m selon les conditions d'assemblage et selon la ligne de lumière où l'expérience est réalisée. L'assemblage est ensuite déclenché à l'aide d'un stopped-flow en mélangeant la solution contenant

la protéine Cp149, le tampon de la protéine, la solution de sel en présence des modulateurs à différentes concentrations ou du DMSO. Chaque cinétique est réalisée au minimum 2 fois et dure entre 5 et 15 minutes. Des mesures en statiques sont également réalisées en fin de cinétique.

De plus, des tests de dilutions sont préalablement effectués avec du lysozyme et des dimères afin de s'assurer du bon fonctionnement du stopped-flow et de la ligne synchrotron avant de commencer les cinétiques d'assemblage.

Comme pour le SAXS statique, les images de 2D sont moyennées pour obtenir des courbes de diffusion et les intensités sont converties en unité absolue après soustraction du tampon. Les intensités  $I(0)$  ainsi que les rayons de giration moyen  $\langle R_g \rangle_w$  sont ensuite déterminées à l'aide des logiciels AUTORG et PRIMUS de la suite ATSAS en respectant la condition  $qR_g < 1.3$ , qui définit la région de Guinier. Le  $I(0)$  est ensuite utilisé pour déterminer le nombre moyen d'agrégation  $\langle N \rangle_w$  selon la formule suivante :

$$\langle N \rangle_w (t) = \frac{c^s I_0(t)}{c I_0^s}$$

où  $I_0^s$  représente l'intensité de la diffusion directe des sous-unités mesurée séparément à une concentration molaire  $c^s$  et  $c$  est la concentration molaire des sous-unités lors de l'expérience de cinétique d'assemblage. Ce nombre correspond au nombre moyen de dimères par objets.

Les temps de croissance des objets selon le  $R_g$  et le  $\langle N \rangle_w$  sont ensuite obtenues à l'aide d'une régression exponentielle sur MATLAB.

## Microscopie électronique

Les capsides déposées sur grilles sont généralement assemblées à 1 mg/ml et observées à une concentration d'environ 0,2 mg/ml. Après avoir ionisé les grilles de cuivre Formvar/Carbon film S162-3 300 mesh, 5  $\mu$ l de la solution contenant les capsides sont déposés sur la grille pendant 5 min. Puis l'excès de solution est retiré à l'aide d'un papier filtre et une goutte d'une solution à 1% de molybdate d'ammonium

préalablement filtrée est déposée pendant 45 secondes sur la grille. Enfin l'excédent de contrastant est retiré à l'aide d'un papier filtre.

Les grilles sont ensuite observées au microscope JEOL JEM-2010 à 2  $\mu\text{m}$  de défocalisation et un grossissement de  $\times 50\,000$ . Des images sont enregistrées avec la caméra GATAN Ultrascan 4K CCD et traitées grâce au logiciel d'analyse ImageJ.

## **Cryomicroscopie électronique et reconstruction de la capsid**

L'assemblage est déclenché manuellement en mélangeant 30  $\mu\text{M}$  (1 mg/ml) de dimères Cp149 contenu dans un tampon composé de 1 mM DTT et 20 mM Tris HCl à pH 7,5 en présence ou non des modulateurs à différentes concentrations. Les capsides assemblées sont ensuite incubées à 37°C pendant 1 h puis concentrées au minimum 10 fois à l'aide d'un tube Amicon de seuil de coupure 100 kDa. Les capsides formées en présence de Bay 41-4109 sont incubées pendant 24 h à 37°C et ne sont pas concentrées.

### **Congélation des grilles**

La congélation des grilles est réalisée à l'aide du Vitrobot FEI. Dans un premier temps, la grille Quantifoil R2/2 ou R2/1 (R2/2 pour les capsides assemblées sans modulateurs ou assemblées avec les CAM de type E, et R2/1 pour les capsides assemblées avec Bay 41-4109) en cuivre est nettoyée à l'acétate d'éthyle puis ionisée. La grille est ensuite positionnée dans le Vitrobot. 4  $\mu\text{l}$  de solution sont déposés sur la grille et celle-ci est époncée pendant 2 secondes avec une force de niveau 5 avant d'être automatiquement plongée dans l'éthane liquide refroidi par de l'azote liquide. La grille est ensuite stockée dans l'azote liquide avant utilisation.

### **Collecte des images de cryomicroscopie électronique**

Les échantillons congelés sont transférés sur un support Gatan 626 et observés à -180°C sur le microscope JEOL JEM-2010 équipé d'un canon d'émission de champ de 200 kV. Les échantillons ont été imagés avec un grossissement de  $\times 50\,000$  en utilisant un système de dose minimale (MDS). Les images sont ensuite enregistrées avec une caméra GATAN Ultrascan 4K CCD à 2  $\mu\text{m}$  de défocalisation.

Les grilles de cryomicroscopie ainsi que la collecte des images pour l'assemblage de la capsidie en présence de la molécule RS-3427 ont été réalisées par Ana Arteni sur le microscope Glacios II à l'Institut Pasteur.

### **Analyse des images de cryomicroscopie électronique**

Le « *particule picking* », la classification 2D ainsi que les reconstructions 3D sont réalisés à l'aide de protocoles standards du logiciel CryoSPARC. Après la correction de la CTF, une première classification 2D est réalisée à l'aide des particules préalablement sélectionnées et extraites manuellement. Ensuite, à partir des classes 2D sélectionnées, les particules sont vérifiées manuellement, extraites des micrographies et une deuxième série de classification 2D est effectuée. Enfin, les volumes 3D sont obtenus après plusieurs étapes de reconstruction *ab initio* et d'affinement homogène ou hétérogène.

### **Reconstruction de la capsidie de Cp149**

La reconstruction de la capsidie est réalisée à l'aide de protocoles standards du logiciel Phenix. Dans un premier temps, l'unité asymétrique de la capsidie de Cp183, issue du modèle cryo-EM 6HTX, est alignée sur la carte de la capsidie de Cp149, obtenue à partir des analyses sur cryoSPARC. Les acides aminés C-terminaux qui ne sont pas superposables à la carte sont ensuite supprimés, créant ainsi un nouveau modèle. Ce nouveau modèle est ensuite utilisé pour les différentes étapes de reconstruction dans Phenix. Tout d'abord, la carte de densité électronique est légèrement améliorée à l'aide de l'outil « *resolve\_cryo\_em* ». Ensuite, l'outil « *map\_to\_model* » est utilisé pour construire un premier modèle atomique en se basant sur la carte de densité électronique et la séquence en acide aminés de la protéine Cp149. Ensuite, l'outil « *dock\_in\_map* » est utilisé pour s'assurer qu'une partie de la carte de densité est bien similaire au premier modèle construit avec l'outil « *map\_to\_model* ». Après cette étape, le modèle est affiné à l'aide de l'outil « *real\_space\_refine* ». La structure obtenue est ensuite analysée et reconstruite à l'aide du logiciel coot, puis soumise à une nouvelle phase d'affinement dans Phenix.

## Bibliographie

Asor, R., Schlicksup, C. J., Zhao, Z., Zlotnick, A. & Raviv, U. Rapidly Forming Early Intermediate Structures Dictate the Pathway of Capsid Assembly. *J. Am. Chem. Soc.* **142**, 7868–7882 (2020).

Chevreuil, M. *et al.* Nonsymmetrical dynamics of the HBV capsid assembly and disassembly evidenced by their transient species. *J. Phys. Chem. B* **124**, 9987–9995 (2020).

Kra, Kalouna, Siyu Li, Laetitia Gargowitsch, Jéril Degrouard, Javier Pérez, Roya Zandi, Stéphane Bressanelli, et Guillaume Tresset. 2023. « Energetics and Kinetic Assembly Pathways of Hepatitis B Virus Capsids in the Presence of Antivirals ». *ACS Nano* 17 (13): 12723-33. <https://doi.org/10.1021/acsnano.3c03595>.

# Chapitre 5 : Effets des modulateurs d'assemblage sur la cinétique et la structure des capsides formées à partir des protéines Cp183 et des protéines Cp183-EEE

## 1. Introduction

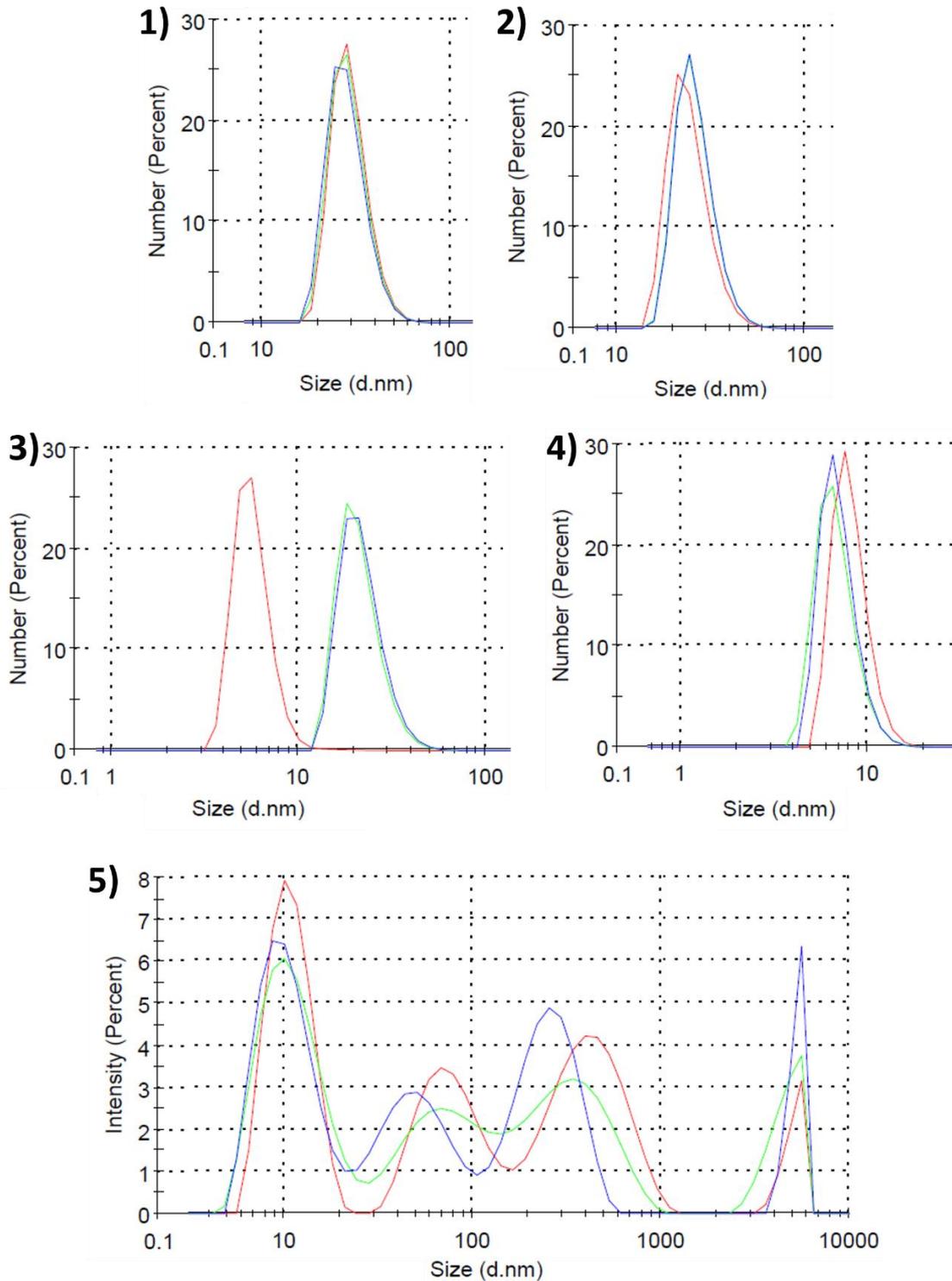
Dans le chapitre précédent, nous avons étudié l'influence des modulateurs d'assemblage sur l'assemblage de la capside du VHB formée à partir des sous-unités dimériques de la protéine Cp149. Cette protéine est une protéine tronquée de la protéine Core, composée uniquement du domaine NTD, permettant l'auto-assemblage de la capside du VHB. Cependant, cette protéine tronquée n'existe pas *in vivo*, dans le cycle viral du VHB. Nous avons donc choisi dans ce chapitre, d'étudier l'impact de cette classe de molécules virales sur deux formes de la protéine Core : la protéine Cp183, qui représente la forme entière de la protéine Core, ainsi que sur un mutant de cette protéine, la protéine Cp183-EEE, qui mime un état de phosphorylation d'assemblage de la capside. Cette protéine correspond à la protéine Cp183 où les trois sérines essentielles lors de l'encapsidation sont mutées en glutamates.

Dans ce chapitre, nous présenterons une étude sur la stabilité du dimère en fonction du tampon de stockage. Ensuite, nous examinerons en détail nos résultats préliminaires concernant l'effet de l'ARN pré-génomique (ARNpg) sur la cinétique d'assemblage et la structure des capsides formées à partir des protéines dimériques de Cp183 et de Cp183-EEE. Ces résultats ont été obtenus à la suite des expériences réalisées par TR-SAXS, par SAXS et par microscopie électronique. Nous discuterons également de l'influence des modulateurs sur la structure des capsides formées à partir de la Cp183 et de la Cp183-EEE. Enfin, la dernière partie de ce chapitre sera consacrée à la conclusion de cette étude et à une comparaison de ces résultats avec ceux obtenus concernant la Cp149.

## **2. Un pH basique et le chlorure de guanidinium sont nécessaires pour maintenir la protéine Cp183-EEE sous sa forme dimérique**

A la fin de la purification, le dimère de la protéine Cp183-EEE (mais également celui de la protéine Cp183) se trouve dans un tampon pH 9, contenant 2 M de chlorure de guanidinium et 0,5 M de chlorure de lithium. L'objectif de cette étude est d'analyser les effets des modulateurs d'assemblage sur les protéines Cp183 et Cp183-EEE. Pour cela, nous avons cherché à travailler dans des conditions aussi proches que possibles des conditions physiologiques.

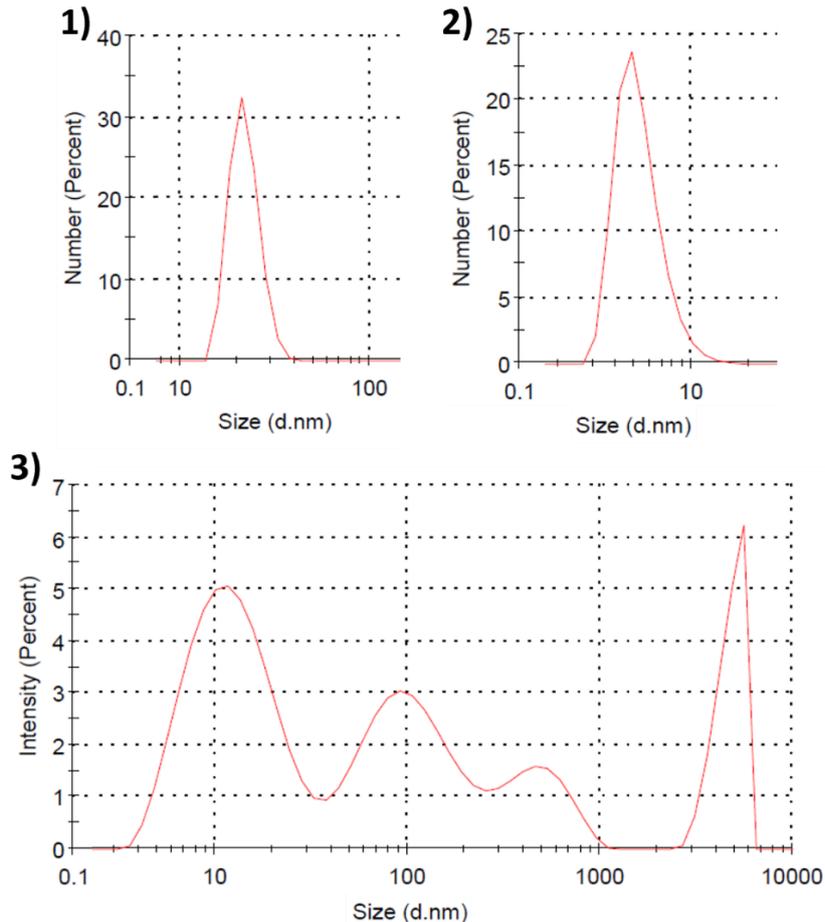
Pour atteindre cet objectif, la première étape a été de changer le tampon de stockage des dimères de ces deux protéines.



**Figure 5.1 : Mesures par DLS du diamètre hydrodynamique dimère de Cp183-EEE dans différents tampons.** 1) Tampon contenant 0,6 M de chlorure de guanidinium. 2) Tampon contenant 1 M de chlorure de guanidinium à pH 7,5. 3) Tampon contenant 1 M de chlorure de guanidinium à pH 9. 4 et 5) Tampon contenant 1,5 M de chlorure de guanidinium à pH 9. La distribution de taille est représentée selon le nombre de particules mesurées (1 à 4) et selon l'intensité des particules mesurées (5). Chaque échantillon est mesuré 3 fois à 37°C. La ligne rouge correspond à la première mesure, et les lignes vertes et bleues à la deuxième et à la troisième mesure respectivement.

La figure 5.1 indique la présence ou l'absence de capsides de la protéine Cp183-EEE après un changement de tampon. Lorsque le tampon contient une concentration faible de chlorure de guanidinium (GuHCl), inférieure à 1 M (Fig. 5.1.1), celle-ci est exclusivement sous forme de capsides (le diamètre hydrodynamique des capsides est d'environ 38 nm). En présence de 1 M de chlorure de guanidinium (Fig. 5.1.2 et 5.1.3), le pH et la température semblent influencer l'état oligomérique de la Cp183-EEE. Le passage d'un pH de 9 à un pH de 7,5, entraîne la formation de capsides. De plus, l'augmentation de la température des échantillons à pH 9 entraîne également la formation de capsides de Cp183-EEE. Nous remarquons que le chlorure de lithium n'est pas nécessaire pour maintenir la protéine sous sa forme dimérique puisqu'il n'est présent dans aucun des tampons testés. Enfin, nous pouvons noter que la condition « 10 mM CHES pH 9 ; 1,5 M GuHCl ; 2 mM DTT » montrée sur les graphiques 4 et 5 représente un compromis approprié. Elle permet de légèrement réduire la concentration en chlorure de guanidinium tout en maintenant la Cp183-EEE majoritairement sous sa forme dimérique (le diamètre hydrodynamique des dimères est d'environ 7 nm).

Afin de réaliser les mesures de SAXS et de TR-SAXS des capsides assemblées à partir des dimères de Cp183/Cp183-EEE, il est nécessaire d'atteindre une concentration finale de dimères supérieure ou égale à 12  $\mu$ M (0,5 g/l) lors de l'assemblage. La deuxième étape de l'optimisation de la solution initiale de dimères a donc consisté à trouver la concentration initiale la plus élevée, qui n'entraîne pas la formation spontanée de capsides.

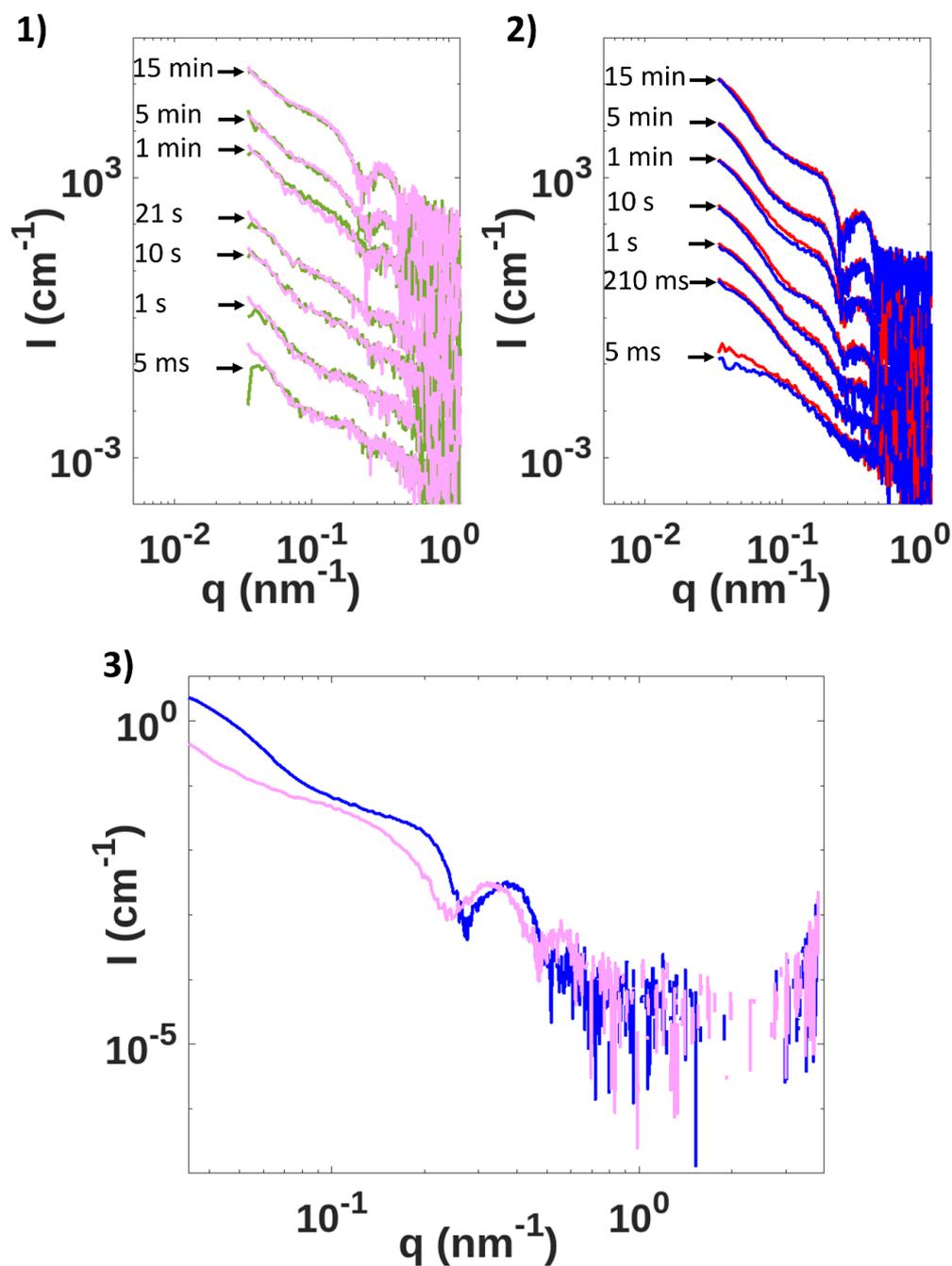


**Figure 5.2 : Mesures par DLS du diamètre hydrodynamique du dimère de Cp183-EEE à différentes concentrations.** 1) Dimères concentrés à 3 g/l. 2) Distribution de taille selon le nombre de particules mesurées du dimère de Cp183-EEE à 1 g/l. 3) Distribution de taille selon l'intensité des particules mesurées du dimère de Cp183-EEE à 1 g/l. Les échantillons sont mesurés à 18°C.

La figure 5.2 montre l'état oligomérique de la protéine Cp183-EEE à différentes concentrations. Lorsque le dimère est concentré à 3 g/l, la formation spontanée de capsides est observée. En revanche, à une concentration de 1 g/l, aucune capsid n'est présente dans l'échantillon. Cependant la présence de gros objets pouvant être des oligomères de dimères et d'agrégats est observée.

Pour les prochaines expériences, nous décidons donc de dialyser les dimères de Cp183-EEE mais également de Cp183 contre un tampon contenant « 10 mM CHES pH 9 ; 1,5 M GuHCl ; 2 mM DTT ». Avant les différents tests d'assemblage, nous ne dépasserons pas une concentration initiale de 1 g/l (24  $\mu$ M). Enfin, les tests d'assemblage où les objets seront ensuite analysés par SAXS et microscopie électronique se feront à température ambiante, tandis que les cinétiques d'assemblages seront réalisées à 15°C.

### 3. *In vitro*, les capsides formées à partir des protéines Cp183 et Cp183-EEE sont de type T=3

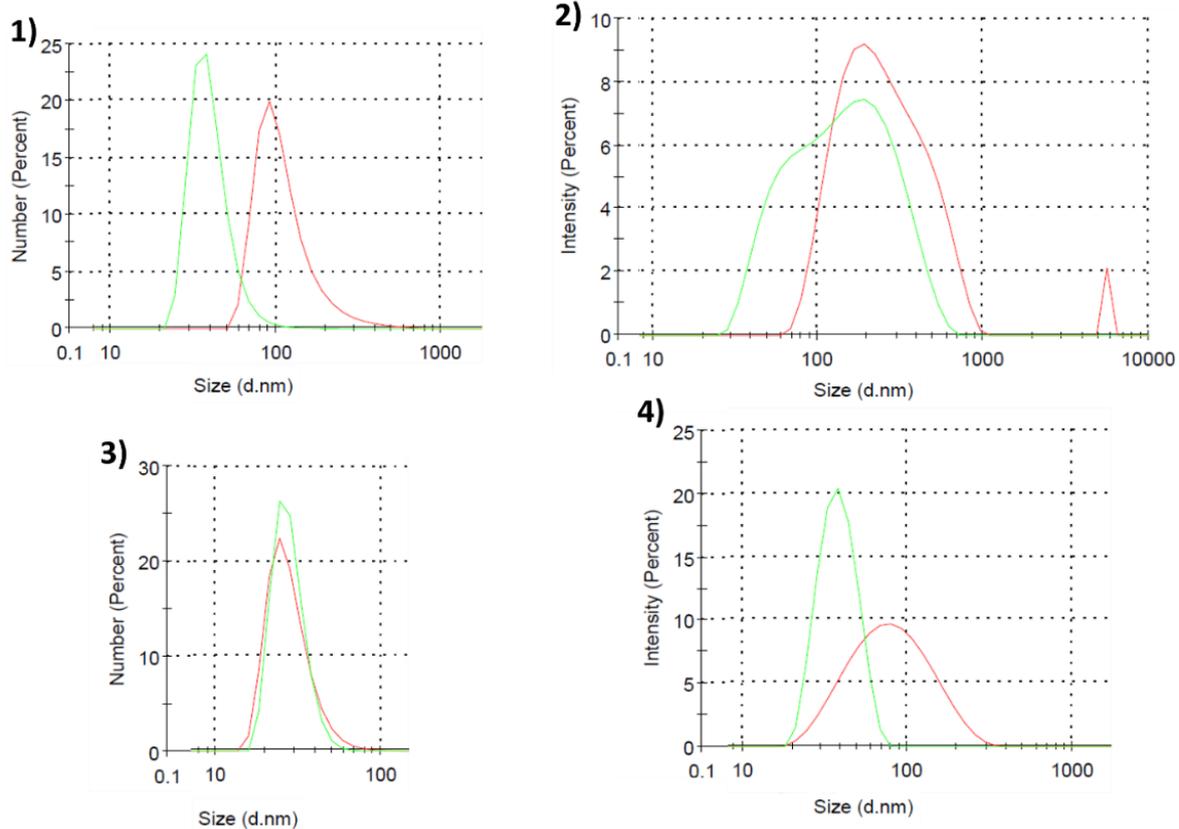


**Figure 5.3 : Mesures TR-SAXS de l'assemblage de capsides de Cp183-EEE.** 1) Courbes TR-SAXS de deux expériences identiques d'assemblage (lignes verte et rose) de 12  $\mu\text{M}$  de Cp183-EEE sans ARNpg. 2) Courbes TR-SAXS de deux expériences identiques d'assemblage (lignes bleue et rouge) de 12  $\mu\text{M}$  de Cp183-EEE en présence de 95 nM d'ARNpg. 3) Courbes SAXS de 12  $\mu\text{M}$  de Cp183-EEE en présence (ligne bleue) ou en l'absence (ligne rose) d'ARNpg, 30 minutes après le début de la cinétique.

Lors de l'expérience de TR-SAXS sur l'influence des molécules GLP-26 et RS-3427 sur la cinétique d'assemblage de la Cp149, nous avons également réalisé une étude préliminaire de la cinétique d'assemblage des capsides de Cp183-EEE en présence ou non d'ARNpg à 15°C.

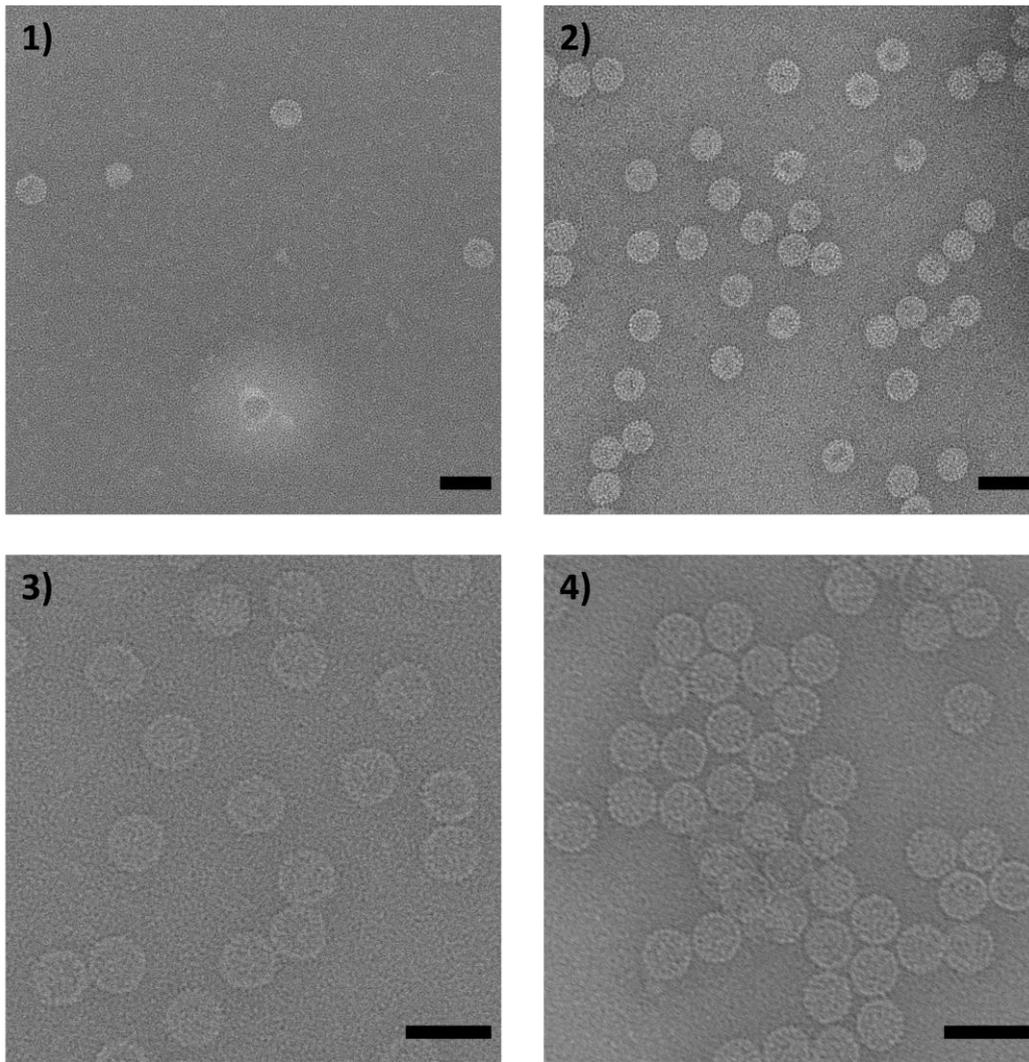
12  $\mu\text{M}$  (0,5 g/l) de dimères de Cp183-EEE ont été assemblés en présence ou non de 95 nM (0,1 g/l) d'ARNpg à 15°C. La concentration d'ARNpg finale est établie à 95 nM afin de respecter la condition 1 ARNpg pour 120 dimères de protéines Core et la concentration finale en chlorure de guanidinium est égale à 750 mM. Les courbes SAXS des cinétiques de formation des capsides assemblées dans ces conditions sont représentées sur la figure 5.3 ci-dessus. Dans les deux conditions, des objets sont formés et les cinétiques sont reproductibles. En absence d'ARNpg, la première oscillation apparaît au bout de 21 s. Cette oscillation est observée à  $q = 0,25 \text{ nm}^{-1}$ . En présence d'ARNpg, la première oscillation apparaît après 210 ms d'assemblage et est observée à  $q = 0,28 \text{ nm}^{-1}$ . La présence d'ARNpg influence la cinétique d'assemblage et la morphologie des objets formés au cours du temps. L'assemblage semble accéléré, et la valeur minimale du nombre d'onde  $q$  de la première oscillation suggère que les objets formés sont plus petits en présence d'ARNpg. Cependant, que ce soit en présence ou en l'absence d'ARNpg, les objets formés semblent être plus petits que des capsides T=4 du VHB.

Malheureusement, l'augmentation de la valeur de l'intensité mesurée aux petits angles ne nous a pas permis d'approfondir l'analyse des cinétiques. Pour les prochaines expériences d'assemblage, nous avons donc décidé de réaliser un assemblage plus progressif par dialyse, contre un tampon contenant 50 mM Tris HCl pH 7,5 et 150 mM d'acétate d'ammonium. Les objets formés seront ensuite centrifugés pour éliminer les éventuels agrégats et analysés à l'aide de la DLS, de la microscopie électronique et du SAXS.



**Figure 5.4 : Mesures par DLS des objets formés après l'assemblage de la Cp183 et de la Cp183-EEE en présence ou non d'ARNpg.** Distribution de taille de 12  $\mu$ M de Cp183 assemblés en présence (vert) ou en l'absence (rouge) d'ARNpg selon le nombre 1) et l'intensité 2) des particules mesurées. 3) et 4) Pareil que 1) et 2) pour la Cp183-EEE.

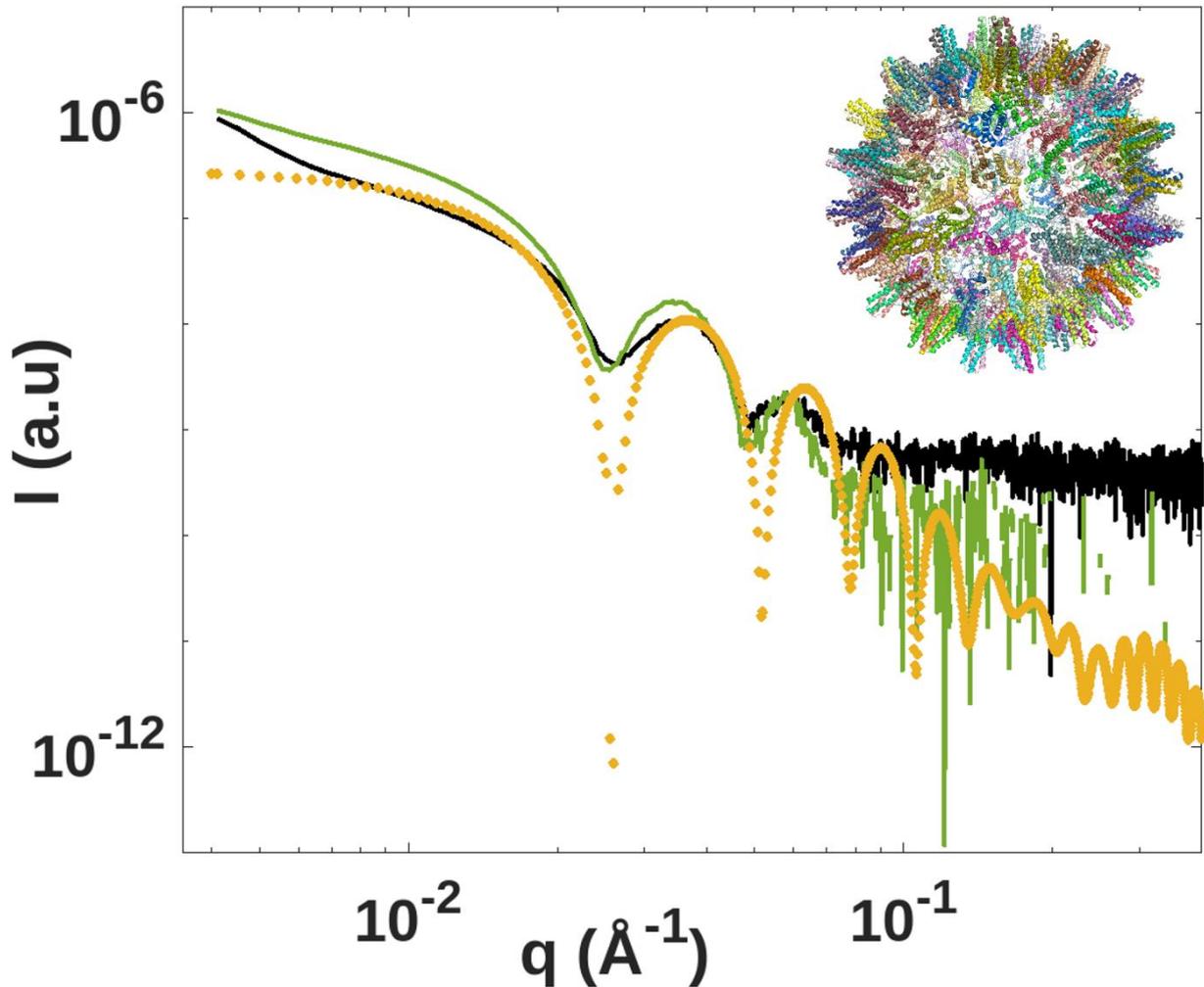
12  $\mu$ M de Cp183 et de Cp183-EEE ont été assemblés par dialyse en présence ou en l'absence de 95 nM d'ARNpg, à température ambiante. Les résultats des mesures de DLS (Figure 5.4) révèlent une grande polydispersité dans les échantillons mesurés. Selon le paramètre retenu pour le calcul de la distribution de taille, le diamètre hydrodynamique des objets formés varie. Lorsque la distribution de taille est calculée uniquement en fonction de l'intensité des particules mesurées, des pics très larges sont observés. Cependant lorsque la distribution de taille est établie en se basant sur le nombre de particules présentes dans l'échantillon, un unique pic relativement fin est observé. Il indique que le diamètre hydrodynamique des objets mesurés est de 28 nm pour l'échantillon contenant les capsides assemblées à partir de la Cp183-EEE et en présence d'ARNpg. Ce résultat est en adéquation avec le TR-SAXS qui indiquait que les objets formés étaient plus petits que des capsides T=4 du VHB. Cependant, le diamètre hydrodynamique est égal à 40 nm et 116 nm, dans le cas d'un assemblage de la Cp183 en présence ou en l'absence d'ARNpg respectivement.



**Figure 5.5 : Micrographies des objets formés lors de l'assemblage de la Cp183 et de la Cp183-EEE et observés en microscopie électronique par coloration négative.** Objets formés à partir de la Cp183 en présence 1) ou en l'absence 2) d'ARNpg. Objets formés à partir de la Cp183-EEE en présence 3) ou en l'absence 4) d'ARNpg. Les échantillons sont observés sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN K2. Les barres d'échelles sont égales à 50 nm.

Les objets formés à partir des différentes protéines et en présence ou non de l'ARNpg ont ensuite été observés en microscopie électronique par coloration négative. Les images ci-dessus (Figure 5.5) indiquent que les objets formés à partir des deux protéines ont des formes similaires aux capsides T=4 du VHB. Cependant, leur taille est d'environ 30 nm. Aucun agrégat n'a été observé, et très peu de capsides endommagées sont présentes sur les micrographies. Enfin, nous n'observons pas de différence morphologique entre les capsides formées avec de l'ARNpg et des capsides formées sans ARNpg.

Nous décidons donc pour la suite de travailler uniquement sur les capsides assemblées par dialyse et en présence d'ARNpg.



**Figure 5.6 : Courbes SAXS des capsides formées en présence de l'ARNpg.** Courbes SAXS expérimentales des capsides formées à partir de 12  $\mu\text{M}$  de dimères de Cp183 (ligne noire) et de 12  $\mu\text{M}$  de dimères de Cp183-EEE (ligne verte) en présence de 95 nM d'ARNpg et courbe SAXS (Ronds oranges) calculée sur CRYSOLOG, à partir de la structure cristallographique (en haut à droite sur l'image) de la capside T=3 de la Cp183 (code PDB 6BVN).

La figure 5.6 montre les courbes SAXS des objets formés lors de l'assemblage par dialyse de 12  $\mu\text{M}$  de protéines Cp183 ou Cp183-EEE en présence de 95 nM d'ARNpg. Des oscillations peuvent être observées sur cette figure. Ces oscillations indiquent la présence de capsides dans les deux échantillons. Les valeurs minimums du nombre d'onde  $q$  de la première oscillation sont égales à 0,025  $\text{\AA}^{-1}$  et 0,026  $\text{\AA}^{-1}$  pour la Cp183-EEE et la Cp183 respectivement. Ces valeurs sont très proches de la valeur théorique obtenue pour une capside T=3 du HBV, qui est égale à 0,025  $\text{\AA}^{-1}$ . Cela indique donc que les objets formés à partir des protéines Cp183 et Cp183-EEE en présence

d'ARNpg sont bien des capsides T=3. Nous remarquons cependant que contrairement à un assemblage classique de la Cp149 dans les mêmes concentrations, les courbes SAXS sont plus bruitées et moins d'oscillations sont observées. Enfin, l'augmentation de l'intensité aux petits angles, traduisant la présence d'agrégats dans les échantillons ne nous a pas permis d'approfondir l'analyse de ces résultats.

#### 4. Les modulateurs d'assemblage perturbent la structure des capsides Cp183 et Cp183-EEE

1)

<i>Conditions d'assemblage Cp183</i>	<i>Diamètre hydrodynamique (nm) selon l'intensité des particules</i>		<i>Diamètre hydrodynamique (nm) selon le nombre de particules</i>
<i>Sans modulateurs</i>	70	1596	37
<i>90 µM GLP-26</i>	76	794	63
<i>10 µM CAM-A</i>	125	794	92
<i>100 µM CAM-A</i>	2612	4894	2010
<i>20 µM CAM-E</i>	53	203	36
<i>200 µM CAM-E</i>	74	334	44
<i>90 µM RS-3427</i>	191	4416	36

2)

<i>Conditions d'assemblage Cp183-EEE</i>	<i>Diamètre hydrodynamique (nm) selon l'intensité des particules</i>		<i>Diamètre hydrodynamique (nm) selon le nombre de particules</i>
<i>Sans modulateurs</i>	45	1286 4809	32
<i>90 µM GLP-26</i>	49	240	34
<i>10 µM CAM-A</i>	115	913	88
<i>100 µM CAM-A</i>	1307		1262
<i>20 µM CAM-E</i>	55	230	41
<i>200 µM CAM-E</i>	160		39
<i>90 µM RS-3427</i>	51	224	39

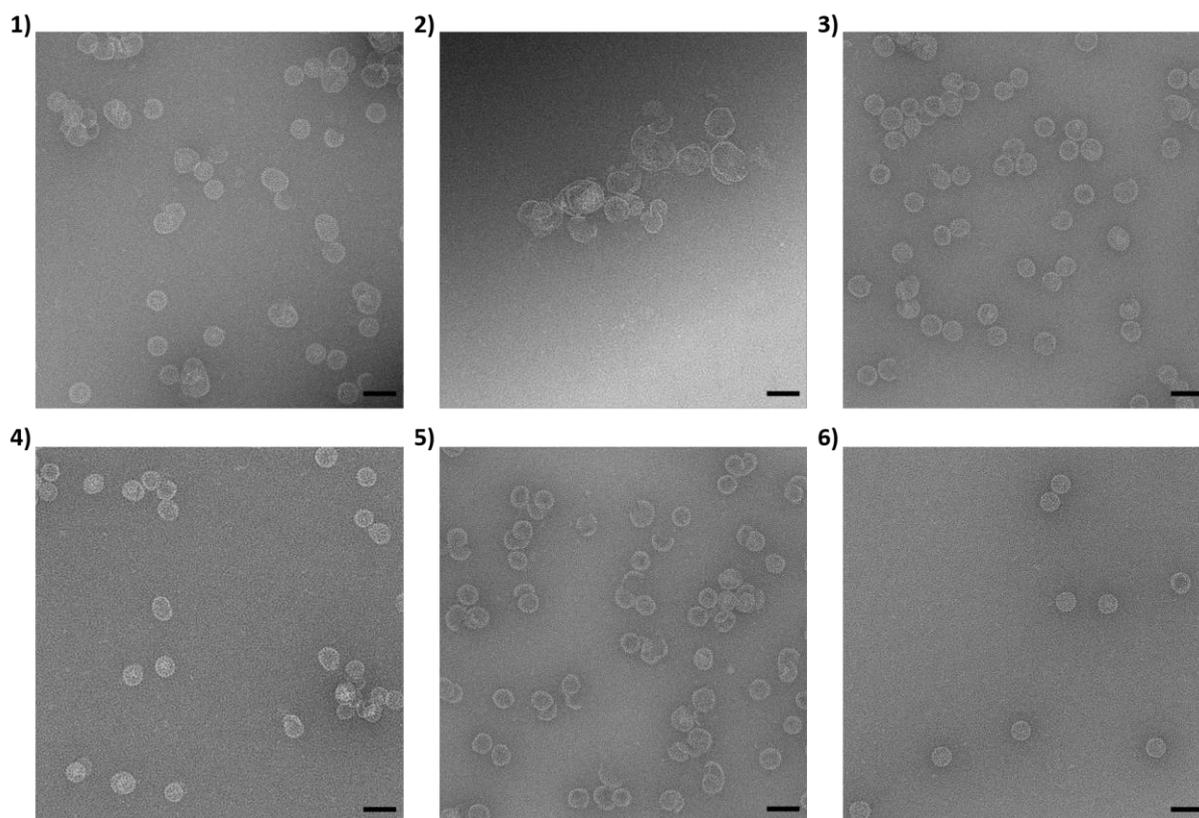
**Tableau 5.1 : Diamètre hydrodynamique selon l'intensité et le nombre de particules mesurées.**

1) Diamètre hydrodynamique des objets formés lors de l'assemblage de la Cp183 en présence ou en l'absence de modulateurs. 2) Diamètre hydrodynamique des objets formés lors de l'assemblage de la Cp183-EEE en présence ou en l'absence de modulateurs. Les valeurs surlignées en bleu indiquent la présence d'agrégats.

12 µM de Cp183 ou de Cp183-EEE ont été assemblés par dialyse, en présence de 95 nM d'ARNpg et de plusieurs modulateurs d'assemblage différents : Bay 41-4109, un modulateur de la classe A, JNJ-632 un modulateur de la classe E, GLP-26 et RS-3427, deux nouvelles molécules antivirales. Les objets formés ont ensuite été mesurés par DLS à 18°C. Le tableau ci-dessus indique les valeurs obtenues pour le diamètre hydrodynamique des objets présents dans chaque échantillon. Nous

remarquons que dans toutes les conditions, de gros objets voire des agrégats sont présents. Pour l'analyse des résultats, nous avons choisi de nous concentrer sur la distribution de taille calculée en fonction du nombre de particules présentes dans l'échantillon. Selon notre point de vue, les valeurs obtenues dans ce cas, sont plus proches de la réalité que celles obtenues en se basant sur l'intensité des particules présentes dans l'échantillon. Le calcul selon l'intensité, nous permet de déterminer s'il y a des agrégats dans nos échantillons et le calcul selon le nombre de particules, d'estimer le diamètre hydrodynamique des objets majoritaires dans cet échantillon.

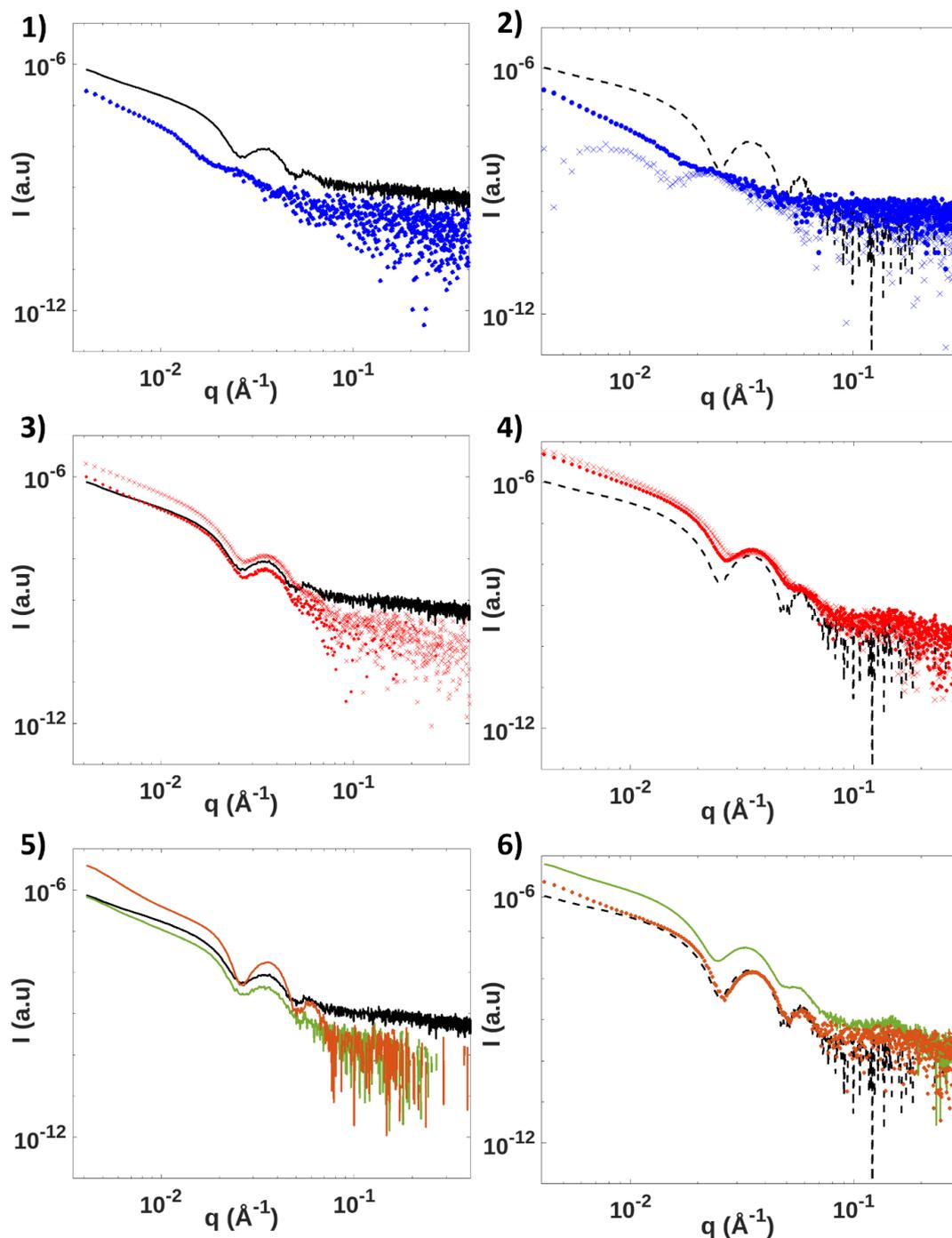
En présence de la CAM-A, les objets formés à partir des deux protéines n'ont plus la morphologie de capsides. Ils sont plus gros en présence de 10  $\mu\text{M}$  de la CAM-A et sont complètement agrégés en présence de 100  $\mu\text{M}$  de la CAM-A. En présence de la CAM-E, les objets formés sont légèrement plus gros, mais les tailles restent similaires aux tailles observées pour les objets formés sans modulateurs. Le diamètre estimé varie entre 36 nm et 44 nm selon la protéine et la concentration en CAM-E utilisée lors de l'assemblage. Les mêmes résultats sont également observés lors d'un assemblage en présence de la molécule RS-3427. Le diamètre hydrodynamique des objets formés à partir de la Cp183 est environ 36 nm et celui des capsides formées à partir de la Cp183-EEE est environ 39 nm. En présence de la molécule GLP-26, des effets différents sont observés. Lors de l'assemblage en présence de la protéine Cp183, les objets formés sont beaucoup plus gros qu'une capsidie formée sans modulateurs. Le diamètre hydrodynamique est environ 63 nm. Cependant, lors de l'assemblage à partir de la Cp183-EEE, les objets formés ont une taille similaire aux objets formés sans modulateurs. Le diamètre hydrodynamique des objets est environ 34 nm pour cette condition d'assemblage.



**Figure 5.7 : Micrographies des objets formés lors de l’assemblage de la Cp183 et observés en microscopie électronique par coloration négative.** 12  $\mu\text{M}$  de dimères de Cp183 ont été assemblés avec 95 nM d’ARNpg en présence de modulateurs. Assemblage en présence de 1) 10  $\mu\text{M}$  de CAM-A, 2) 100  $\mu\text{M}$  de CAM-A, 3) 20  $\mu\text{M}$  de CAM-E, 4) 200  $\mu\text{M}$  de CAM-E 5) 90 $\mu\text{M}$  de GLP-26 et 6) 90  $\mu\text{M}$  de RS-3427. Les échantillons sont observés sur un microscope JEOL JEM-2010 et les images enregistrées avec une caméra GATAN K2. Les barres d’échelles sont égales à 50 nm.

Les objets formés à partir de la Cp183 sont ensuite analysés par microscopie électronique, en coloration négative. Les images ci-dessus indiquent que la présence des modulateurs lors de l’assemblage influence la structure des objets formés. L’assemblage de la Cp183 en présence de 10  $\mu\text{M}$  de la CAM-A (Fig. 5.7.1) conduit à la formation d’une population d’objets très hétérogène. Certaines capsides ont la taille et la morphologie classique des capsides du VHB tandis que certaines sont plus grandes, avec des tailles qui varient entre 35 et 60 nm. En présence de 100  $\mu\text{M}$  de la CAM-A (Fig. 5.7.2), les capsides formées sont fortement endommagées et agrégées. La présence de capsides endommagées est également observée sur les micrographies montrant des capsides formées en présence de la CAM-E et de GLP-26 (Fig. 5.7.3 à 5.7.5). Nous observons également la présence de capsides légèrement plus grandes et plus allongées dans ces conditions. Enfin, l’assemblage

en présence de la molécule RS-3427 (Fig. 5.7.6) conduit à la formation de capsides très similaires aux capsides formées lors d'un assemblage classique. Toutefois, il est à noter que le nombre d'objets retrouvés sur la grille pour cette condition d'assemblage est relativement faible.



**Figure 5.8 : Courbes SAXS des objets formés en présence des modulateurs.** Courbes SAXS expérimentales des capsides formées à partir de  $12 \mu\text{M}$  de dimères de Cp183 (ligne pleine noire, graphique 1, 3, 5) et de  $12 \mu\text{M}$  de dimères de Cp183-EEE (ligne pointillée noire, graphique 2, 4, 6) en présence de  $95 \text{ nM}$  d'ARNpg et de  $10 \mu\text{M}$  (rond bleu) ou  $100 \mu\text{M}$  (croix bleu) de la CAM-A,  $20 \mu\text{M}$  (rond

rouge) ou 200  $\mu\text{M}$  (croix rouge) de la CAM-E, 90  $\mu\text{M}$  de GLP-26 (ligne verte) ou 90  $\mu\text{M}$  de RS-3421 (ligne orange).

Enfin, ces objets ont été analysés par SAXS. La figure X montre les courbes SAXS des objets formés à partir de la Cp183 ou de la Cp183-EEE en présence d'ARNpg et des différents modulateurs. En présence de la CAM-A (Fig. 5.8.1 et 5.8.2), aucune oscillation n'est observée. Cela traduit l'absence de capsides dans les échantillons analysés. Les intensités mesurées sont également beaucoup plus faibles que celles des capsides classiques. Après l'assemblage par dialyse, les échantillons sont centrifugés afin d'éliminer les agrégats. Une quantité élevée d'agrégats réduira la quantité et donc la concentration de protéines mesurées dans l'échantillon. En présence des modulateurs, les résultats obtenus à partir des mesures DLS ont révélé la formation d'agrégats lors de l'assemblage. Cela signifie que les valeurs d'intensité mesurées par SAXS ne peuvent être analysées de manière fiable, car elles dépendront de la quantité d'agrégats formés puis éliminés après la centrifugation. Il est important de noter que la quantité d'agrégats formés varie probablement selon le type de modulateurs utilisé mais également selon sa concentration. Pour la suite de l'analyse, nous nous intéressons donc uniquement à la forme des oscillations et à la valeur minimum du nombre d'onde  $q$  de la première oscillation.

En présence de la CAM-E (Fig. 5.8.3 et 5.8.4), les courbes obtenues sont similaires à celle de la capsid Cp183 classique. Cependant, les oscillations se déplacent vers les grandes valeurs de  $q$  et cet effet est accentué avec l'augmentation de la concentration en modulateur. En présence de la molécule RS-3427 (Fig. 5.8.5 et 5.8.6), l'assemblage de la Cp183-EEE conduit à la formation d'objets identiques à ceux formés lors de l'assemblage classique. En présence de la Cp183, les oscillations sont mieux définies, indiquant que les objets formés semblent plus structurés. La courbe est très similaire à celle obtenue lors de l'assemblage de la Cp183-EEE. Enfin, en présence de la molécule GLP-26 (Fig. 5.8.5 et 5.8.6), les oscillations sont moins profondes, et se déplacent vers les petites valeurs de  $q$  lors de l'assemblage avec les deux protéines.

Dans toutes les conditions, l'augmentation de l'intensité aux petits angles ne nous a pas permis d'approfondir l'analyse de ces résultats.

Ces résultats montrent que les modulateurs d'assemblage ont bien des effets sur la structure des capsides de Cp183 et de Cp183-EEE. Ils conduisent majoritairement à la formation de capsides aberrantes, voire d'agrégats. Les modulateurs de la classe E conduisent à la formation de capsides vides, mais la technique utilisée ici ne permet pas de déterminer si les capsides formées contiennent ou non de l'ARN pré-génomique.

## 5. Conclusion du chapitre

Dans ce chapitre, nous avons analysé les effets de l'ARN pré-génomique mais également des modulateurs d'assemblage sur l'auto-assemblage de la capside Cp183 et de la capside Cp183-EEE du VHB. Nos approches utilisées ici ont permis d'apporter des premiers éléments de réponses sur les effets de ces différentes molécules sur la cinétique mais également la structure des capsides du VHB.

Dans un premier temps, nous avons établi qu'une forte concentration en chlorure de guanidinium et la conservation à basse température étaient des éléments nécessaires pour maintenir la protéine Cp183-EEE sous une forme dimérique et empêcher une formation spontanée de capsides. Puis, dans un deuxième temps, nous avons étudié l'assemblage de la Cp183 et de la Cp183-EEE en présence d'ARNpg. Nos résultats ont démontré que l'ARNpg induit des modifications de la structure et de la cinétique d'assemblage des capsides formées à partir de la Cp183-EEE à 15°C. En présence d'ARNpg, l'assemblage des capsides semble s'accélérer et les objets formés sont légèrement plus petits. Par la suite, lorsque l'assemblage est réalisé par dialyse à température ambiante, les objets formés en présence de d'ARNpg et de Cp183 ou de Cp183-EEE ont la même morphologie. Les données obtenues via l'expérience de SAXS nous ont également permis de conclure que, contrairement à un assemblage de la Cp149 dans des conditions classiques, l'assemblage des protéines Cp183 et Cp183-EEE en présence d'ARNpg conduit dans notre contexte à la formation de capsides de type T=3.

Ensuite, nous avons mis en évidence les effets des modulateurs d'assemblage sur le processus d'auto-assemblage *in vitro* des capsides de Cp183 et de Cp183-EEE. Tout comme lors de l'assemblage de la Cp149 en présence des modulateurs, la CAM-A conduit à une modification drastique de la morphologie des capsides de la Core entière. De même, la CAM-E et la molécule GLP-26 conduisent également à une

altération structurale de ces capsides. Enfin, la molécule RS-3427 semble avoir des effets comparables à ceux observés lors de l'assemblage de la Cp149. Elle conduit à la formation d'un nombre plus faible de capsides très similaires à celles observées lors d'un assemblage *in vitro* de la capside Core entière sans modulateurs.

Il convient néanmoins de noter que les résultats obtenus ici sont des résultats préliminaires et pas optimaux. La présence d'agrégats dans toutes les conditions d'assemblage, que ce soit par mélange rapide ou par dialyse, en présence ou en absence d'ARNpg et de modulateurs, ne permet pas une analyse approfondie de la cinétique d'assemblage et de la structure ou de la morphologie des objets formés. De plus, l'assemblage en présence ou en absence d'ARNpg conduit à la formation de capsides T=3. Cependant, *in vivo*, la capside de HBV est une capside T=4. Des expériences d'optimisations de la stabilité des dimères de Cp183 et de Cp183-EEE dans des tampons contenant moins de chlorure de guanidine ainsi que de l'assemblage *in vitro* seront donc nécessaires pour approfondir ces observations.

## 6. Matériels et méthodes

Tous les protocoles utilisés pour exprimer, purifier et assembler la protéine Cp183-EEE sont les mêmes que ceux utilisés pour la protéine Cp183.

### Synthèse de l'ARN pré-génomique

#### Préparation du plasmide linéaire pCHT-3101

Le plasmide pCHT-3101 nous a été fourni par Michael Nassal (Université de Freiburg).

10 µl de bactéries contenant le plasmide sontensemencés dans 5 ml de LB contenant 75 µg/ml d'ampicilline. La préculture est ensuite incubée à 37°C et sous agitation à 200 rpm toute la nuit. Le lendemain matin, après avoir vérifié que la préculture a poussé, 3 ml de la préculture sont dilués dans 300 ml de LB contenant 75 µg/ml d'ampicilline. Le milieu est mis en culture toute la nuit à 37°C et sous agitation à 200 rpm. Les cultures sont ensuite centrifugées à 6000 g à 4°C pendant 30 minutes et le surnageant est éliminé.

L'ADN est extrait des bactéries en réalisant une maxipreparation à l'aide du kit NucleoBond Xtra Midi conçu pour la purification rapide de plasmides. Le protocole est le suivant :

- 1- Re-suspendre le culot bactérien avec 12 ml de tampon de re-suspension RES + RNase A.
- 2- Lyser les bactéries avec 12 ml du tampon de lyse. Remuer doucement 5 fois, en renversant le tube contenant le culot. Incuber le mélange pendant 5 minutes à température ambiante.
- 3- Pendant ce temps, équilibrer la colonne NucleoBond Xtra contenant une résine de silice et un filtre permettant l'élimination des précipités, avec 25 ml de tampon d'équilibration EQU.
- 4- Ajouter 12 ml de tampon EQU dans le tube contenant les bactéries. Mélanger immédiatement le lysat en renversant le tube jusqu'à la disparition de la couleur bleue et la formation d'un précipité blanc.
- 5- Mélanger 3 fois le tube puis filtrer la solution sur la colonne. On peut éventuellement centrifuger le lysat à 5000 g pendant 10-20 minutes pour enlever le précipité blanc.

- 6- Laver le filtre avec 15 ml de tampon EQU puis retirer le filtre.
- 7- Laver la colonne avec 25 ml de tampon de lavage.
- 8- Eluer l'ADN plasmidique avec 15 ml de tampon d'éluion ELU. Collecter l'éluât dans un tube de centrifugation.
- 9- Ajouter 10,5 ml d'isopropanol à l'éluât et mélanger à l'aide d'un vortex. Centrifuger 40 minutes, à 15 000 g et 4°C.
- 10- Retirer le surnageant et laver le culot avec 14 ml d'éthanol à 70% (préparé avec de l'eau RNase free). Centrifuger la solution pendant 5 minutes à 15 000 g et à température ambiante.
- 11- Eliminer l'éthanol et sécher le culot sur la paillasse à température ambiante pendant 30 minutes.
- 12- Reprendre le culot avec 600 µl d'eau RNase free et laisser l'ADN se dissoudre. Mesurer la concentration en ADN plasmidique au NanoDrop.
- 13- Digérer 1 g/l d'ADN plasmidique avec 1X de tampon CutSmart et l'enzyme Xma1 pendant 2h à 37°C. Vérifier la digestion sur gel d'agarose. L'ADN est ensuite purifié.

### **Purification de l'ADN plasmidique linéarisé**

La purification de l'ADN plasmidique est réalisée de la manière suivante :

- 1- Ajouter un volume identique d'ADN plasmidique et de phénol/chloroforme à pH 7.9 dans un tube Eppendorf et bien mélanger.
- 2- Centrifuger la solution à température ambiante pendant 5 minutes à 14 000 rpm.
- 3- Récupérer la phase aqueuse supérieure, la transférer dans un nouveau tube et y ajouter un volume égal de phénol/chloroforme. Bien mélanger.
- 4- Centrifuger la solution pendant 5 minutes à 14 000 rpm et à température ambiante.
- 5- Récupérer la phase aqueuse supérieure et la transférer dans un nouveau tube.
- 6- Ajouter 2,5 volumes d'éthanol 100 % et 1/10 du volume final d'acétate de sodium à 3 M pH 5.2 pour précipiter l'ADN.
- 7- Incuber la solution à -80°C pendant 20 minutes et centrifuger à 15 000 rpm pendant 20 minutes à 4°C.

- 8- Retirer le surnageant et re-suspendre le culot d'ADN avec 800 µl d'éthanol 70 %.
- 9- Centrifuger la solution à 14 000 rpm pendant 10 minutes et à 4°C.
- 10-Retirer le surnageant et sécher le culot d'ADN pendant 20 minutes sous la hotte. Re-suspendre l'ADN avec de l'eau RNase free.
- 11-Mesurer la concentration en ADN et si la valeur du ratio  $A_{260}/A_{280}$  n'est pas entre 1,7 et 1,8, refaire l'extraction au phénol/chloroforme.
- 12-Ajuster la concentration à 1 g/l en ajoutant la quantité appropriée d'eau RNase free.
- 13-Vérifier l'intégrité de l'ADN grâce à une électrophorèse sur gel d'agarose 0,8%.

### **Transcription et purification de l'ARN pré-génomique**

La transcription de l'ARN pré-génomique est réalisée de la manière suivante :

- 1- Ajouter les différentes solutions de nucléotides (ATP, CTP, GTP et UTP) dans un tube contenant l'ADN linéarisé à température ambiante.
- 2- Ajouter ensuite l'eau RNase free ainsi que le tampon de réaction.
- 3- Mélanger vigoureusement puis incuber pendant 4 à 6 h à 37°C.
- 4- Ajouter 1 µl de TURBO DNase par réaction dans le tube contenant l'ARN pré-génomique.
- 5- Incuber pendant 15 minutes à 37°C.
- 6- Ajouter ensuite 25 mM d'acétate d'ammonium pour stopper la réaction et mélanger.

L'ARN pré-génomique est ensuite purifié en suivant le protocole ci-dessous :

- 1- Ajouter un volume équivalent de phénol/chloroforme/alcool isoamyle pH 4,7 et bien mélanger.
- 2- Centrifuger le mélange 12 minutes à température ambiante et 12 000 g.
- 3- Récupérer la phase aqueuse supérieure et la mélanger avec un volume de chloroforme.
- 4- Centrifuger à nouveau 2 minutes à température ambiante et récupérer la phase aqueuse qui contient l'ARN.
- 5- Mélanger cette phase avec 1 volume équivalent d'isopropanol et laisser reposer 15 minutes à -80°C.
- 6- Centrifuger ensuite le mélange pendant 15 minutes à 4°C et 12 000 g.

- 7- Retirer le surnageant et laisser sécher le culot d'ARN.
- 8- Re-suspendre le culot avec du tampon 50 mM Tris HCl pH 7,5 RNase free.
- 9- Mesurer la concentration en ARN pré-génomique au NanoDrop.
- 10-Réaliser une électrophorèse sur gel d'agarose 0,8 % afin de vérifier l'intégrité de l'ARN.
- 11-Aliquoter et conserver l'ARN pré-génomique à -80°C.

### **Electrophorèse sur gel d'agarose**

Faire fondre 0,4 g d'agarose dans 50 ml de tampon TAE 1X, laisser refroidir légèrement et y ajouter 2 µl de SYBR Safe. Mélanger et verser la solution dans la cassette. Ajouter le peigne formant les puits et laisser le gel se solidifier.

Préparer les échantillons en mélangeant 1 µg à 2 µg d'ARN avec un volume équivalent de tampon. Incuber pendant 15 min à 70°C puis déposer environ 10 µL d'échantillons dans chaque puit et faire migrer pendant 1 h à 120 V. La lecture du gel se fait à l'aide du logiciel Gel doc go de Biorad.

### **Expression des capsides**

Les étapes de préculture et de culture sont exactement les mêmes que pour l'expression de la Cp149. Cependant, après l'induction, les cultures sont incubées à 18°C au lieu de 25°C.

### **Purification des capsides**

L'étape de purification des capsides de Cp183 et de Cp183-EEE est la même que pour les capsides de Cp149. Cependant, la composition du tampon de lyse est légèrement différente. Il s'agit ici d'un tampon TN50 composé de 50 mM Tris HCl pH 7,5, 50 mM NaCl, 5 mM EDTA et 5 mM DTT. La précipitation au sulfate d'ammonium est réalisée avec une saturation de 60%.

### **Purification des dimères**

La solution contenant les capsides est dialysée deux fois pendant 1 h, puis toute la nuit sous agitation à 4°C contre un tampon de désassemblage composé de 20 mM Tris HCl pH 7,5, 2 M chlorure de guanidinium (pour précipiter l'ARN), 0,5 M de chlorure de lithium et 10 mM DTT.

Après la dialyse, la concentration de l'échantillon est mesurée au NanoDrop et les agrégats qui forment un précipité blanc sont éliminés. Pour cela, l'échantillon est centrifugé à 1500 g et 4°C pendant 8 minutes. Le surnageant est ensuite récupéré, filtré et la concentration en protéines est à nouveau mesurée.

La solution est ensuite injectée sur une colonne d'exclusion de taille (la Superdex S200 10/300) préalablement équilibrée avec le tampon de désassemblage. Les fractions contenant les dimères sont récupérées et leur concentration est mesurée au NanoDrop. Les fractions pures, ne contenant plus de matériel génétique sont rassemblées. La concentration finale en protéines est mesurée et la taille des dimères ainsi que l'homogénéité de la solution sont vérifiés par DLS. Les protéines sont ensuite aliquotées et congelées à -80°C jusqu'à leur utilisation.

### **Assemblage des capsides en présence ou non des CAMS**

Après décongélation des dimères à 4°C, ceux-ci sont dialysés contre un tampon contenant 1,5 M de chlorure de guanidinium, 2 mM DTT et 10 mM CHES pH 9 sous agitation deux fois pendant 1 h à 4°C. Puis les dimères sont filtrés à l'aide d'un filtre de 0,22 µm et leur concentration est mesurée.

Plusieurs conditions d'assemblages sont testées et l'assemblage (sauf pour les cinétiques) est réalisé par dialyse.

La méthode d'assemblage consiste à mélanger tous les composants (la protéine, en présence ou non d'ARN pré-génomique et en présence ou non de modulateurs d'assemblage) dans une cassette de dialyse de seuil de coupure 3,5 kDa et de les dialyser contre un tampon contenant 50 mM Tris HCl pH 7,5 et 150 mM d'acétate d'ammonium à température ambiante pendant 1 h puis toute la nuit après avoir changé le bain de dialyse. Après la dialyse, la solution est retirée de la cassette et centrifugée à 8000 rpm et 4°C pendant 10 min pour retirer les agrégats. La concentration des objets est ensuite mesurée au NanoDrop.

### **DLS et méthodes structurales**

Le protocole de mesure des objets par DLS est le même que celui de la Cp149.

Les protocoles de mesure des objets par SAXS et d'observation en microscopie électronique sont les mêmes que ceux des capsides de Cp149.

## **TR-SAXS**

Pour les expériences de TR-SAXS, avant la manipulation, les dimères de Cp183-EEE sont dialysés contre le tampon contenant 1,5 M de chlorure de guanidinium, 2 mM DTT et 10 mM CHES pH 9 sous agitation deux fois pendant 1 h à 4°C puis concentrés à 1 g/l. Lors de l'expérience, le détecteur est placé à une distance de 2 mètres. L'assemblage est déclenché à l'aide d'un stopped-flow en mélangeant la solution contenant les dimères de Cp183-EEE, le tampon de la protéine, l'ARNpg et un tampon contenant 50 mM Tris HCl pH 7,5. Chaque cinétique est réalisée au minimum 2 fois à 15°C et dure 15 minutes. Des mesures additionnelles sont également réalisées en fin de cinétique.

## Bibliographie

Schlicksup, Christopher John, Patrick Laughlin, Steven Dunkelbarger, Joseph Che-Yen Wang, et Adam Zlotnick. 2020. « Local Stabilization of Subunit–Subunit Contacts Causes Global Destabilization of Hepatitis B Virus Capsids ». *ACS Chemical Biology* 15 (6): 1708-17. <https://doi.org/10.1021/acscchembio.0c00320>.

Schlicksup, Christopher John, Joseph Che-Yen Wang, Samson Francis, Balasubramanian Venkatakrishnan, William W Turner, Michael VanNieuwenhze, et Adam Zlotnick. 2018. « Hepatitis B virus core protein allosteric modulators can distort and disrupt intact capsids ». Édité par Karla Kirkegaard. *eLife* 7 (janvier): e31473. <https://doi.org/10.7554/eLife.31473>.

Stray, S. J., C. R. Bourne, S. Punna, W. G. Lewis, M. G. Finn, et A. Zlotnick. 2005. « A Heteroaryldihydropyrimidine Activates and Can Misdirect Hepatitis B Virus Capsid Assembly ». *Proceedings of the National Academy of Sciences* 102 (23): 8138-43. <https://doi.org/10.1073/pnas.0409732102>.

## Chapitre 6 : Conclusions et perspectives

Durant ces trois années, j'ai étudié les effets des modulateurs d'assemblage sur la structure et la cinétique d'assemblage des capsides du virus de l'hépatite B. Pour cela, j'ai utilisé plusieurs techniques, telles que le TR-SAXS et la cryomicroscopie électronique en transmission. Cette étude a été réalisée conjointement avec l'Institut de Biologie Intégrative de la Cellule et le Laboratoire de Physique des Solides, faisant intervenir les synchrotrons SOLEIL et l'ESRF ainsi que la plateforme de cryomicroscopie électronique de l'Institut Pasteur. Cette thèse avait deux objectifs principaux :

- Déterminer quelles sont les perturbations apportées par les modulateurs aux chemins cinétiques d'assemblage en capsides de la Cp149 et identifier les modifications structurales induites par ces molécules sur les capsides formées dans ces conditions.
- Etablir l'effet du CTD sur l'assemblage de la capside et la réponse aux CAMs en étendant ces études à deux formes de la protéine Core incluant le CTD et mimant deux états de phosphorylation vis-à-vis de la formation de la capside : la Cp183 et la Cp183-EEE.

L'étude de l'assemblage de la capside du virus de l'hépatite B formée à partir de la protéine Cp149 en présence des modulateurs par DLS, photométrie de masse, (TR) SAXS et cryomicroscopie électronique en transmission m'a permis de répondre au premier objectif. Cette version recombinante de la protéine Core contient uniquement le domaine NTD responsable de l'autoassemblage. Dans un premier temps, j'ai confirmé qu'il est possible d'utiliser un autre sel que le chlorure de sodium pour déclencher l'assemblage de la Cp149. Puis j'ai montré que, les modulateurs d'assemblage accélèrent la cinétique d'assemblage de la capside et modifient la structure des capsides formées. Les modulateurs de type A conduisent bien à la formation d'objets aberrants, très gros par rapport à une capside de triangulation  $T=4$ . De plus, nous avons révélé que contrairement à ce qui est énoncé dans la littérature scientifique, les CAM-E tels que JNJ-632 et GLP-26 peuvent également altérer la structure de la capside de Cp149, mais de manière plus subtile que les CAM-A (Kra et al. 2023). Enfin, j'ai démontré que la molécule RS-3427, étudiée dans le cadre d'une

collaboration avec l'équipe de Raymond Schinazi à l'université d'Emory aux Etats-Unis, serait un nouveau type de modulateur.

Les effets de l'ARNpg et des CAMs sur l'assemblage de deux protéines : la cp183 et la cp183-EEE, ont ensuite été étudiés. Ces deux protéines contiennent le domaine NTD mais également le domaine CTD qui interagit avec l'ARNpg et est nécessaire pour son encapsidation. De plus, la protéine Cp183-EEE qui est un mutant de la protéine Cp183, mime un état de phosphorylation de la capsidite ressemblant à celui des nucléocapsides immatures, après l'assemblage mais avant la transcription inverse. Les résultats obtenus ont apporté des premiers éléments de réponse concernant le deuxième objectif de ce projet. Ils ont révélé que l'ARNpg induit des changements structuraux des capsides formées et accélère la cinétique d'assemblage de la Cp183-EEE. Cependant, contrairement à ce qui est montré dans la littérature, les capsides formées ici, sont des capsides de triangulation  $T = 3$ , sans doute à cause des conditions physicochimiques utilisées qui étaient trop éloignées de celles prévalant dans le cytoplasme. Ces capsides, qui sont composées de 180 dimères, sont donc plus petites que les capsides  $T = 4$  du virus natif de l'hépatite B. Puis j'ai montré que les modulateurs d'assemblage perturbent également les capsides formées à partir des protéines Cp183 ou Cp183-EEE en présence d'ARNpg. Enfin, lors d'une collaboration avec l'équipe de Fabien Montel à l'ENS Lyon, j'ai confirmé que les CAM-A peuvent également perturber les capsides de Cp183 déjà formées.

Toutes les techniques utilisées durant cette thèse sont très complémentaires et permettent d'obtenir de nombreux éléments de réponse sur les effets des CAMs sur l'autoassemblage de la capsidite du virus de l'hépatite B. Dans un premier temps, les techniques de caractérisation optique m'ont permis d'avoir une première hypothèse sur l'effet des modulateurs sur la taille et la masse des capsides du virus de l'hépatite B. Ensuite, l'utilisation du SAXS et du TR-SAXS a permis d'obtenir les premières informations structurales des effets des CAMs mais également des informations précises sur la cinétique d'assemblage des capsides en présence des CAMs. Enfin, la cryomicroscopie électronique en transmission a permis d'affiner notre vision de ces modifications structurales, observées grâce aux techniques de caractérisation optique et au SAXS.

Cependant, de nombreuses questions restent à clarifier.

Tout d'abord, dans le cas de l'assemblage de la capsid Cp149 en présence du modulateur de type A, la formation d'objets de tailles et de formes différentes est observée. Ces objets, dont la cinétique d'assemblage comporte une étape tardive, très lente n'ont pas été pleinement analysés. Une étude par TR-SAXS de la cinétique d'assemblage sur une durée d'acquisition plus longue, aurait facilité l'identification d'intermédiaires d'assemblage stables mais également d'autres phases du chemin d'assemblage. De plus, une étude par cryotomographie électronique aurait amélioré la caractérisation structurale de ces objets. Elle aurait apporté des informations sur le degré de facettage de ces objets, permis de vérifier l'uniformité de toutes les faces, ou encore de vérifier que la structure du dimère de dimère n'y est pas altérée, comme le suggèrent nos données SAXS.

Puis, en ce qui concerne l'assemblage de la capsid Cp149, en présence des deux modulateurs de type E, JNJ-632 et GLP-26, des intermédiaires précoces ou tardifs n'ont pas été détectés lors des expériences de TR-SAXS. D'après l'analyse de décomposition en valeur singulière, seulement deux types d'espèces sont présentes lors de l'assemblage en présence de JNJ-632. Nous pouvons donc supposer qu'il s'agit des dimères et des capsides de Cp149. Cependant, cette analyse n'a pas été effectuée pour l'assemblage en présence de GLP-26. Ainsi, nous ne pouvons pas conclure sur la présence de plusieurs espèces lors de l'assemblage avec GLP-26 et, le cas échéant, lesquelles.

Ensuite, lors de l'assemblage de la Cp149 en présence de la molécule RS-3427, les résultats obtenus suggèrent que cette molécule serait un nouveau type de modulateur. Notre recherche ne permet pas de conclure sur l'interaction entre ce modulateur et les dimères. Mais, du fait de sa structure chimique, très similaire à la molécule GLP-26, nous pouvons faire l'hypothèse que cette molécule interagit avec les dimères de Cp149. Dans ce cas, elle pourrait conduire à la formation d'un grand nombre de noyaux composés d'un faible nombre de dimères, diminuant ainsi le nombre de dimères disponibles pour la formation de capsides complètes, ou au contraire modifier l'arrangement inter-dimères et ainsi empêcher totalement la formation de capsides par les dimères auxquels elle se fixe. Dans la première hypothèse, une seule molécule RS-3427 ou un petit nombre de ces molécules seraient présentes dans les capsides formées. Dans la seconde, elles en seraient complètement absentes. Notre

reconstruction à l'échelle quasi-atomique des capsides formées en présence de RS-3427 établit que ces capsides sont parfaitement icosaédriques, mais ne nous a pas permis de conclure pour l'instant quant à la présence ou l'absence de RS-3427. Des expériences à l'échelle de la molécule unique pourraient apporter une meilleure compréhension du mécanisme d'action de cette molécule.

De plus, des reconstructions atomiques à haute résolution devront être réalisées afin de caractériser le site de fixation des molécules GLP-26 et RS-3427 sur la capside du virus de l'hépatite B. Pour cela, la collecte d'un nombre conséquent de micrographies de cryomicroscopie électronique, et donc d'un nombre suffisant de particules est nécessaire afin de faire ces reconstructions sans imposer de symétrie. En effet, il est vraisemblable que ces molécules ne se fixent pas sur la capside en suivant une symétrie icosaédrique.

Enfin, dans le cas de l'assemblage de la protéine entière, mutée ou non, et en présence de l'ARNpg, la formation de capsides de triangulation  $T = 3$  et la présence d'agrégats quelles que soient les méthodes d'assemblage sont observées. Ces objets ont été assemblés par mélange progressif par dialyse à température ambiante ou par mélange rapide à l'aide d'un stopped flow à 15°C. *In vivo*, les capsides du virus de l'hépatite B sont des capsides  $T = 4$  (Porterfield et al. 2010; Wang, Dhason, et Zlotnick 2012). Ce changement du nombre de triangulation pourrait s'expliquer par une modification de la cinétique d'assemblage. Il est possible que cette modification de la cinétique soit due à plusieurs facteurs. Parmi ces facteurs, nous pouvons mentionner une concentration trop élevée en protéines et trop faible en sel, l'utilisation du chlorure de guanidine ou encore un pH trop basique. Il serait ainsi judicieux d'effectuer des assemblages à 37 °C en utilisant des concentrations de protéines plus faibles et des concentrations en sel plus importantes, mais également des concentrations d'ARN pré-génomique différentes afin de modifier le rapport entre le nombre d'ARNpg et de dimères lors de l'assemblage. Puis, les capsides formées devront être analysées par SAXS afin de déterminer leur nombre de triangulation. De plus, il n'est pas possible de distinguer les capsides vides, des capsides contenant de l'ARN pré-génomique lors des expériences de SAXS. Des expériences de diffusion des neutrons aux petits angles pourraient également être envisagées afin de quantifier le ratio entre le nombre

de capsides vides et le nombre de capsides pleines selon les conditions d'assemblage.

L'ensemble des expériences de cette thèse a été réalisé avec de fortes concentrations de dimères de Core, de l'ordre du micromolaire et dans un milieu non encombré. Cependant, *in cellulo*, la concentration de la protéine Core est plus faible que celles utilisées dans nos expériences et l'assemblage de la capside a lieu dans le cytoplasme, qui est un milieu encombré. Ainsi, d'autres techniques facilitant une observation à l'échelle de la particule unique telle que la photométrie de masse (Wagner et al. 2023; Wu et al. 2022), mais également la microscopie à force atomique (Buzón et al. 2021), pourraient être employées pour une meilleure compréhension de l'assemblage de la capside du VHB dans des conditions plus proches de la réalité. En effet, l'utilisation de ces techniques permettrait de réaliser des assemblages avec une très faible concentration de dimères de Core. De plus, des assemblages dans un milieu plus visqueux, qui mimerait le cytoplasme permettraient d'étudier les effets de l'encombrement du milieu sur l'assemblage de la capside et l'influence de ces modulateurs sur l'assemblage dans ces conditions.

Les CAMs perturbant l'assemblage de la capside et conduisant à la formation d'objets différents, il serait intéressant de réaliser des études sur les effets des CAMs sur les propriétés physiques et mécaniques des capsides en réalisant une étude par *thermal shift assay* (Chevreuil et al. 2020; Senaweera et al. 2021) ou en mesurant la rigidité de ces objets par microscopie à force atomique.

Dans son ensemble, ce projet de thèse apporte les tout premiers éléments de réponse concernant l'influence des modulateurs d'assemblage sur la structure et la cinétique d'assemblage des capsides du virus de l'hépatite B. De plus, il met en évidence les limites de l'étude du mode de fonctionnement de ces modulateurs et les nombreuses questions encore non élucidées. La compréhension de ces mécanismes d'action devrait conduire à une amélioration des molécules synthétisées et donc à l'avenir à une amélioration des traitements proposés aux patients atteints du virus de l'hépatite B.

## Bibliographie

Buzón, Pedro, Sourav Maity, Panagiotis Christodoulis, Monique J. Wiertsema, Steven Dunkelbarger, Christine Kim, Gijs J. L. Wuite, Adam Zlotnick, et Wouter H. Roos. 2021. « Virus Self-Assembly Proceeds through Contact-Rich Energy Minima ». *Science Advances* 7 (45): eabg0811. <https://doi.org/10.1126/sciadv.abg0811>.

Chevreuril, M. *et al.* Nonsymmetrical dynamics of the HBV capsid assembly and disassembly evidenced by their transient species. *J. Phys. Chem. B* **124**, 9987–9995 (2020).

Kra, Kalouna, Siyu Li, Laetitia Gargowitsch, Jéril Degrouard, Javier Pérez, Roya Zandi, Stéphane Bressanelli, et Guillaume Tresset. 2023. « Energetics and Kinetic Assembly Pathways of Hepatitis B Virus Capsids in the Presence of Antivirals ». *ACS Nano* 17 (13): 12723-33. <https://doi.org/10.1021/acsnano.3c03595>.

Porterfield, J. Zachary, Mary Savari Dhason, Daniel D. Loeb, Michael Nassal, Stephen J. Stray, et Adam Zlotnick. 2010. « Full-Length Hepatitis B Virus Core Protein Packages Viral and Heterologous RNA with Similarly High Levels of Cooperativity ». *Journal of Virology* 84 (14): 7174-84. <https://doi.org/10.1128/JVI.00586-10>.

Senaweera, Sameera, Haijuan Du, Huanchun Zhang, Karen A. Kirby, Philip R. Tedbury, Jiashu Xie, Stefan G. Sarafianos, et Zhengqiang Wang. 2021. « Discovery of New Small Molecule Hits as Hepatitis B Virus Capsid Assembly Modulators: Structure and Pharmacophore-Based Approaches ». *Viruses* 13 (5): 770. <https://doi.org/10.3390/v13050770>.

Wagner, Christina, Felix F. Fuchsberger, Bernd Innthaler, Martin Lemmerer, et Ruth Birner-Gruenberger. 2023. « Quantification of Empty, Partially Filled and Full Adeno-Associated Virus Vectors Using Mass Photometry ». *International Journal of Molecular Sciences* 24 (13): 11033. <https://doi.org/10.3390/ijms241311033>.

Wang, Joseph C.-Y., Mary S. Dhason, et Adam Zlotnick. 2012. « Structural Organization of Pregenomic RNA and the Carboxy-Terminal Domain of the Capsid Protein of Hepatitis B Virus ». *PLOS Pathogens* 8 (9): e1002919. <https://doi.org/10.1371/journal.ppat.1002919>.

Wu, Di, Philsang Hwang, Tiansen Li, et Grzegorz Piszczek. 2022. « Rapid characterization of adeno-associated virus (AAV) gene therapy vectors by mass photometry ». *Gene therapy* 29 (12): 691-97. <https://doi.org/10.1038/s41434-021-00311-4>.

## **Annexes**

**Supporting Information : Energetics and kinetic assembly pathways of Hepatitis B virus capsids in the presence of antivirals**

# Supporting Information

## Energetics and kinetic assembly pathways of Hepatitis B virus capsids in the presence of antivirals

*Kalouna Kra,<sup>1,2</sup> Siyu Li,<sup>3</sup> Laetitia Gargowitsch,<sup>1</sup> Jéril Degrouard,<sup>1</sup> Javier Pérez,<sup>4</sup> Roya Zandi,<sup>3</sup> Stéphane Bressanelli<sup>2\*</sup> and Guillaume Tresset<sup>1\*</sup>*

<sup>1</sup>Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, 91405 Orsay, France.

<sup>2</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC),  
91198 Gif-sur-Yvette, France.

<sup>3</sup>Department of Physics and Astronomy, University of California, Riverside, California  
92521, United States.

<sup>4</sup>SOLEIL Synchrotron, 91192 Gif-sur-Yvette, France.

\*Corresponding authors : [stephane.bressanelli@i2bc.paris-saclay.fr](mailto:stephane.bressanelli@i2bc.paris-saclay.fr);

[guillaume.tresset@universite-paris-saclay.fr](mailto:guillaume.tresset@universite-paris-saclay.fr)

## Supporting Text

### Singular value decomposition (SVD) analysis

Let  $\mathbf{I}$  be a matrix in which the columns are the scattering intensities  $I(q)$  collected at successive time points  $t$ . The SVD provides a unique decomposition of  $\mathbf{I}$ :<sup>55</sup>

$$\mathbf{I} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (1)$$

where  $\mathbf{U}$  and  $\mathbf{V}$  are unitary matrices, and  $\mathbf{\Sigma}$  is a matrix with real, nonnegative singular values  $\Sigma_k$  on the diagonal and zeros off the diagonal, which are sorted in a rapidly decreasing order. It should be noted that the columns of  $\mathbf{U}$ , called basis vectors, have the same nature as scattering intensities even though they are not physically meaningful. The matrix of intensities  $\mathbf{I}$  can be approximated by retaining the first  $r$  columns of  $\mathbf{U}$  and  $\mathbf{V}$  as well as the first  $r$  largest singular values  $\Sigma_k$ , the resulting truncated matrix thus being the best least-square approximation of rank  $r$  of  $\mathbf{I}$ .

An optimal hard threshold for the singular values was proposed by Gavish and Donoho<sup>56</sup> in such a way that the truncated matrix approximates the original matrix within noise level. Here, instead of using directly  $\mathbf{I}$ , the scattering intensities were normalized to their uncertainties yielding a matrix  $\tilde{\mathbf{I}}$ , the elements of which had noise with unitary standard deviation. Notice that scattering intensities and uncertainties must be decorrelated as much as possible, otherwise a significant part of the information is lost.  $\tilde{\mathbf{I}}$  being a  $n$ -by- $m$  matrix, with  $m > n$ , the optimal threshold was given by

$$\Sigma_* = \lambda_*(\beta)\sqrt{n} \quad (2)$$

with  $\beta = m/n$  and

$$\lambda_*(\beta) = \sqrt{2(\beta + 1) + \frac{8\beta}{\beta + 1 + \sqrt{\beta^2 + 14\beta + 1}}} \quad (3)$$

being the optimal hard threshold coefficient. The optimal truncation rank  $r_*$  was deduced such that for any  $k \leq r_*$ ,  $\Sigma_k \geq \Sigma_*$ . It represents the number of basis vectors, and subsequently of singular values, required to approximate the scattering intensities within noise level.

### Classical nucleation theory

A classical nucleation theory was proposed by Zandi and coworkers<sup>41</sup> to account for the self-assembly kinetics of spherical viral capsids. A spherical cap is growing by sequential addition of subunits due to a favourable binding energy but, at the same time, is subject to an energy frustration arising from the line tension of the rim exposed to the solvent. The dimensionless free energy of the cap made of  $n$  subunits, in the limit of a large number of subunits in a complete shell  $N$ , reads

$$\beta\Delta G(n) = -n \ln \frac{c}{c_*} + 2\beta\gamma_{\text{rim}} \sqrt{\frac{n}{N} \left(1 - \frac{n}{N}\right)} \quad (4)$$

where  $\beta^{-1} = k_B T$  with  $k_B$  the Boltzmann constant and  $T$  the temperature,  $c$  the dimensionless concentration of subunits during assembly (note that  $c$  varies with time),  $c_*$  denotes the dimensionless critical aggregation concentration, and  $\gamma_{\text{rim}}$  is the line tension of the rim.  $c_*$  is related to the subunit binding energy  $\Delta g$  via  $c_* \sim \exp(\beta\Delta g)$ .  $\beta\gamma_{\text{rim}}$  can be related to the subunit binding energy by assimilating subunits with discs, which yields  $\beta\gamma_{\text{rim}} \sim (\pi/2)|\beta\Delta g|\sqrt{N}s$ ,<sup>42</sup> with  $s$  a parameter depending on the contact interface between subunits. Note that the disc hypothesis is not required since  $s$  is actually a fitting parameter and we could have chosen directly  $\beta\gamma_{\text{rim}}$  instead, it just allowed us to relate analytically the line tension of the rim to the subunit binding energy. Eq. 4 exhibits an energy barrier that gives rise to a nucleation process: once a spherical cap exceeds the critical nucleus size  $n_*$  for which the free energy is maximal,

it rapidly grows until its completion into a shell of  $N$  subunits. If the height of the energy barrier is sufficiently large compared to the thermal energy, i.e.,  $\beta\Delta G_* \gg 1$ , then the nucleation rate takes the following simple form:

$$J \sim k_{\text{on}} Z c e^{-\beta\Delta G_*} \quad (5)$$

with  $k_{\text{on}}$  the forward constant rate and  $Z$  the so-called Zeldovitch factor. Defining the supersaturation as

$$\Gamma = \frac{\sqrt{N} \ln(c/c_*)}{\pi |\beta\Delta g| s} \quad (6)$$

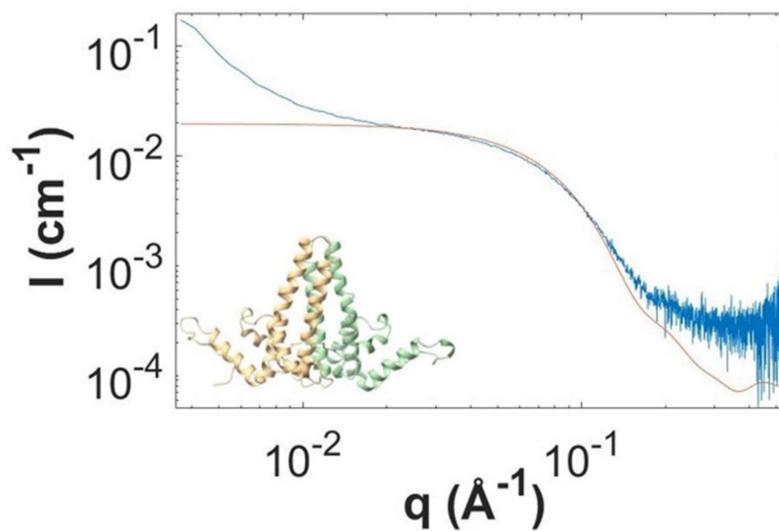
we can express the height of the energy barrier as well as the Zeldovitch factor as follows:

$$Z = \sqrt{\frac{|\beta\Delta g| s}{N^{3/2}}} (1 + \Gamma^2)^{3/4} \quad (7)$$

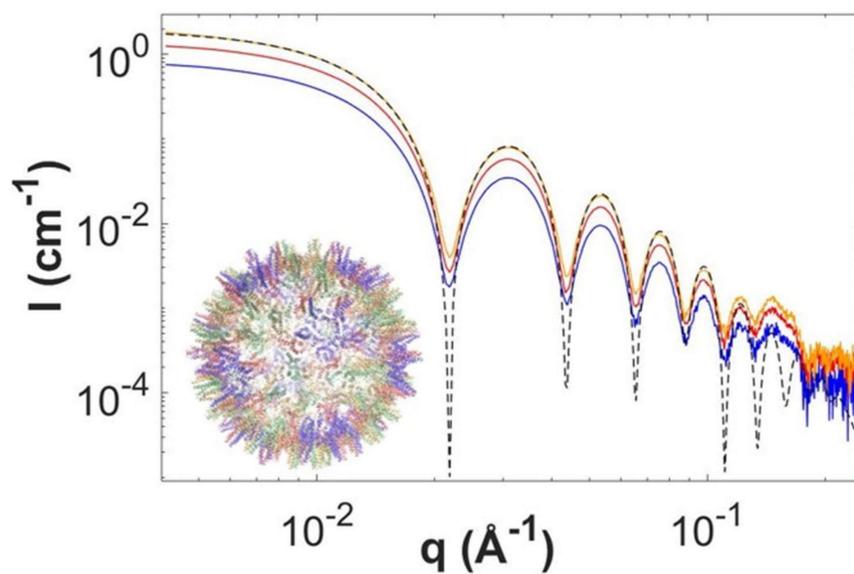
$$\beta\Delta G_* = \frac{\sqrt{N}}{2} \pi |\beta\Delta g| s (\sqrt{1 + \Gamma^2} - \Gamma) \quad (8)$$

The evolution of the dimensionless concentration of free subunits is then governed by  $dc/dt = -J$  and the fraction of subunits in capsids is deduced by mass conservation.

Figure S13 depicts the fractions of subunits in capsids  $\phi_m(t)$  calculated from the classical nucleation theory (dashed lines) for capsid assembly without modulator, as well as with CAM-E and CAM-A, and they are compared with the semi-experimental fractions  $\phi_{se}(t)$  (discs). The fitting parameters  $k_{on}$ ,  $s$  and  $-\beta\Delta g$  were obtained by a least-squares minimization on the scattering intensities of capsid assembly without modulator. In that case, the form factors of subunits and capsids were calculated by CRY SOL<sup>1</sup> with the crystal structures.  $\chi^2$  was 1.4, which was better than with ‘model A’, because three fitting parameters allowed more flexibility to reproduce the data than only two as in ‘model A’. However, the height of the energy barrier at the beginning of the assembly was  $\beta\Delta G_* = 0.41$ , which invalidates the use of the classical nucleation theory – recall that it requires  $\beta\Delta G_* \gg 1$ . By fixing  $k_{on}$  and  $s$ , and using the values of  $-\beta\Delta g$  found with ‘model A’, we can see that the classical nucleation theory could nicely reproduce the fractions  $\phi_{se}$  for capsid assembly with CAM-E and CAM-A (Figure S13), in particular over the early time steps. The form factors were the same as those used with ‘model A’. In both cases,  $\chi^2$  was less than 2.0, but once again  $\beta\Delta G_*$  at the beginning of the assembly was lower than unity, which renders the classical nucleation theory not appropriate for the modelling.

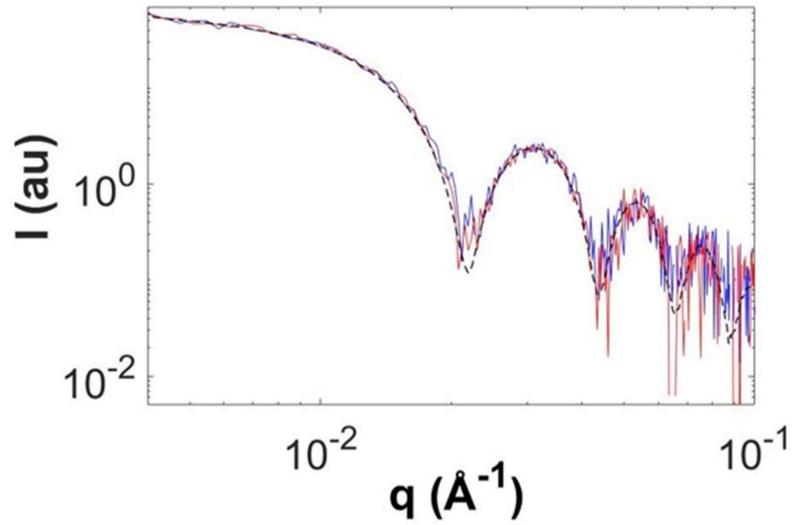


**Figure S1. SAXS patterns of free subunits.** Experimental SAXS pattern (blue) measured with 30  $\mu\text{M}$  of subunits in 1 mM DTT, 50 mM Tris-HCl pH 7.5, and SAXS pattern (orange) calculated with CRY SOL<sup>47</sup> from the crystal structure (bottom left) of a Cp149 dimer (Protein Data Bank code 1QGT).

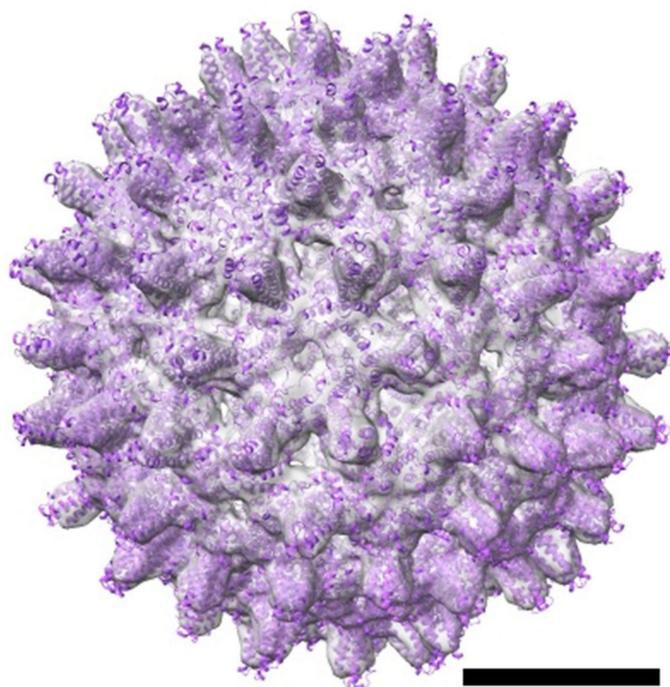


**Figure S2. SAXS patterns of capsids assembled at different subunit concentrations.**

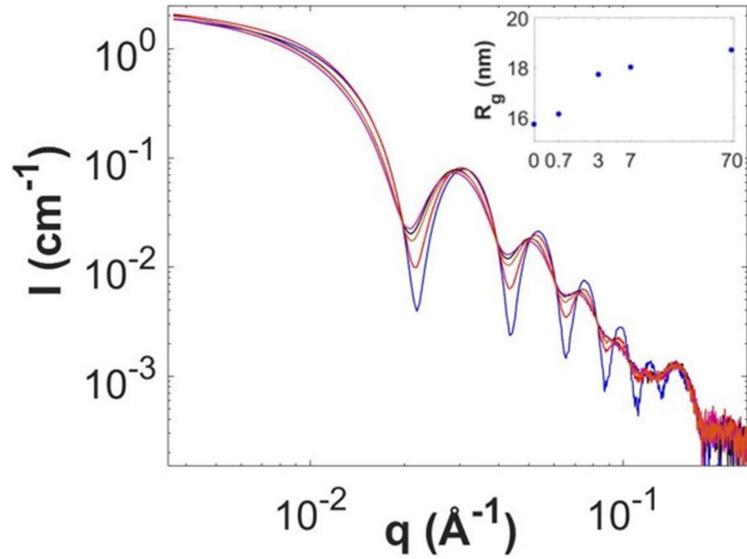
Experimental SAXS patterns of capsids formed with 30  $\mu\text{M}$  (orange line), 22.5  $\mu\text{M}$  (red line), and 15  $\mu\text{M}$  (blue line) of subunits, and SAXS pattern (black dashed line) calculated with CRY SOL<sup>1</sup> from the crystal structure (bottom left) of a Cp149  $T = 4$  capsid (Protein Data Bank code 1QGT). The assembly was triggered with 150 mM of ammonium acetate.



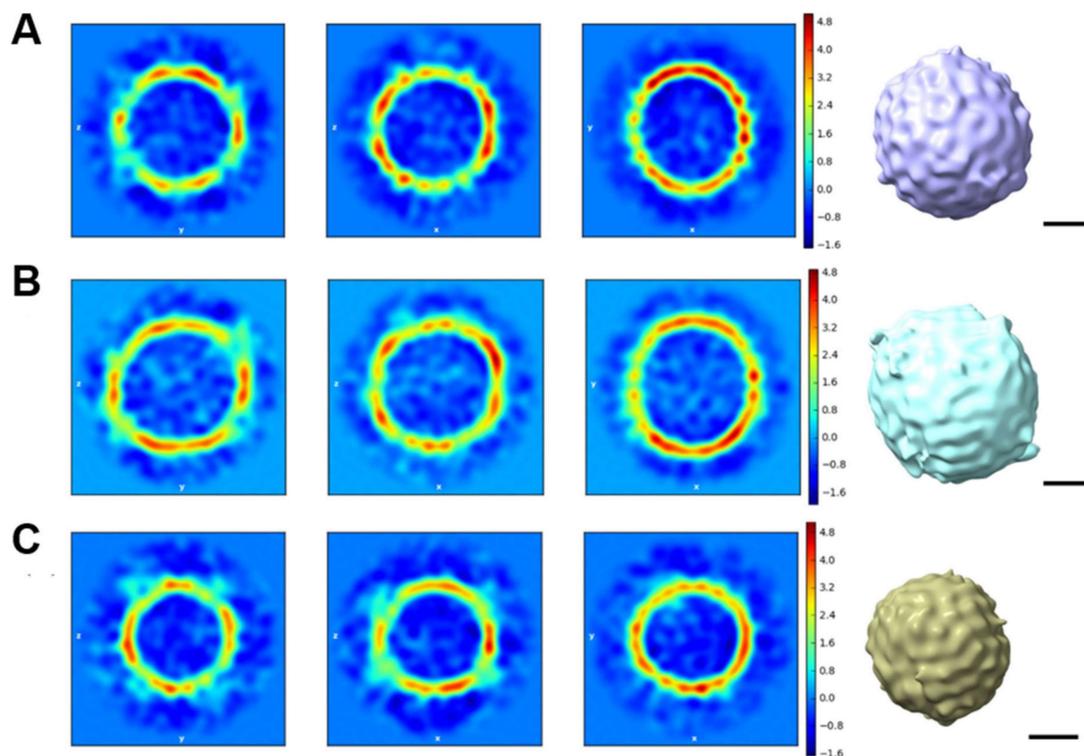
**Figure S3. SAXS and TR-SAXS patterns of assembled capsids.** Experimental TR-SAXS patterns measured at 4 s (blue line) and at 850 s (red line) by using a stopped-flow mixer, and experimental SAXS pattern of capsids at equilibrium (black dashed line) via manual mixing. Subunit concentration was 30  $\mu\text{M}$  in all cases.



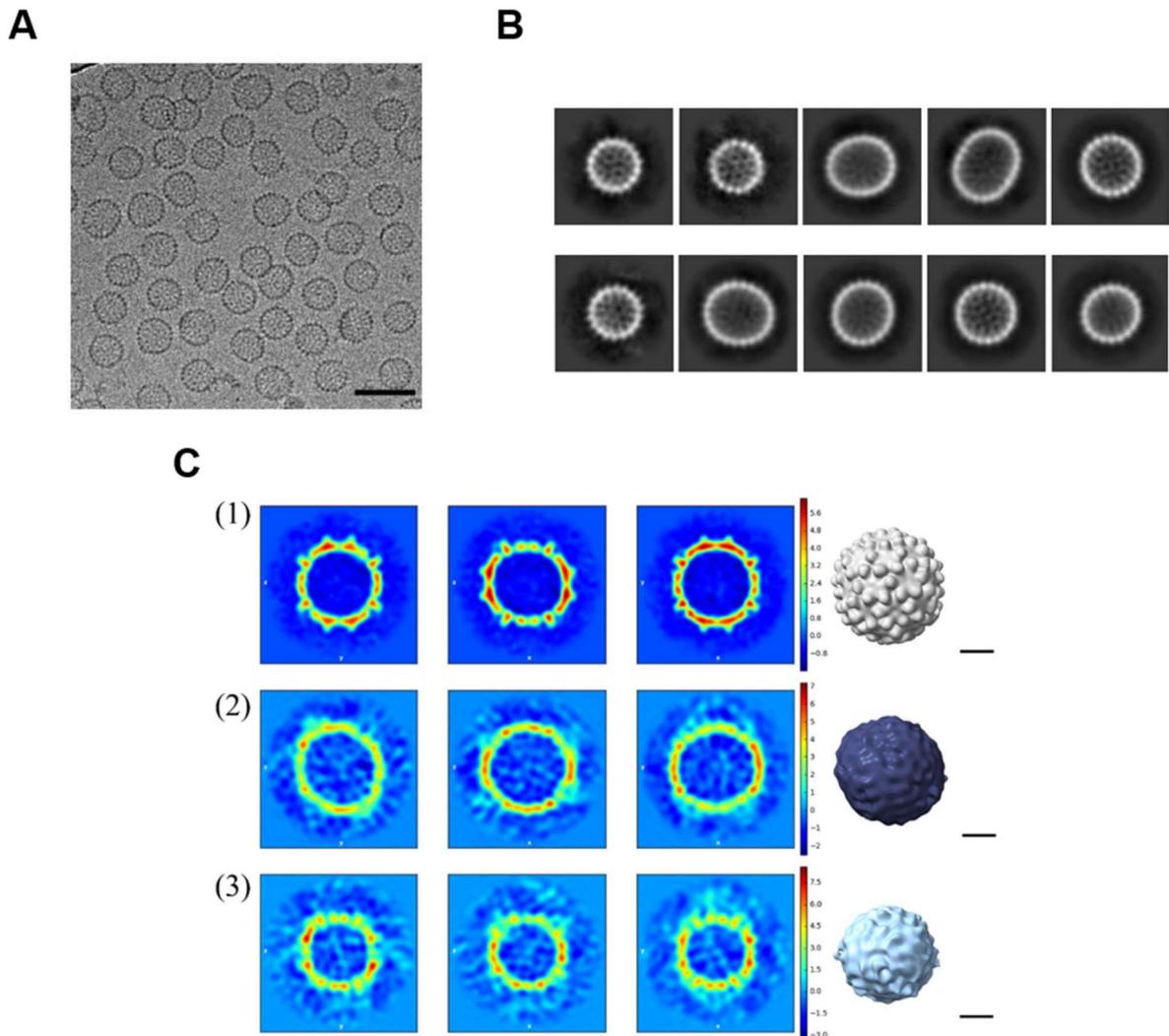
**Figure S4. Fit of the cryoTEM 3D map of assembled capsids with the  $T = 4$  crystal structure (Protein Data Bank 1QGT).** Capsids were assembled with a subunit concentration of 30  $\mu\text{M}$ , then concentrated 10 times. 3D map resolution after reconstruction with icosahedral symmetry is 8.6  $\text{\AA}$  (limited by the microscope used), the scale bar is 10 nm, and the map was obtained with 3,478 particles.



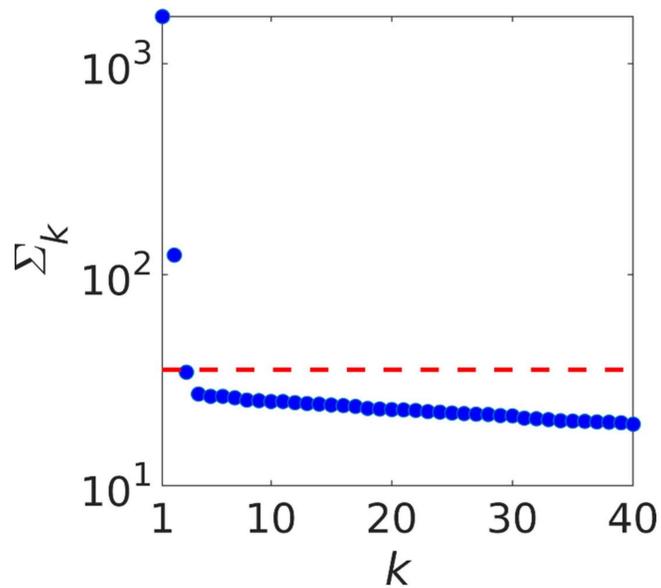
**Figure S5. SAXS patterns of capsids assembled with different molar ratios of CAM-E.** The assembly was triggered with 150 mM of ammonium acetate and modulator-to-subunit molar ratios  $\rho$  of 0 (blue), 0.7 (red), 3.0 (orange), 7.0 (black) and 70 (magenta). The inset shows the radius of gyration  $R_g$  as a function of  $\rho$ .



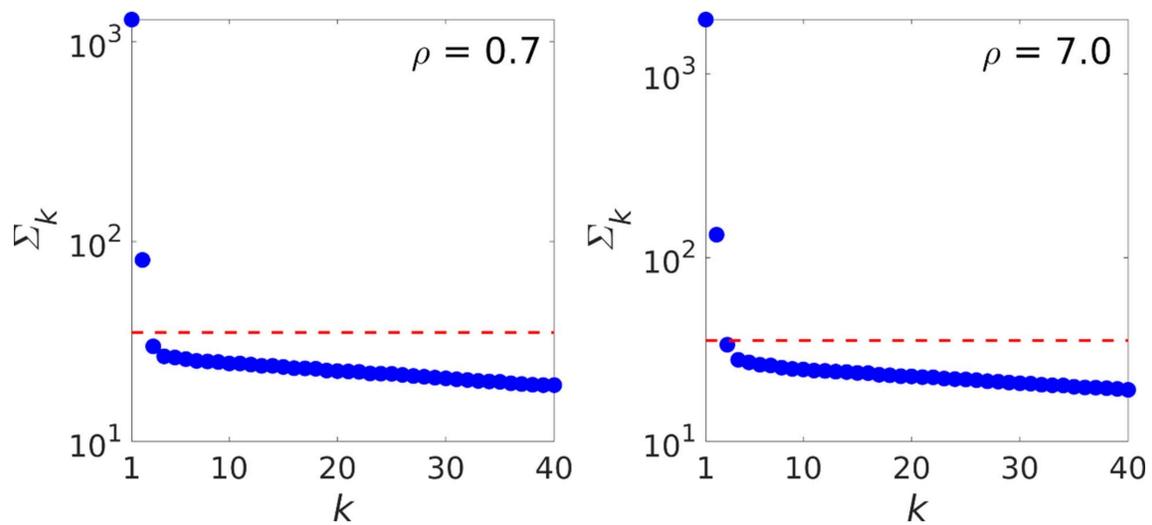
**Figure S6. Real space slices from the 3D cryoSPARC reconstruction and 3D map of capsids assembled in the presence of CAM-E at  $\rho = 7.0$ .** Capsids were assembled with a subunit concentration of  $30 \mu\text{M}$ , then concentrated 11-fold. 3D map resolution after reconstruction without icosahedral symmetry is  $26 \text{ \AA}$  from 1,397 particles (**A**),  $30 \text{ \AA}$  from 892 particles (**B**) and  $30 \text{ \AA}$  from 721 particles (**C**). The box size is  $50 \text{ nm}$  and scale bars are  $10 \text{ nm}$ .



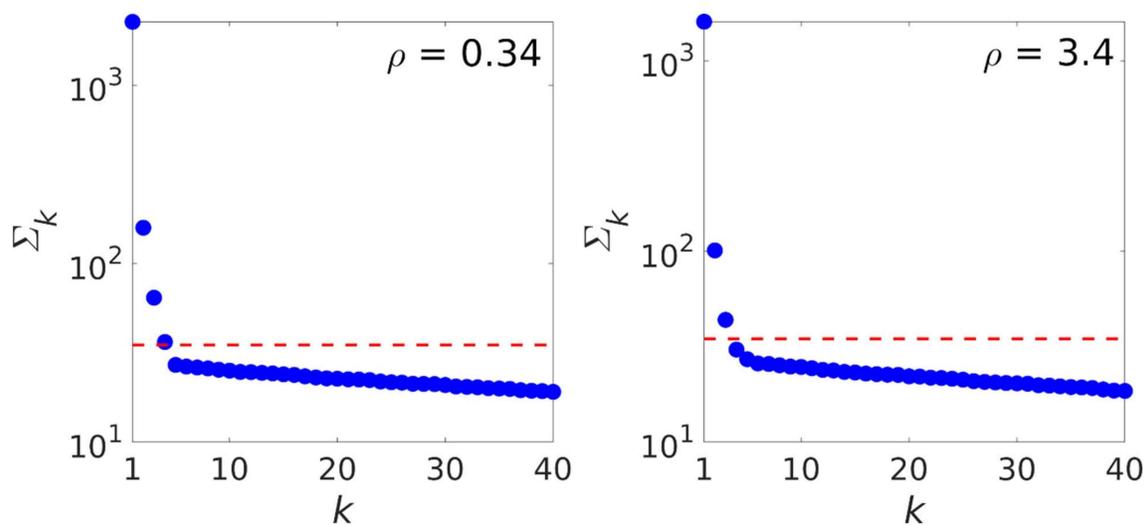
**Figure S7. CryoTEM data for capsid assembly in the presence of CAM-E at  $\rho = 0.7$ .** (A) CryoTEM micrograph of assembled capsids after 11-fold concentration. The initial subunit concentration for assembly was  $30 \mu\text{M}$ . The scale bar is 50 nm. (B) 2D classification of various particles identified by cryoSPARC.<sup>2</sup> The box size is 50 nm. (C) Real space slices from 3D cryoSPARC<sup>49</sup> reconstruction and 3D map of capsids. The 3D map resolution after reconstruction without icosahedral symmetry is 23 Å from 2,953 particles (1), 33 Å from 576 particles (2) and 32 Å from 383 particles (3). Scale bars are 10 nm.



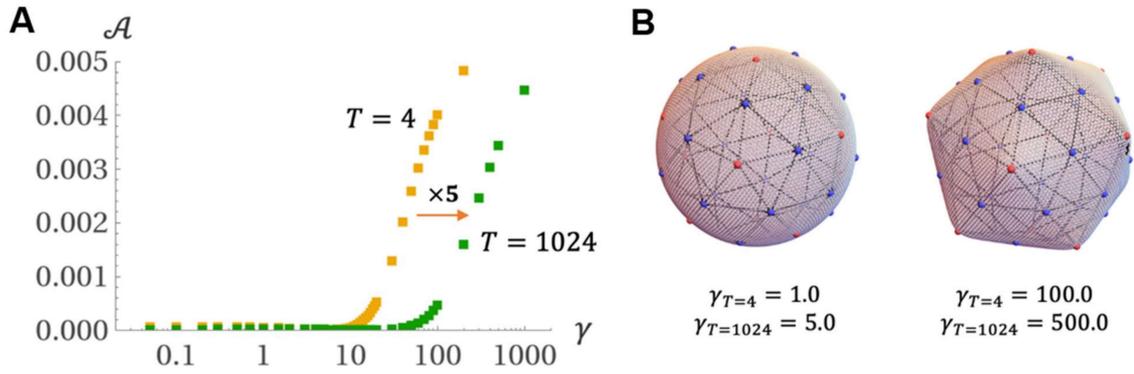
**Figure S8. SVD analysis for capsid assembly without modulator.** Singular values  $\Sigma_k$  (blue discs) as a function of rank  $k$  computed from the normalized scattering intensities measured at 30  $\mu\text{M}$  in subunits without modulator. The red dashed line represents the optimal hard threshold  $\Sigma_*$  described in Supplementary Methods. Only the first 40 singular values are shown for clarity. The optimal rank was then  $r_* = 2$ , which means that two basis vectors were sufficient to approximate the scattering intensities.



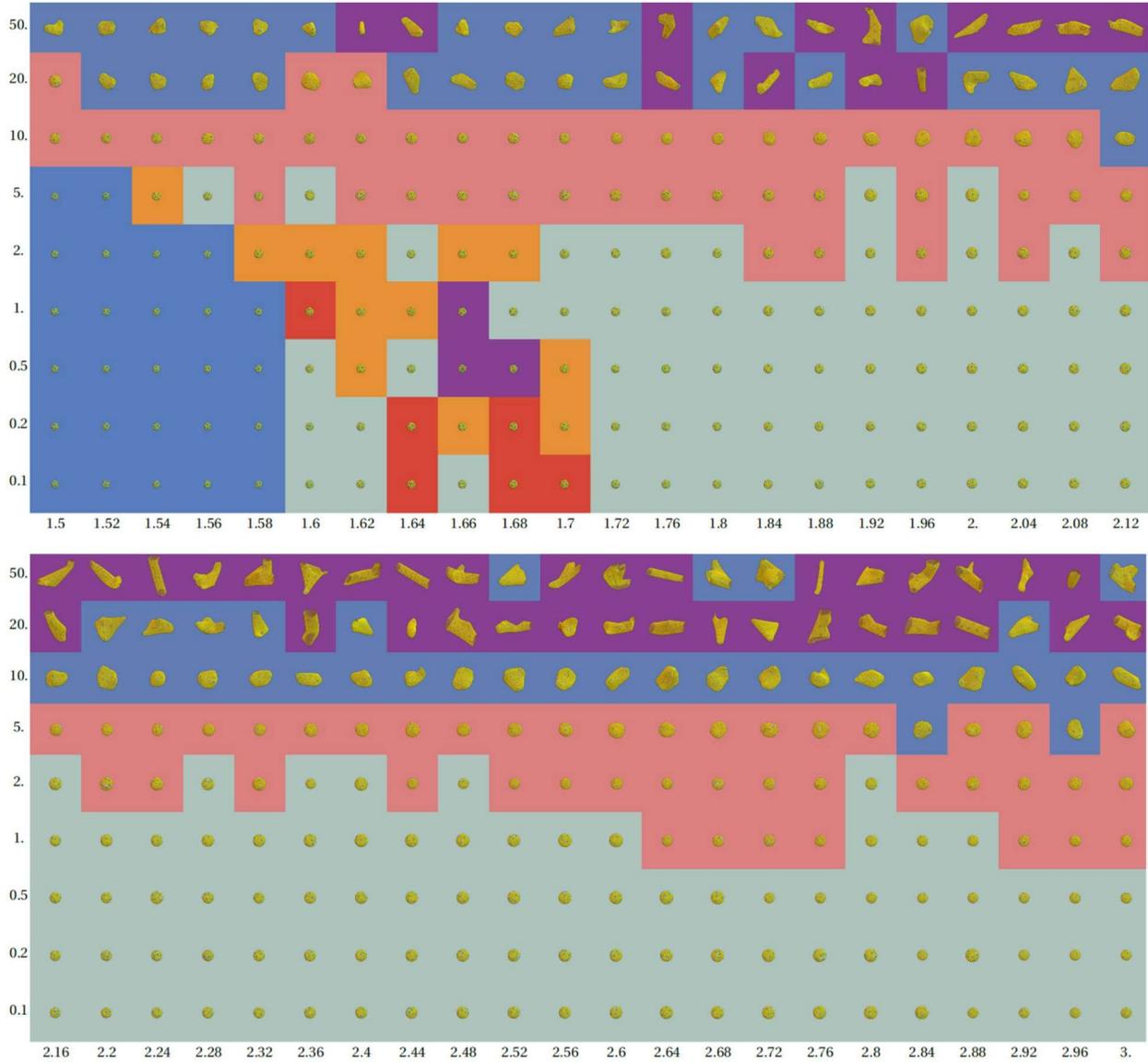
**Figure S9. SVD analysis for capsid assembly with CAM-E.** Singular values  $\Sigma_k$  (blue discs) as a function of rank  $k$  computed from the normalized scattering intensities measured at 30  $\mu\text{M}$  in subunits, with CAM-E at molar ratios  $\rho$  of 0.7 and 7.0. The red dashed line represents the optimal hard threshold  $\Sigma_*$ . Only the first 40 singular values are shown for clarity. The figures show that two basis vectors were sufficient to reproduce the scattering intensities at both ratios.



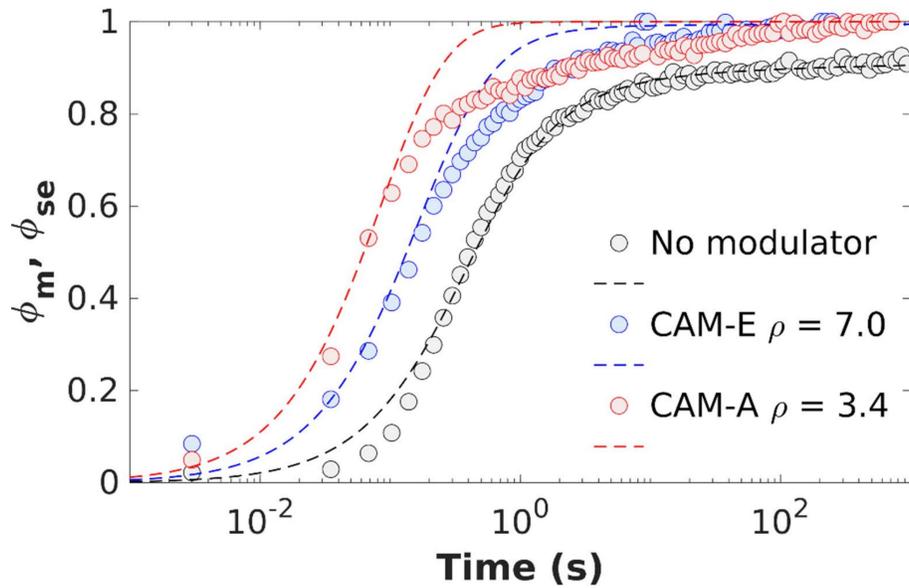
**Figure S10. SVD analysis for capsid assembly with CAM-A.** Singular values  $\Sigma_k$  (blue discs) as a function of rank  $k$  computed from the normalized scattering intensities measured at 30  $\mu\text{M}$  in subunits, with CAM-A at molar ratios  $\rho$  of 0.34 and 3.4. The red dashed line represents the optimal hard threshold  $\Sigma_*$ . Only the first 40 singular values are shown for clarity. The figures show that more than two basis vectors were necessary to reproduce the scattering intensities.



**Figure S11. Mapping between discrete  $\gamma$  and continuum FvK number.** (A) Comparison of the asphericities between two discrete closed shell models with triangulation numbers  $T = 4$  and  $T = 1024$ . The  $T = 1024$  model was obtained by refining the  $T = 4$  mesh, which then could be compared with the results in the continuum limit. Based on the asphericity of the two models, i.e.,  $T = 4$  and  $T = 1024$ , we found similar asphericities for  $\gamma_{T=1024} \sim 5\gamma_{T=4}$ . Considering that FvK number is  $YR^2/\kappa$ , or  $4/3R^2\gamma_{T=1024}$ ,<sup>3</sup> we could estimate  $\gamma_{T=4}$  in our simulations to be  $\sim 0.05\text{FvK}$ . (B) Illustration of the closed shells for  $T = 4$  (black mesh) and  $T = 1024$  (pink mesh). The dots with two-fold (blue) and five-fold (red) symmetry were chosen to calculate the asphericity presented in Eq. 9. Note that the asphericity was calculated after the shells were fully relaxed with the elastic energy given by Eq. 7, where the parameters used were  $l_0 = 1$  for  $T = 4$  and  $l_0 = 0.0625$  for  $T = 1024$ , while  $R_0 = 1.7$  in both cases.



**Figure S12. Full phase diagram computed by coarse-grained Monte Carlo simulation.** The shapes were calculated as a function of the spontaneous radius of curvature  $R_0$  in  $x$  axis and  $\gamma$ , which is proportional to the Föppl-von-Kármán (FvK) number, in  $y$  axis. The  $x$ -axis corresponds in the top graph to  $1.5 \leq R_0 \leq 2.24$ , and in the bottom graph to  $2.26 \leq R_0 \leq 3.0$ . The range of  $\gamma$  values is the same for both graphs, namely,  $0.1 \leq \gamma \leq 50$ . The color code is the same as that of Figure 6 in the main text.



**Figure S13. Kinetic modelling of TR-SAXS data with the classical nucleation theory.** Fraction of subunits in capsids versus time for capsid assembly without modulator (black), as well as with CAM-E (blue) and CAM-A (red), the modulator-to-subunit ratios  $\rho$  being 7.0 and 3.4, respectively. Discs are the fractions  $\phi_{se}$  obtained with the semi-experimental model (see Methods for details) while dashed lines represent the fractions  $\phi_m$  estimated through the classical nucleation theory (see Supplementary Methods). The parameters of the classical nucleation theory were:  $k_{on} = 3 \text{ s}^{-1}$ ,  $s = 0.07$ , and  $-\beta\Delta g$  was set to 9.0, 12 and 18 without modulator, with CAM-E and with CAM-A, respectively. In all cases, the subunit concentration was  $30 \text{ }\mu\text{M}$ .

## Supporting References

55. Henry, E. R. The use of matrix methods in the modeling of spectroscopic data sets. *Biophys. J.* **72**, 652–673 (1997).

56. Gavish, M. & Donoho, D. L. The Optimal Hard Threshold for Singular Values is  $\sqrt{3}$ . *IEEE Trans. Inf. Theory* **60**, 5040–5053 (2014).

**Article : Relationships between RNA topology and nucleocapsid structure in a model icosahedral virus**

# Relationships between RNA topology and nucleocapsid structure in a model icosahedral virus

Laurent Marichal,<sup>1</sup> Laetitia Gargowitsch,<sup>1</sup> Rafael Leite Rubim,<sup>1</sup> Christina Sizun,<sup>2</sup> Kalouna Kra,<sup>1,3</sup> Stéphane Bressanelli,<sup>3</sup> Yinan Dong,<sup>4</sup> Sanaz Panahandeh,<sup>4</sup> Roya Zandi,<sup>4</sup> and Guillaume Tresset<sup>1,\*</sup>

<sup>1</sup>Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides, Orsay, France; <sup>2</sup>Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, Gif-sur-Yvette, France; <sup>3</sup>Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell, Gif-sur-Yvette, France; and <sup>4</sup>Department of Physics and Astronomy, University of California, Riverside, California

**ABSTRACT** The process of genome packaging in most of viruses is poorly understood, notably the role of the genome itself in the nucleocapsid structure. For simple icosahedral single-stranded RNA viruses, the branched topology due to the RNA secondary structure is thought to lower the free energy required to complete a virion. We investigate the structure of nucleocapsids packaging RNA segments with various degrees of compactness by small-angle x-ray scattering and cryotransmission electron microscopy. The structural differences are mild even though compact RNA segments lead on average to better-ordered and more uniform particles across the sample. Numerical calculations confirm that the free energy is lowered for the RNA segments displaying the larger number of branch points. The effect is, however, opposite with synthetic polyelectrolytes, in which a star topology gives rise to more disorder in the capsids than a linear topology. If RNA compactness and size account in part for the proper assembly of the nucleocapsid and the genome selectivity, other factors most likely related to the host cell environment during viral assembly must come into play as well.

**SIGNIFICANCE** Many single-stranded RNA viruses package their genome through a self-assembly process occurring within the host cell. Their survival depends upon the capability of capsid proteins to selectively capture the segments of viral RNA among a sea of host cell molecules and, concomitantly, to build up a defectless capsid. By using a model icosahedral virus, we observed that nucleocapsids packaging compact RNA segments were slightly better assembled than with loosely folded segments, whereas the trend was opposite for capsids packaging synthetic polyelectrolytes with compact and extended topologies. As a consequence, the genome selectivity cannot rely solely on the RNA topology, and environmental factors at the viral assembly site must ensure the production of viable virions.

## INTRODUCTION

Positive-sense, single-stranded (ss)RNA viruses are ubiquitous in the environment, with members infecting all kingdoms of life apart from Archaea (1). Some of them cause important public health issues because of a lack of vaccines (e.g., hepatitis C virus), curative treatments (e.g., coronaviruses), or most often both (e.g., dengue virus). Part of the difficulty of fighting the associated diseases comes from their life cycle, which alternates between well-protected virions and unprotected but intracellular, biologically active viral RNAs and proteins that hijack the biological systems of the host cell. Every ssRNA virus possesses a genome,

either in the form of one RNA molecule or split into several molecules in the case of segmented genomes. The genome is protected by a protein shell called capsid, the composition and structure of which vary but is generally remarkably simple. In the simplest cases, the building blocks of these capsids, called capsid proteins (CPs), have identical sequences and can self-assemble to form objects with spherical or diverse geometries (2–4). One of the most remarkable shared characteristics of these viruses is their ability to spontaneously self-assemble within the host cell. Contrary to double-stranded DNA and RNA viruses, which generally package their genome inside preformed capsids using non-spontaneous, ATP-consuming mechanisms, ssRNA viruses package their genome during the capsid assembly (5,6). This co-assembly process has been the center of many studies, and a general understanding of the process has been achieved (7–10). Being spontaneous, the nucleocapsid

Submitted March 24, 2021, and accepted for publication August 13, 2021.

\*Correspondence: [guillaume.tresset@universite-paris-saclay.fr](mailto:guillaume.tresset@universite-paris-saclay.fr)

Editor: Wilma Olson.

<https://doi.org/10.1016/j.bpj.2021.08.021>

© 2021 Biophysical Society.



self-assembly relies only on RNA-protein and protein-protein interactions. Yet, the presence of RNAs is not mandatory for the capsid formation because many *in vitro* studies have shown the formation of empty capsids (11–13). These studies, as well as simulation studies, determined the local forces ruling protein-protein interactions, especially the existence of hydrophobic interaction between subunits (14–17). However, the biological relevance of these experiments is somewhat questionable because RNAs (either viral or nonviral) are always present in the infected cell, and the physicochemical conditions, in particular pH and ionic strength, used to trigger the self-assembly of empty capsids are often nonphysiological. Moreover, RNA can lower the energy barrier to assembled capsids and thus favors the formation of virions (18). The resulting nucleocapsids generally possess a highly regular structure with very few defects. The main driving force arises from the electrostatic attractive interactions between the negatively charged nucleotides and the positively charged, RNA-binding domain of the CPs protruding toward the interior of the capsid (19,20).

However, many details of the co-assembly mechanisms are still elusive. In particular, how viruses selectively package their own genome and not the host RNAs is still a matter of debate (6). Genome replication and transcription, as well as nucleocapsid self-assembly, generally occur in the cytoplasm (21). In this crowded environment, viral RNAs are outnumbered by the large number of host RNAs present. A common strategy adopted by ssRNA viruses is to compartmentalize their genome and possibly proteins into organelle-like structures, generally in the form of membrane invaginations or vesicles (21). The spatial proximity is likely to reduce the packaging competition between host and viral RNAs. Nevertheless, significant amounts of nonviral RNAs were found to be packaged during the nucleocapsid assembly (22), indicating that despite compartmentalization, viral RNAs have to compete with nonviral ones.

To package their own genome, ssRNA viruses developed several strategies of packaging selectivity (6). One of the most important strategies is the presence of sequences and/or structures embedded in viral RNAs that can specifically interact with CPs of the cognate virus. These specific elements—coined “packaging signals”—have been shown to be necessary, but not always sufficient, to explain the packaging selectivity observed during an infection. Many RNA sequences have been identified as packaging signals, and sometimes several of them can be found on a given RNA molecule (6,23). As Twarock et al. stated (24), this illustrates “the crucial roles of multiple dispersed, specific interactions between viral genomes and coat proteins in capsid assembly.”

Packaging signals are often recognized by viral proteins from the secondary structures they can form (e.g., hairpin-like structures). However, nonspecific RNA structures can also play a large role for packaging selectivity. For instance,

RNA length can contribute greatly to the packaging selectivity. Several competition studies determined that longer RNAs were preferentially packaged (25,26). However, there seems to be an optimal length—which is usually close to the length of the viral RNAs—because longer RNAs are shared between multiple, nonclosed capsids (27,28). The length factor is relevant *in vivo* because viral genomic RNAs tend to be larger than RNAs of the host cell (29). Besides, the natural tendency of RNA for basepairing leads to the constitution of structures akin to branched polyelectrolytes. The level of branching depends upon RNA sequence, and viruses have evolved to have highly branched and compact RNAs (30,31). Compactness has been shown to increase the packaging efficiency both in *in vitro* experiments (26) and in simulation studies (32,33). Moreover, increased compactness seems to allow longer RNA molecules to be packaged by viral CPs (33–35). However, the relationship between compactness and selectivity is not straightforward because it was shown in some conditions that linear RNAs could be packaged more efficiently than branched RNAs of the same length (28). An explanation could be that a high level of branching can lead to an increased stiffness that may be detrimental for packaging selectivity (36). The connection between RNA structure and packaging selectivity is then far from being thoroughly understood.

Despite the discovery of various strategies of packaging selectivity, their relative importance during an infection is unknown and is likely to be virus specific. Here, we focus our work on the cowpea chlorotic mottle virus (CCMV), a model virus of the *Bromoviridae* family. This multipartite virus has a segmented genome of four RNA molecules split into three nucleocapsids, each of them hosting around 3000 nucleotides (37). The capsids are homomultimers of CPs that follow the Caspar-Klug quasiequivalence principle of icosahedral symmetry (38,39). They are natively under the shape of  $T = 3$  capsids, i.e., comprising 180 CPs (40). CCMV packaging selectivity strategy is not clearly established because no specific packaging signal has been found so far (6), although there is evidence that RNA length is an important factor (26). On the contrary, studies done on the closely related brome mosaic virus (BMV) showed the existence of packaging signals (41). Besides, competition experiments showed that BMV RNAs were packaged preferentially over CCMV RNAs into CCMV capsids (26). This improved selectivity for noncognate RNAs tends to confirm the lack of packaging signals for CCMV RNAs. However, because of the high similarities between BMV and CCMV genomes, shared packaging selectivity strategies may still be present.

In our study, we compare the structure of CCMV nucleocapsids packaging RNA segments belonging to CCMV and BMV, as well as a nonrelated segment derived from rotavirus. All segments are similar in length, but numerical calculations reveal that they display different numbers of branch points. Secondly, we investigate the structure of

capsids packaging synthetic polyelectrolytes with linear and star topologies. By investigating the structures and thermodynamics of nucleocapsid assemblies as a function of the cargo topology in physiological conditions, we advance our understanding of the nonspecific packaging selectivity of icosahedral ssRNA viruses.

## MATERIALS AND METHODS

### CCMV protein purification

CCMV virions were purified from infected cowpea leaves (*Vigna unguiculata*) by following the purification method previously developed (9,42). CCMV CPs were then isolated from their viral RNAs by an ultracentrifugation-based method. Briefly, virions were dialyzed against 500 mM CaCl<sub>2</sub>, 1 mM EDTA, 1 mM dithiothreitol (DTT), 0.5 mM phenylmethylsulfonyl fluoride, and 50 mM Tris-HCl (pH 7.5). The dialyzed solution (~10 mL) was centrifuged at 150,000 × *g* for 18 h using an Optima XPN-80 ultracentrifuge (Beckman Coulter Life Sciences, Indianapolis, IN). 0.5 mL aliquots were collected, and their protein purity was measured by spectrophotometry using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA). Aliquots with  $A_{260}/A_{280} < 0.66$  were deemed pure and were pooled then concentrated using an Amicon Ultra-4 (10 kDa) centrifugal filter unit. CP concentrations were estimated by ultraviolet absorbance at 280 nm. Virions were stored at -80°C and CPs kept at 4°C until use.

### RNA production

RNAs were produced by *in vitro* transcription from recombinant plasmids. Plasmids coding for each RNA were transformed in 10-β competent *E. coli* cells (New England Biolabs, Ipswich, MA), produced and purified by a NucleoBond Xtra plasmid DNA purification kit (Macherey-Nagel, Düren, Germany), and linearized using the appropriate restriction enzyme. The linearized plasmid containing a T7 promoter was then transcribed using a MEGAscript T7 Transcription Kit (Thermo Fisher Scientific). Newly produced RNAs were finally purified using a MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific) and placed in a moderate ionic strength buffer (50 mM NaCl, 1 mM EDTA, and 50 mM Tris-HCl (pH 7.5)). RNA concentration was measured using the RNA mode of a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). RNA length and purity were checked by running an agarose gel electrophoresis (Fig. S1). UltraPure DNase/RNase-Free Distilled Water (Invitrogen, Carlsbad, CA) was used for every experiment involving RNA. Three RNA segments were used in this study: RNA 1 of BMV (B1), RNA 2 of CCMV (C2), and RNA 2 of the bovine rotavirus strain RF (RF2). Nucleoprotein complexes were produced by dialyzing a mixture of CPs initially in high ionic strength buffer (500 mM NaCl, 0.5 mM DTT, 1 mM EDTA, 50 mM Tris-HCl (pH 7.5)) and RNA against a moderate ionic strength buffer. Final concentrations were estimated by spectrophotometric measurements following the method proposed by Porterfield and Zlotnick (43).

### Deuterated poly(styrene sulfonic acid)

Deuterated linear and star polystyrene (Polymer Source, Dorval, Canada) were used. Sulfonation was performed by adapting a method previously developed (44). 1.1 g of phosphorus pentoxide was dissolved in 5 mL of concentrated sulfuric acid. 100 mg of deuterated polystyrene dissolved in 15 mL of cyclohexane was then added before agitation for 2 h at 50°C. The reaction setup was placed in argon to prevent any presence of water. The mixture was cooled down, and 10 g of ice was added. The aqueous phase was collected, placed in a dialysis tube, and thoroughly dialyzed

against pure sterile water for 2 weeks. Pure deuterated poly(styrene sulfonic acid) dPSS was finally freeze-dried before storage. Our previous investigations revealed that the method yields a degree of sulfonation around 95% (9).

## SAXS

Small-angle x-ray scattering (SAXS) measurements were performed at the SWING beamline of the SOLEIL synchrotron facility (Saint-Aubin, France). The wavelength of the monochromatic beam was set to  $\lambda = 1.03$  Å, and two-dimensional (2D) scattering images were recorded using an Eiger X 4M detector (Dectris, Philadelphia, PA) with a sample-to-detector distance of 2 m, which provided scattering wavenumbers  $Q$  ( $Q = 4\pi/\lambda \sin(\theta/2)$ , where  $\theta$  is the scattering angle) ranging from  $3.2 \times 10^{-3}$  to  $0.57$  Å<sup>-1</sup>. The temperature was maintained at 20°C using a thermostated circulating water bath. For each sample, 40 measurements were taken with a beam exposure time of 1 s. The scattering intensities were converted into absolute units after subtracting the contribution of the buffer solution. Experimental uncertainties were calculated after circular averaging using the Foxtrot software package.

The forward scattering intensity ( $I_0$ ) of a mixture of subunits (dimeric CPs) and RNA is given by

$$I_0 = \Delta b_S^2 X_S + \sum_{N=0}^{+\infty} X_N (\Delta b_{\text{RNA}} + N \Delta b_S)^2, \quad (1)$$

with  $\Delta b_{\text{RNA}}$  and  $\Delta b_S$  the excess scattering lengths of RNA and subunits, respectively;  $X_N$  the molar concentration of nucleoprotein complexes comprising  $N$  subunits ( $X_0 \equiv X_{\text{RNA}}$  the molar concentration of bare RNA); and  $X_S$  the molar concentration of free subunits. The mean number of subunits ( $N$ ) cannot be calculated without the knowledge of the standard deviation (SD) of  $N$ . However, an upper limit  $\langle N \rangle_{\text{up}}$  can be inferred from  $I_0$  regardless of the probability distribution of  $N$  (9) and reads

$$\langle N \rangle_{\text{up}} \equiv \sqrt{\Gamma^2 + \frac{I_0 - I_0^*}{\Delta b_S^2 c_{\text{RNA}}}} - \Gamma, \quad (2)$$

with  $\Gamma \equiv \frac{\Delta b_{\text{RNA}}}{\Delta b_S} - \frac{1}{2}$  and  $I_0^* \equiv \Delta b_{\text{RNA}}^2 c_{\text{RNA}} + \Delta b_S^2 c_S$  the forward scattering intensity produced by a mixture of noninteracting subunits and RNA.  $c_{\text{RNA}}$  and  $c_S$  are the total molar concentrations of RNA and subunits, respectively.  $\Delta b_{\text{RNA}}$  and  $\Delta b_S$  were estimated experimentally from solutions of purified virions and subunits by using Eq. 1.

## Cryotransmission electron microscopy

4 μL of solution was deposited onto a holey carbon grid (Quantifoil R2/2; Großlobichau, Germany) ionized by glow discharge. The grid was blotted with a filter paper for 1.5 s and directly plunged into liquid ethane cooled down by liquid nitrogen using an FEI Vitrobot (FEI, Hillsboro, Oregon) operated at room temperature and 100% relative humidity. The grids were stored in liquid nitrogen until use. Frozen samples were transferred into a Gatan 626 cryo-holder (Gatan, Pleasanton, CA) and observed at -180°C via a JEOL 2010F microscope (Tokyo, Japan) equipped with a 200 kV field emission gun. The samples were imaged with a magnification of ×50,000 using a minimal dose system, and the images were collected with a Gatan Ultrascan 4K charge coupled device (CCD) camera at 2.5 μm of nominal defocus.

Image analysis was performed with cryoSPARC (45). After contrast transfer function (CTF) correction, a first round of automated particle picking (cryoSPARC “blob picking”) was performed. This initial picking was manually curated to confirm real particles and remove clear nonprotein artifacts (tubular structures were retained). Particles were extracted in a 256-pixel

box (nominal pixel size: 2.36 Å), and two rounds of 2D classification were performed (see the numbers of particles used in Table 1. For the sake of clarity, only the five most populated classes of particles were displayed.

## ITC

Isothermal titration calorimetry (ITC) measurements were carried out using a MicroCal PEAK-ITC (Malvern Panalytical, Malvern, UK). The reaction cell (250 μL) was loaded with an RNA solution ( $c_{\text{RNA}} = 0.2 \mu\text{M}$ , 50 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl (pH 7.5)) containing B1, C2, or RF2, and the syringe was filled with a CP solution ( $c_{\text{S}} = 100 \mu\text{M}$ , 500 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl (pH 7.5)). The experiments were done in triplicate at 25°C by injecting 34 μL of CP solution (split into 17 injections) into the cell with an equilibration interval of 200 s. To take into account the salt dilution as well as the protein-protein interactions, control experiments without RNA were performed and subtracted to the titration curves before curve fitting. Variation enthalpy ( $\Delta H$ ), stoichiometry of the reaction ( $N$ ), and dissociation constant ( $K_D$ ) were obtained by nonlinear least-squares fitting of the experimental data using the single set of independent binding sites model of the Origin software provided with the instrument.

## SANS

Small-angle neutron scattering (SANS) measurements were performed on the D22 spectrometer at the Institut Laue Langevin (Grenoble, France). The samples were contained in 1-mm-thick quartz cells, and all measurements were done at room temperature. Data were normalized by using the direct beam, and the buffer contribution was subtracted from these corrected data. The scattering wavenumbers  $Q$  covered the range between  $4 \times 10^{-3}$  and  $0.45 \text{ \AA}^{-1}$  by setting two sample-to-detector distances of 11.2 and 2 m, respectively, with a wavelength of 6 Å. Exposure time was  $\sim 1$  h at large sample-to-detector distance and 10 min at short distance.

Porod volumes were estimated by using GNOM and DATPOROD from the ATSAS suite (46). The mean number of subunits  $\langle N \rangle_{\text{Porod}}$  within capsids was then inferred from the Porod volume of a sample of interest and from that of a solution of native virions in 68% D<sub>2</sub>O (to contrast match the RNA genome), given that the capsids of native virions possess 90 subunits.

The scattering data were fitted with the SASfit package (47) by a model of vesicles of core radius  $R$  with a gaussian polydispersity characterized by an SD  $\Delta R$  on the core radius. In the particular case of capsids packaging linear dPSS, a structure factor of sticky hard spheres was implemented to reproduce the increase of intensity at very small  $Q$ -values. The model was supplemented with a Guinier law and a constant background to account for free subunits in solution.

## Numerical calculations

After the experimental observations, we introduce a model to study the spontaneous packaging of both C2 and RF2 by CPs in a  $T = 3$  icosahedral structure. We present the capsid as a hollow sphere with 180 positively charged N-terminal tails (modeled as solid cylinders) extended toward its

interior. As for the genome, we model it as a flexible negatively charged polyelectrolyte, interacting attractively with the positive charges on the N-terminal domains of capsid proteins.

To consider the impact of the secondary structure of RNA, we used field-theoretical techniques developed for branched polymers. Because the strength of RNA basepairing is relatively weak, we consider the annealed case in which the genome branchiness can easily be changed as a result of its interaction with the positive charges on the CPs (48). Using the ground-state dominance approximation (49,50), in which case only the dominant contribution to the genome partition function is considered, the free energy of the genome-capsid complex in a salt solution at the mean-field level can be written as (33,35,51–53)

$$\beta F = \int d^3r \left[ \frac{a^2}{6} |\nabla \Psi(\mathbf{r})|^2 + W[\Psi(\mathbf{r})] - \frac{\beta^2 e^2}{8\pi\lambda_B} \right] \times |\nabla \Phi(\mathbf{r})|^2 - 2\mu \cosh[\beta e \Phi(\mathbf{r})] + \beta \tau \Phi(\mathbf{r}) \Psi^2(\mathbf{r}) + \int d^2r [\beta \rho(\mathbf{r}) \Phi(\mathbf{r})], \quad (3)$$

where  $\beta$  is the inverse of temperature in the units of energy,  $a$  is the Kuhn length of the polymer,  $e$  is the elementary charge,  $\mu$  is the density of monovalent salt ions, and  $\tau$  is the linear charge density of polyelectrolyte. The Bjerrum length  $\lambda_B = e^2/4\pi\epsilon$  is  $\sim 0.7$  nm for water at room temperature. A very detailed derivation of Eq. 3 can be found in (53,54).

The field  $\Psi(\mathbf{r})$  in Eq. 3 represents the monomer density field and  $\Phi(\mathbf{r})$  the electrostatic potential. The quantity  $\rho(\mathbf{r})$  denotes the density of positive charges on the N-terminal tails of capsid proteins. The first term in Eq. 3 corresponds to the entropic cost of deviation from a uniform polymer density. The last two terms of Eq. 3 are associated with the electrostatic interactions between the polymer segments, the N-terminal capsid proteins, and the salt ions at the level of the Poisson-Boltzmann theory (51,55–57). The term  $W[\Psi]$  is associated with the free energy density of an annealed branched polymer with the excluded volume interaction (58–60),

$$W[\Psi] = -\frac{1}{\sqrt{a^3}} \left( f_e \Psi + \frac{a^3}{6} f_b \Psi^3 \right) + \frac{1}{2} \nu \Psi^4, \quad (4)$$

where  $f_e$  and  $f_b$  are the fugacities of the end and branched points of the annealed chain, respectively (54), and  $\nu$  is the effective excluded volume of each monomer. It is worth mentioning that in this model, the stem-loop or hairpin configurations of RNA are considered as end points. The parameters  $C_e(r) = \frac{1}{\sqrt{a^3}} f_e \Psi$  are associated with the density of end points and  $C_b(r) = \frac{\sqrt{a^3}}{6} f_b \Psi^3$  with the density of branch points.

Using the fugacities  $f_e$  and  $f_b$ , the expectation number of end and branched points,  $N_e$  and  $N_b$ , can be written as

$$N_e = -\beta f_e \frac{\partial F}{\partial f_e} \quad \text{and} \quad N_b = -\beta f_b \frac{\partial F}{\partial f_b} \quad (5)$$

It is important to note that the number of end and branched points are related as follows,

$$N_e = N_b + 2, \quad (6)$$

because there is only one chain in each capsid and no closed loops within the secondary structure of the genome. The chain is linear if  $f_b = 0$ , and the number of branched points increases as  $f_b$  increases. Using Eqs. 5 and 6, we vary  $f_b$  and find  $f_e$ . To this end,  $f_e$  is not a free parameter.

**TABLE 1** Number of particles used in the cryoSPARC analysis

Sample	Number of micrographs	Number of manual picks	Number of particles after 2D cleanup
CP-B1	34	2890	1911
CP-C2	14	1226	1076
CP-RF2	44	3680	3472

By minimizing the free energy with respect to the fields  $\Psi(\mathbf{r})$  and  $\Phi(\mathbf{r})$  subject to the constrain that the total number of monomers inside the capsid is constant, we obtain three self-consistent nonlinear coupled equations for the interior and exterior of the capsid,

$$\frac{a^2}{6}\nabla^2\Psi(\mathbf{r}) = -E\Psi(\mathbf{r}) + \tau\beta\Phi_{in}(\mathbf{r})\Psi(\mathbf{r}) + \frac{1}{2}\frac{\partial W}{\partial\Psi}, \quad (7a)$$

$$\nabla^2\Phi_{in}(\mathbf{r}) = \frac{1}{\lambda_D^2}\sinh[\Phi_{in}(\mathbf{r})] - \frac{\tau}{2\lambda_D^2\mu\beta e^2}\Psi^2(\mathbf{r}) - \frac{1}{2\lambda_D^2\mu\beta e^2}\rho(\mathbf{r}), \quad (7b)$$

and

$$\nabla^2\Phi_{out}(\mathbf{r}) = \frac{1}{\lambda_D^2}\sinh[\Phi_{out}(\mathbf{r})], \quad (7c)$$

where  $\lambda_D = 1/\sqrt{8\pi\lambda_B\mu}$  is the (dimensionless) Debye screening length and  $E$  is Lagrange multiplier enforcing the fixed monomer number inside the shell. The polymer concentration in the exterior of the capsid is zero,  $\Psi = 0$ . The details of the derivations of all equations given above can be found in (35,54).

Solving the three coupled above equations, we obtain the genome density profiles for C2 and RF2 that minimize the free energy inside the capsid. For the numerical calculation, we set the radius of the  $T = 3$  capsid to  $R = 12$  nm and the number of negative charges on the capsid to 1800. Similar to the experimental condition, the salt concentration is 100 mM.

## RESULTS

### RNA compactness

Three RNA segments were used for this study: RNA1 of BMV (denoted B1), RNA2 of CCMV (C2), and RNA2 of the bovine rotavirus strain RF (RF2) (61). These RNAs were chosen for their similar length (3234, 2767, and 2687 nucleotides, respectively) and RF2 also because it is not packaged as such by the rotavirus capsid, in contrast to B1 by the BMV CP and C2 by the CCMV CP. Thus, the secondary structure of RF2 has no reason for being optimized by evolutionary processes for packaging purpose. Moreover, because each of them codes for a different protein, their sequence identities are very low, with very few regions of sequence similarity. Structural characterization of these RNA segments was performed by SAXS. This ensemble-averaging technique gives information about the size, shape, and dispersity of suspended objects. RNAs are rather disordered objects, and their form factors differ from those of well-organized structures. To obtain information about RNA conformations, the SAXS data were converted into a dimensionless Kratky plot (Fig. 1). This representation emphasizes the differences between the pattern of compact objects and that of random chains (62). The three RNA segments had neatly distinct patterns. C2 RNA displayed a bell-shaped curve in the low to intermediate  $QR_g$  that was related to the over-

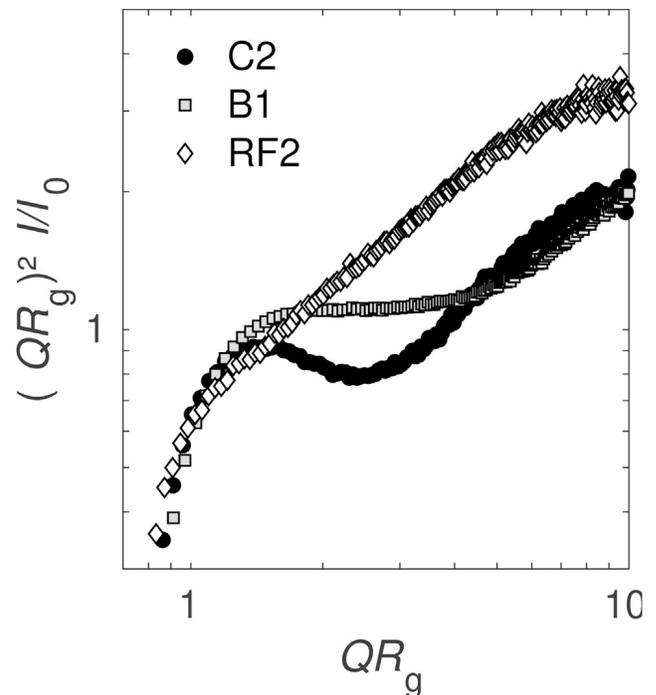


FIGURE 1 SAXS curves of RNA segments. RNA samples were dialyzed in a moderate ionic strength buffer (50 mM NaCl, 1 mM EDTA, 50 mM Tris-HCl (pH 7.5)), and their final concentration was  $c_{\text{RNA}} = 0.5 \text{ g} \cdot \text{L}^{-1}$ . The dimensionless Kratky plot representation was chosen to compare the degrees of compactness.  $R_g$  is the radius of gyration,  $I$  the scattering intensity, and  $I_0$  the value of  $I$  extrapolated at  $Q = 0$ .  $R_g$  of B1 was estimated to be 12 nm, whereas that of C2 and RF2 was  $\sim 10$  nm.

all compactness of the molecule (63). B1 RNA had a pattern similar to C2 but showed a plateau instead of a decrease of intensity in the intermediate  $QR_g$  region ( $R_g$  denotes the radius of gyration). A bell-shaped curve, as seen with C2, is attributed to a partially folded conformation, whereas the presence of a plateau, as seen with B1, is characteristic of a less compact structure (63). The Kratky plot of RF2 displayed a continuous increase of intensity, which is reminiscent of a rigid, elongated object. Therefore, RF2 is likely hybridized over longer segments on average than C2 and B1.

### Ensemble-averaged structures of nucleocapsids packaging various RNA segments

Each CCMV CP is a 190-amino-acid-long polypeptide (20.3 kDa) (40). In the absence of RNA, CCMV CPs can form a wide variety of polymorphs (e.g., dimers, hollow capsids, tubes), depending on the pH and ionic strength of the solution (12). The building block of all these polymorphs is the dimer of CP, which will be hereafter referred to as the subunit. Recently, a strong attention was directed toward the formation of nucleocapsids (8,9,64). A reason for that is the fact that the physicochemical conditions used to form nucleocapsids in vitro are not always

biologically relevant (e.g., acidic pH). In this study, we chose to be close to physiological conditions (pH 7.5, ionic strength of 100 mM), and we probed the structural differences of nucleocapsids packaging different RNA segments.

First, we studied the structure of particles formed by mixing CCMV CPs and RNA segments at different concentrations using SAXS. A scattering curve of native virions is given for reference (Fig. 2 B, blue curve) and showed pronounced oscillations, characteristic of monodisperse sphere-like nucleocapsids. In comparison, the SAXS curves of CP-C2 mixtures (Fig. 2 B) showed oscillations that became more pronounced when the CP concentration increased but were always less marked than the reference curve. Differences in form factor could be quantified by fitting the experimental curves with a polydisperse vesicle model and by estimating polydispersity values (Table S1). From the lowest to the highest concentration, there was a clear decrease in polydispersity ( $\Delta R/R$  going from 32 to 26%,  $R$  being the inner radius of the vesicle), but even the most concentrated sample was still more polydisperse than native virions measured in the same physicochemical condition ( $\Delta R/R = 20\%$ ). The decrease in polydispersity means that the objects became more ordered upon increased concentration. The disorder-order transition of CCMV nucleocapsids was established by Monte Carlo simulations (65), which highlighted the crucial role of elastic energy during assembly. CP-B1 curves (Fig. 2 A) show form factors that are highly similar to CP-C2 curves, indicating the presence of sphere-like objects. However, the increase of concentration only slightly reduced polydispersity ( $\Delta R/R$  going from 28 to 26% in the concentration range studied). This is due to the fact that at low concentration, oscillations were already highly pronounced, indicating the presence of well-formed nucleocapsids. Thus, under identical physicochemical condition and similar concentration, the non-

cognate B1 RNA appeared to be better at forming nucleocapsids. However, it is worth mentioning that at even lower concentration and higher ionic strength, CP-B1 mixtures formed amorphous particles instead of closed nucleocapsids (66). Finally, CP-RF2 curves showed totally distinct scattering patterns (Fig. 2 C). In the low  $Q$  region ( $Q \leq 0.01 \text{ \AA}^{-1}$ ), the intensity did not reach a plateau, indicating the existence of long-range intermolecular interactions (structure factor). In the intermediate  $Q$  region ( $0.01 \leq Q \leq 0.1 \text{ \AA}^{-1}$ ), oscillations were absent, which indicated that the objects were nonhomogeneous, partially assembled nucleocapsids. For these reasons, we did not apply any model to fit CP-RF2 curves.

The presence of a plateau in the scattering curves of B1 and C2 mixtures gave us access to the forward scattering intensities  $I_0$  that are directly related to the mass of the particles. Using an analytical method previously developed (9), an approximation of the number of bound subunits per RNA  $\langle N \rangle_{\text{up}}$  could be deduced (Fig. 3). Native virions placed under the same physicochemical condition were used as a reference for  $\langle N \rangle_{\text{up}}$  determination. Their scattering curve was typical of swollen nucleocapsids (Fig. 2 B), and a  $\langle N \rangle_{\text{up}}$ -value of 90 subunits was imposed, as expected for  $T = 3$  icosahedral capsids. By comparison, C2 and B1 mixtures had on average fewer bound subunits per RNA segment (Fig. 3, gray and black disks, respectively), indicating that under this condition, most of the objects were partially assembled nucleocapsids. In both cases, there was a clear increase of  $\langle N \rangle_{\text{up}}$  upon increase of both CP and RNA concentrations, which confirmed the existence of nonspecific interactions. However, at similar concentrations,  $\langle N \rangle_{\text{up}}$ -values of CP-B1 mixtures were always higher than those of CP-C2 mixtures. This could be explained by the slightly larger length of B1 compared to C2, which resulted in more binding sites.

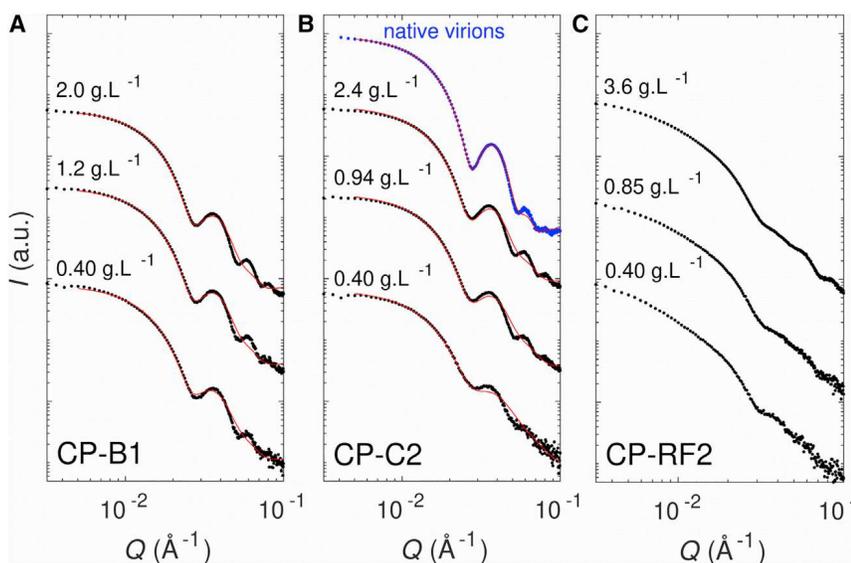


FIGURE 2 SAXS curves of nucleocapsids. CCMV CPs were mixed with RNAs and dialyzed in a moderate ionic strength buffer (50 mM NaCl, 50 mM Tris-HCl (pH 7.5)): B1 (A), C2 (B), and RF2 (C). For each system, three concentrations were measured with the CP/RNA mass ratio kept constant. Exact CP and RNA concentrations can be found in Table S1. Curves were fitted by a polydisperse vesicle model (solid red lines). Native virions dialyzed in the same buffer were used as a reference (blue curve). To see this figure in color, go online.

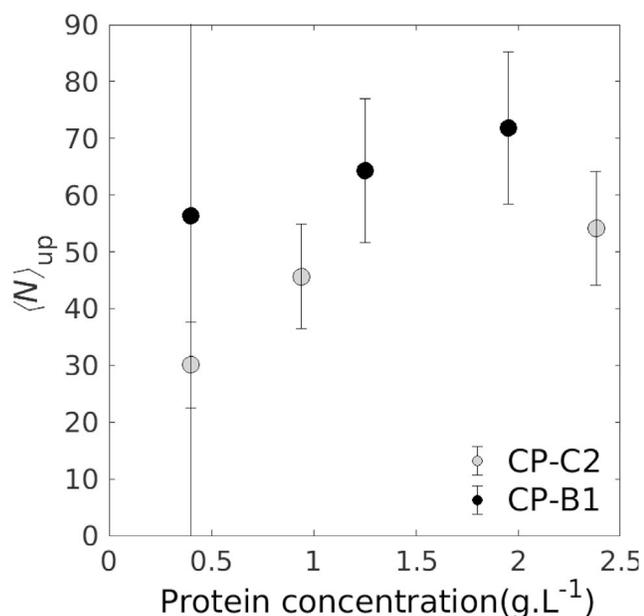


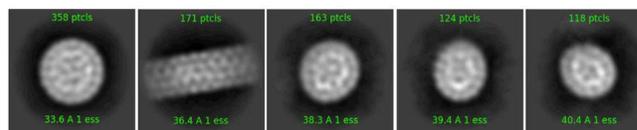
FIGURE 3 Mean number of bound subunits per RNA segment  $\langle N \rangle_{\text{up}}$  calculated from the scattering curves in Fig. 2. CP-C2 mixtures (gray disks) and CP-B1 mixtures (black disks) of varying protein concentrations are indicated. The protein/RNA mass ratio was around 3.5 and 4.0 for CP-C2 and CP-B1, respectively. Error bars are standard errors of the mean.

### Morphology at the single-particle level

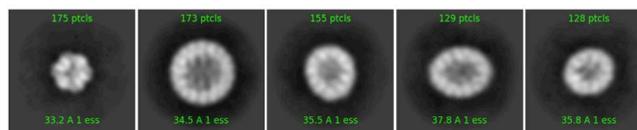
We investigated the morphology of the nucleocapsids made with each of the three RNA segments by cryotransmission electron microscopy (cryoTEM). To do so, we carried out 2D classification over more than a thousand particles for each kind of nucleocapsid (Fig. 4). It turned out that in all cases, some sphere-like particles were formed, and their size was consistent with  $T = 3$  icosahedral nucleocapsids (see Table S2). The rest of the identified particles looked like the pieces of nucleocapsids seen under various orientations. In the case of CP-B1 and CP-RF2 mixtures, nanotubes comprising some defects were identified as well. It must be noted here that, unlike SAXS measurements, this 2D classification did not reflect the actual composition of the samples. Only those particles that were structurally similar enough were grouped into homogeneous classes by the cryoSPARC software, whereas amorphous and loosely ordered particles were assigned to junk classes and not retained for further analysis (see Fig. S2 for typical cryoTEM views of the samples). Nonetheless, Fig. 4 indicates that even though RF2 induced a lower degree of order and/or more variability than B1 and C2 across a large set of particles as evidenced by SAXS data (Fig. 2) at the single-particle level, a significant number of spherical nucleocapsids were completed or partially assembled.

The presence of nanotubes coexisting with the CP-B1 and CP-RF2 particles, but not with the CP-C2 ones, was unexpected. Because in this cryoTEM study, we used a CP/RNA mass ratio around 6 and the stoichiometry in a native

### CP-B1



### CP-C2



### CP-RF2

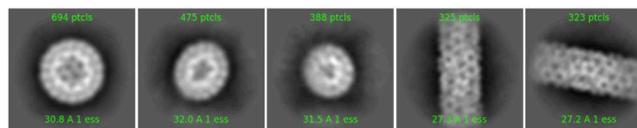


FIGURE 4 CryoTEM 2D classification for nucleocapsids with B1, C2, and RF2. CP concentrations were 4, 4, and 3.6 g · L<sup>-1</sup>, and CP/RNA mass ratios were 6, 6, and 5.5, respectively. The size of each box is 604 Å. To see this figure in color, go online.

virion is of 3.6, CPs were in excess in the solution. CPs alone are known to form nanotubes under these pH and ionic strength conditions (12), so we could indeed expect the CPs in excess to make nanotubes. The fact that they were not seen in the presence of C2 might simply be an artifact caused by the preparation of the microscopy grid; it happens from time to time that large anisotropic objects are sorted out and gathered in a particular location on the grid during the sample blotting and freezing.

### RNA binding affinity

To assess differently the binding affinity of CPs for the various RNA segments, we carried out ITC experiments (Fig. 5). This technique measured the heat exchange of an RNA solution during the addition of a CP solution. The large positive peaks seen after each injection (upper panels) were mainly due to the dilution of the salts present in the CP solution. Once the control subtracted, peaks became negative (exothermic) and were directly linked to interactions between CPs and RNAs. Fitting the titration curves gave us information about the stoichiometry  $N$  (number of subunits per RNA molecule), the dissociation constant  $K_D$  (inversely correlated to the affinity) and the variation of enthalpy  $\Delta H$ .

First, the three systems studied had a negative  $\Delta H$ , confirming the electrostatic nature of the interactions at play (67). Secondly, the  $N$ -values obtained by ITC could in principle be compared to  $\langle N \rangle_{\text{up}}$  as measured by SAXS. The  $N$ -values were sensibly lower than  $\langle N \rangle_{\text{up}}$  for each CP-RNA mixture.

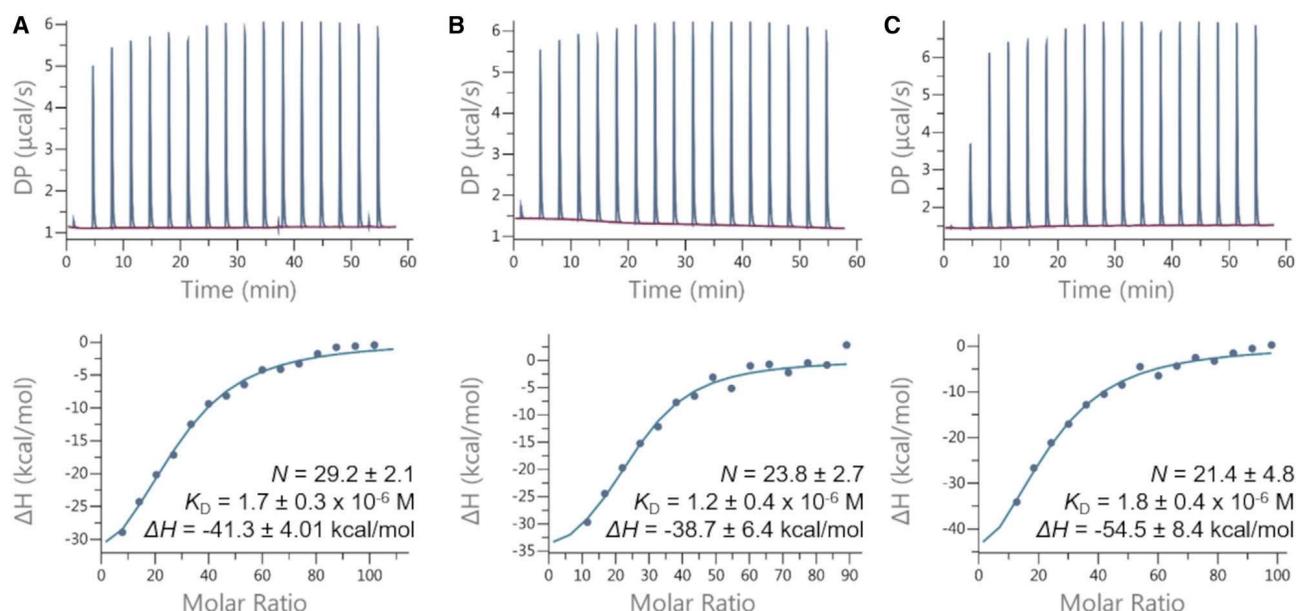


FIGURE 5 ITC experiments of CCMV CPs mixed with RNA segments: B1 (A), C2 (B), and RF2 (C). Upper panels: raw data of the heat exchange. Each spike corresponds to one injection of a CP solution (100  $\mu\text{M}$ ) in an RNA solution (0.2  $\mu\text{M}$ ). Lower panels: integrated heat of each injection after subtraction of the dilution signal and titration curve fitted according to a model of independent binding sites. Stoichiometry ( $N$ ), dissociation constant ( $K_D$ ), and variation of enthalpy ( $\Delta H$ ) are shown. Each value was the average of three replicas. To see this figure in color, go online. Uncertainties are standard errors of the mean.

For instance, B1 and C2  $N$ -values were 29.2 and 23.8, whereas the lowest  $\langle N \rangle_{\text{up}}$  were 56.3 and 30.1, respectively. Because ITC and SAXS experiments were done using commensurate CP and RNA concentrations (with different CP/RNA molar ratios), concentration effects could not be the cause of this discrepancy. However, an explanation can come from the fact that ITC cannot probe athermic reactions or reactions whose heat variations cancel out. In our case, conformational changes of CPs, which are expected to happen during nucleocapsid assembly, may interfere with the accurate titration of CP-RNA interactions. This can lead to underestimations of the actual stoichiometry of the reaction. Such a bias was observed for proteins interacting on hydroxyapatite (68) and silica surfaces (69). However, despite this discrepancy, a relative comparison could be done and similar trends between  $N$  and  $\langle N \rangle_{\text{up}}$  could be observed. Both values were neatly lower than the stoichiometry of a complete nucleocapsid, i.e., 90, which indicates that a large fraction of the particles were partially assembled. B1 interacted with more subunits than C2 and RF2. Again, this can partially be explained by the slightly longer length of the B1 RNA compared to the others. Furthermore, although we could not calculate  $\langle N \rangle_{\text{up}}$ -values for CP-RF2 mixtures using SAXS data,  $N$ -values could be obtained and indicated that C2 and RF2 interacted with the same number of subunits.

Finally,  $K_D$  measured for each system were highly similar from one another, around 1.5  $\mu\text{M}$ . This is in agreement with  $K_D$ -values for CCMV CPs interacting with small RNAs measured by fluorescent correlation spectroscopy ( $K_D = 2\text{--}3 \mu\text{M}$ ) (64). This is also consistent with NMR measure-

ments (Fig. S3), which revealed an intermediate to fast exchange rate (from 2  $\text{s}^{-1}$  up to 1000  $\text{s}^{-1}$  and  $>1000 \text{s}^{-1}$ , respectively) between the bound and free states of CPs. Then, despite forming different assemblies, the three RNAs did not exhibit any difference in terms of binding affinity with CCMV CPs. As a consequence, the topological differences between the studied RNA segments do not seem to influence the binding affinity.

### Calculated density profiles of branch points

The results of numerical calculations are given in Fig. 6 as a plot of the density profile of branch points and nucleotide density (*inset graph*) as a function of  $r$ , the distance from the center of the capsid; see the section Numerical calculations for the details. The black dashed line is the density profile of branch points for C2, and the red solid line is for RF2. The number of branch points was obtained through RNASubopt, a program in the Vienna RNA package (70). Generating an ensemble of secondary structures for sequences of C2 and RF2, we calculated the thermally averaged number of branch points from the secondary structures of each RNA. We found that whereas the number of monomers for C2 and RF2 were close, i.e., 2767 and 2687, respectively, the difference between the number of branch points was more significant: 60 for C2 and 48 for RF2. The number of branch points for B1 was calculated previously (53) and was estimated to be 65.

Fig. 6 clearly shows that the branch point density for C2 was larger than that for RF2, whereas their nucleotide densities

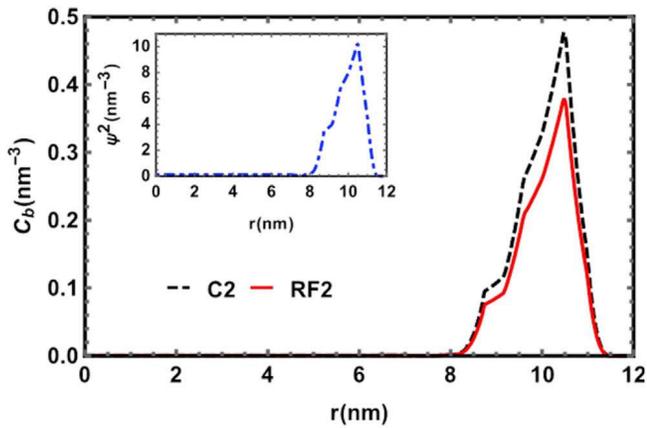


FIGURE 6 Branch density profile versus  $r$ , the distance from the center of the capsid. The dashed black line and red solid line correspond to the branch density profiles for C2 and RF2, respectively. The dot-dashed blue line (*inset graph*) is the nucleotide density profile for C2 and RF2, which is almost the same for both. The capsid radius is  $R = 12$  nm, N-terminal tail length is 4 nm, and the salt concentration 100 mM. The total number of positive charges on the capsid interior is 1800. To see this figure in color, go online.

(*inset*) were almost the same. Using the profiles for RF2 and C2, we calculated the encapsulation free energy and found that it was lower for C2 ( $-3367 k_B T$ , with  $k_B$  the Boltzmann constant and  $T$  the temperature) than for RF2 ( $-3358 k_B T$ ). Because C2 and RF2 had almost the same nucleotide density profile inside the capsid, the difference of their encapsulated free energy was mainly due to the difference in their number of branch points. This slight difference in free energy explains why, experimentally, the nucleocapsids packaging C2 were statistically more uniform and better ordered across the sample than those packaging RF2.

### Packaging linear versus star polyelectrolyte

To get further insight into the physical role of the branched topology in the packaging mechanism of RNA, we assembled virus-like particles containing linear and star synthetic polyelectrolytes. PSS is a negatively charged, flexible homopolymer with a linear charge density close to that of ssRNA ( $\sim -0.33 e \cdot \text{\AA}^{-1}$ ). It is an excellent model to test the selective capabilities of CPs without the specific molecular recognition arising from RNA sequence. Our linear PSS had a molecular mass of 162 kDa (758 repeat units), and star PSS had a molecular mass of 160 kDa (749 repeat units) and comprised 17 arms. Thus, both PSSs possessed a similar number of charges—one negative charge per repeat unit—but a different topology. Virus-like particles were formed with various concentrations of CCMV CPs, but the CP/PSS mass ratio was kept around 6 in all cases. The structure of the capsids was investigated by using SANS with the contrast variation method. Data were collected on the D22 spectrometer at the Institut Laue Lan-

gevin, a neutron source delivering one of the highest fluxes worldwide. Virus-like particles packaging dPSS were dialyzed against a buffer solution containing 100%  $D_2O$  in such a way that the contrast of dPSS with solvent vanished and solely the scattering signal arising from the capsids was detected. This methodology enabled us in the past to measure accurately the mass of capsids (71). In principle, the scattering intensity arising from RNA could be contrast matched with 68%  $D_2O$ . However, the contrast between CP and the solvent containing 68%  $D_2O$  was so weak that we could not probe the capsid structure by SANS in the presence of RNA.

Fig. 7 depicts SANS patterns of virus-like particles packaging linear and star dPSS. As mentioned previously, in 100%  $D_2O$ , the scattering curves reflect the morphology of the capsids only. The CP concentrations for each dPSS topology were sufficiently close to allow a direct comparison of the capsid morphologies. Oscillations of the SANS patterns were slightly more pronounced with the linear dPSS (Fig. 7 A) than with the star one (Fig. 7 B). Such oscillations are characteristic of spherically symmetric objects, and the fact that their amplitudes are damped indicated a morphological variability across the population and/or the presence of defects within the capsids.

Table 2 gives the mean number of subunits per capsid inferred from the Porod volume. The values were all around 60, which would suggest the formation of  $T = 2$  icosahedral capsids, comprising, however, a few defects. Notice that the capsids packaging star dPSS contained slightly more subunits

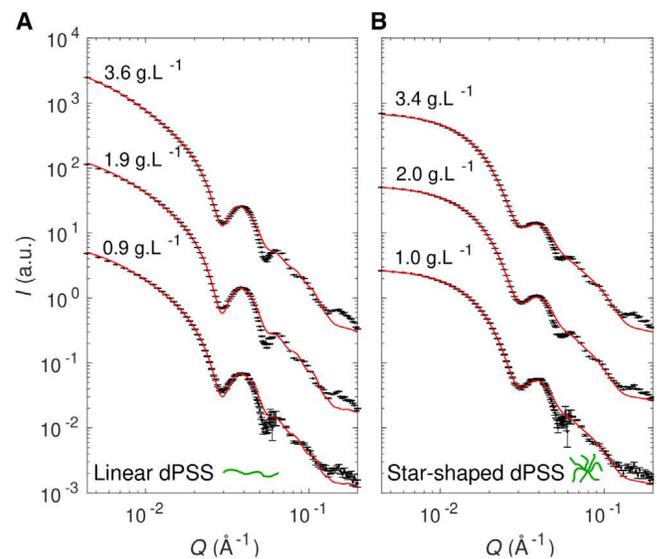


FIGURE 7 Effect of polyelectrolyte topology on capsid structure. SANS patterns of capsids packaging linear (A) and star (B) dPSS in 100%  $D_2O$  in moderate ionic strength buffer (50 mM NaCl, 50 mM Tris-HCl (pD 7.5)). The CP concentrations are indicated next to each curve, and the CP/PSS mass ratio was set to 6. The scattering data were fitted with a polydisperse vesicle model (red solid line; see [Materials and methods](#) for details). To see this figure in color, go online.

**TABLE 2** Characteristics of capsids packaging dPSS inferred from SANS patterns

Sample	CP concentration ( $\text{g} \cdot \text{L}^{-1}$ )	$\langle N \rangle_{\text{Porod}}$	$\Delta R/R$ (%)
Linear dPSS	0.9	58	8.5
	1.9	59	9.5
	3.6	59	16
Star dPSS	1.0	68	30
	2.0	64	34
	3.4	66	41

$\langle N \rangle_{\text{Porod}}$  is the mean number of subunits in the capsids estimated from the Porod volume, and  $\Delta R/R$  is the polydispersity on the core radius of a vesicle model.

than those with linear dPSS. To obtain a measure of this morphological variability, at least in an approximate fashion, the experimental scattering curves were fitted with a model of polydisperse vesicles (*red curves* in Fig. 7). The relative SD on the core radius,  $\Delta R/R$ , turned out to be two or three-fold higher for capsids with star dPSS than those with linear dPSS, which means that the latter were more uniform across the population and/or with fewer defects than the former. Nevertheless, the SANS patterns on Fig. 7A exhibited an increase of intensity at very small  $Q$ -values, which is characteristic of an attractive interaction between the capsids, fairly reproduced by introducing a sticky hard sphere interaction potential in the structure factor of the model of polydisperse vesicles (*red curves* in Fig. 7A). An explanation might be that the long chains of linear dPSS (contour length  $>150$  nm) partly exited out of the capsid and linked the neighboring capsids, which was much less likely with star dPSS (end-to-end distance  $\sim 20$  nm).

## CONCLUSIONS

This study compared the structure of CCMV nucleocapsids packaging various RNA segments with similar lengths but decreasing levels of compactness; C2 belongs to CCMV genome, B1 comes from a close relative, namely BMV, and RF2 is an unrelated segment derived from rotavirus genome. ITC measurements showed that the binding affinity of CCMV CPs for each RNA was of the same magnitude, which indicated that there was no specific mechanism of recognition in the form of packaging signals. Yet, SAXS patterns showed that the nucleocapsids packaging C2 and B1 exhibited a higher degree of order than those packaging RF2. A closer look at the single-particle level by cryoTEM revealed that even in the latter case, some spherical nucleocapsids resembling native virions were formed, whereas the majority of the other particles seemed partially assembled or amorphous. Because RF2 had a lower number of branch points than C2, theoretical calculations revealed that the free energy for packaging RF2 was slightly higher than for packaging C2. The difference was, however, mild and could explain the presence of a significant fraction of spherical nucleocapsids in both cases.

The last set of experiments carried out with synthetic polyelectrolytes demonstrated that a branched topology does not always guarantee a good packaging in terms of morphological structure of capsids. Indeed, capsids packaging linear polyelectrolytes were more ordered and/or less variable than those enclosing star polyelectrolytes carrying the same number of charges. This finding is in good agreement with competition experiments demonstrating that linear, structureless polyU were packaged preferentially in comparison with viral RNA of the same length (28). In a previous work, we reported that an average number of eight linear chains of this molecular weight was packaged within each capsid (71). Therefore, one can imagine that it may be energetically less favorable to coat the interior of a capsid with a few star polyelectrolytes, whereas linear chains can readily cover the surface because each monomer can sit next to the wall. There may be a tradeoff between the compactness offered by a branched topology and the easiness of coating the capsid interior to neutralize the RNA-binding domain of CPs. Additionally, the electrostatic repulsion between star polyelectrolytes should be stronger because of their compactness.

In a biological context, CCMV virions self-assemble inside the host cell near the endoplasmic reticulum (6,21), that is, in a dense and crowded milieu. Because the viral genome lacks packaging signals ensuring its specific recognition by CPs, the virus must have developed a strategy to prevent the uptake of cellular RNA, which would jeopardize its survival. Our work confirms that the compactness of viral RNA can favor its packaging, but it does not imply an all-or-nothing process in which less compact exogenous RNA will be simply rejected. Instead, nucleocapsids packaging nonviral RNA have a larger variability and are mostly partially assembled. Because CPs can rapidly unbind RNA segments as long as capsids are not closed up, it gives a competitive advantage to more compact viral RNA that will capture CPs by completing more virions. Nonetheless, the most powerful driving force for the selectivity of viral genome might be the spatial proximity between the RNA replication site and that of CP translation. The segments of viral RNA might be packaged near the exit of ribosomes by a highly concentrated flow of CPs, thus ensuring the purity of the RNA content and facilitating the proper assembly of icosahedral capsids through the chemical potential (65).

## SUPPORTING MATERIAL

Supporting material can be found online at <https://doi.org/10.1016/j.bpj.2021.08.021>.

## AUTHOR CONTRIBUTIONS

L.M., L.G., R.L.R., C.S., and G.T. performed the experiments. L.M., K.K., C.S., S.B., and G.T. analyzed the data. Y.D., S.P., and R.Z. carried out the simulations. L.M. and G.T. wrote the article.

## ACKNOWLEDGMENTS

Plasmid coding for CCMV RNA 2 was kindly provided by Christian Beren at University of California Los Angeles. Plasmid coding for BMV RNA 1 was provided by Rees Garmann at Harvard University. We are grateful to Didier Poncet for the supply of plasmid coding for rotavirus RF RNA 2 and for his help in RNA production. We thank Jéril Degrouard, Aurélien Thureau and Anne Martel for their technical assistance with cryoTEM, SAXS and SANS, respectively.

L.M., R.L.R., S.B., and G.T. acknowledge financial support from the Agence Nationale de la Recherche (contract ANR-16-CE30-0017-01). K.K. is supported by Agence Nationale de Recherches sur le Sida et les hépatites virales (ANRS) | Maladies infectieuses émergentes. Y.D., S.P., and R.Z. are supported by the National Science Foundation through award DMR-1719550. The electron microscopy imaging is supported by “Investissements d’Avenir” LabEx PALM (contract ANR-10-LABX-0039-PALM). We also acknowledge the SOLEIL synchrotron and the Institut Laue Langevin for allocation of beam time on the SWING and D22 beam-lines, respectively. Financial support from the IR-RMN-THC FR 3050 CNRS for conducting the research is gratefully acknowledged.

## REFERENCES

- Zhang, Y.-Z., Y.-M. Chen, ..., E. C. Holmes. 2019. Expanding the RNA virosphere by unbiased metagenomics. *Annu. Rev. Virol.* 6:119–139.
- Roos, W. H., R. Bruinsma, and G. J. L. Wuite. 2010. Physical virology. *Nat. Phys.* 6:733–743.
- Bruinsma, R. F., and W. S. Klug. 2015. Physics of viral shells. *Annu. Rev. Condens. Matter Phys.* 6:245–268.
- Zandi, R., B. Dragnea, ..., R. Podgornik. 2020. On virus growth and form. *Phys. Rep.* 847:1–102.
- Speir, J. A., and J. E. Johnson. 2012. Nucleic acid packaging in viruses. *Curr. Opin. Struct. Biol.* 22:65–71.
- Comas-Garcia, M. 2019. Packaging of genomic RNA in positive-sense single-stranded RNA viruses: a complex story. *Viruses.* 11:253.
- Perlmutter, J. D., and M. F. Hagan. 2015. Mechanisms of virus assembly. *Annu. Rev. Phys. Chem.* 66:217–239.
- Garmann, R. F., M. Comas-Garcia, ..., W. M. Gelbart. 2016. Physical principles in the self-assembly of a simple spherical virus. *Acc. Chem. Res.* 49:48–55.
- Chevreuil, M., D. Law-Hine, ..., G. Tresset. 2018. Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte. *Nat. Commun.* 9:3071.
- Bruinsma, R. F., G. J. L. Wuite, and W. H. Roos. 2021. Physics of viral dynamics. *Nat. Rev. Phys.* 3:76–91.
- Johnson, J. M., J. Tang, ..., A. Zlotnick. 2005. Regulating self-assembly of spherical oligomers. *Nano Lett.* 5:765–770.
- Lavelle, L., M. Gingery, ..., J. Ruiz-Garcia. 2009. Phase diagram of self-assembled viral capsid protein polymorphs. *J. Phys. Chem. B.* 113:3813–3819.
- Tresset, G., C. Le Coeur, ..., S. Bressanelli. 2013. Norovirus capsid proteins self-assemble through biphasic kinetics via long-lived stable intermediates. *J. Am. Chem. Soc.* 135:15373–15381.
- Ceres, P., and A. Zlotnick. 2002. Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry.* 41:11525–11531.
- Kegel, W. K., and P. van der Schoot. 2004. Competing hydrophobic and screened-coulomb interactions in hepatitis B virus capsid assembly. *Biophys. J.* 86:3905–3913.
- Tresset, G., J. Chen, ..., Y. Lansac. 2017. Two-dimensional phase transition of viral capsid gives insights into subunit interactions. *Phys. Rev. Appl.* 7:014005.
- Chen, J., Y. Lansac, and G. Tresset. 2018. Interactions between the molecular components of the cowpea chlorotic mottle virus investigated by molecular dynamics simulations. *J. Phys. Chem. B.* 122:9490–9498.
- Bruinsma, R. F., M. Comas-Garcia, ..., A. Y. Grosberg. 2016. Equilibrium self-assembly of small RNA viruses. *Phys. Rev. E.* 93:032405.
- Schneemann, A. 2006. The structural and functional role of RNA in icosahedral virus assembly. *Annu. Rev. Microbiol.* 60:51–67.
- Beren, C., Y. Cui, ..., W. M. Gelbart. 2020. Genome organization and interaction with capsid protein in a multipartite RNA virus. *Proc. Natl. Acad. Sci. U S A.* 117:10673–10680.
- den Boon, J. A., A. Diaz, and P. Ahlquist. 2010. Cytoplasmic viral replication complexes. *Cell Host Microbe.* 8:77–85.
- Routh, A., T. Domitrovic, and J. E. Johnson. 2012. Host RNAs, including transposons, are encapsidated by a eukaryotic single-stranded RNA virus. *Proc. Natl. Acad. Sci. USA.* 109:1907–1912.
- Patel, N., S. J. White, ..., P. G. Stockley. 2017. HBV RNA pre-genome encodes specific motifs that mediate interactions with the viral core protein that promote nucleocapsid assembly. *Nat. Microbiol.* 2:17098.
- Twarock, R., R. J. Bingham, ..., P. G. Stockley. 2018. A modelling paradigm for RNA virus assembly. *Curr. Opin. Virol.* 31:74–81.
- Qu, F., and T. J. Morris. 1997. Encapsidation of turnip crinkle virus is defined by a specific packaging signal and RNA size. *J. Virol.* 71:1428–1435.
- Comas-Garcia, M., R. D. Cadena-Nava, ..., W. M. Gelbart. 2012. In vitro quantification of the relative packaging efficiencies of single-stranded RNA molecules by viral capsid protein. *J. Virol.* 86:12271–12282.
- Cadena-Nava, R. D., M. Comas-Garcia, ..., W. M. Gelbart. 2012. Self-assembly of viral capsid protein and RNA molecules of different sizes: requirement for a specific high protein/RNA mass ratio. *J. Virol.* 86:3318–3326.
- Beren, C., L. L. Dreesens, ..., W. M. Gelbart. 2017. The effect of RNA secondary structure on the self-assembly of viral capsids. *Biophys. J.* 113:339–347.
- Yoffe, A. M., P. Prinsen, ..., A. Ben-Shaul. 2008. Predicting the sizes of large RNA molecules. *Proc. Natl. Acad. Sci. USA.* 105:16153–16158.
- Gopal, A., D. E. Eggecioglu, ..., W. M. Gelbart. 2014. Viral RNAs are unusually compact. *PLoS One.* 9:e105875.
- Tubiana, L., A. L. Božič, ..., R. Podgornik. 2015. Synonymous mutations reduce genome compactness in icosahedral ssRNA viruses. *Biophys. J.* 108:194–202.
- Singaram, S. W., R. F. Garmann, ..., A. Ben-Shaul. 2015. Role of RNA branchedness in the competition for viral capsid proteins. *J. Phys. Chem. B.* 119:13991–14002.
- Li, S., G. Erdemci-Tandogan, ..., R. Zandi. 2017. Impact of a nonuniform charge distribution on virus assembly. *Phys. Rev. E.* 96:022401.
- Perlmutter, J. D., M. R. Perkett, and M. F. Hagan. 2014. Pathways for virus assembly around nucleic acids. *J. Mol. Biol.* 426:3148–3165.
- Erdemci-Tandogan, G., J. Wagner, ..., R. Zandi. 2016. Effects of RNA branching on the electrostatic stabilization of viruses. *Phys. Rev. E.* 94:022408.
- Dong, Y., S. Li, and R. Zandi. 2020. Effect of the charge distribution of virus coat proteins on the length of packaged RNAs. *Phys. Rev. E.* 102:062423.
- Sivanandam, V., D. Mathews, ..., A. L. N. Rao. 2016. Functional analysis of the N-terminal basic motif of a eukaryotic satellite RNA virus capsid protein in replication and packaging. *Sci. Rep.* 6:26328.
- Caspar, D. L., and A. Klug. 1962. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27:1–24.
- Lorman, V. L., and S. B. Rochal. 2007. Density-wave theory of the capsid structure of small icosahedral viruses. *Phys. Rev. Lett.* 98:185502.
- Speir, J. A., S. Munshi, ..., J. E. Johnson. 1995. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined

- by X-ray crystallography and cryo-electron microscopy. *Structure*. 3:63–78.
41. Choi, Y. G., and A. L. N. Rao. 2003. Packaging of bromo mosaic virus RNA3 is mediated through a bipartite signal. *J. Virol.* 77:9750–9757.
  42. Ali, A., and M. J. Roossinck. 2007. Rapid and efficient purification of Cowpea chlorotic mottle virus by sucrose cushion ultracentrifugation. *J. Virol. Methods*. 141:84–86.
  43. Porterfield, J. Z., and A. Zlotnick. 2010. A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance. *Virology*. 407:281–288.
  44. Vink, H. 1981. A new convenient method for the synthesis of poly(styrenesulfonic acid). *Makromol. Chem.* 182:279–281.
  45. Punjani, A., J. L. Rubinstein, ..., M. A. Brubaker. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*. 14:290–296.
  46. Franke, D., M. V. Petoukhov, ..., D. I. Svergun. 2017. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Cryst.* 50:1212–1225.
  47. Breßler, I., J. Kohlbrecher, and A. F. Thünemann. 2015. SASfit: a tool for small-angle scattering data analysis using a library of analytical expressions. *J. Appl. Cryst.* 48:1587–1598.
  48. van der Schoot, P., and R. Zandi. 2013. Impact of the topology of viral RNAs on their encapsulation by virus coat proteins. *J. Biol. Phys.* 39:289–299.
  49. de Gennes, P.-G. 1979. *Scaling Concepts in Polymer Physics*. Cornell University Press, Ithaca, NY.
  50. Li, S., H. Orland, and R. Zandi. 2018. Self consistent field theory of virus assembly. *J. Phys. Condens. Matter*. 30:144002.
  51. Borukhov, I., D. Andelman, and H. Orland. 1998. Random polyelectrolytes and polyampholytes in solution. *Eur. Phys. J. B*. 5:869–880.
  52. Siber, A., and R. Podgornik. 2008. Nonspecific interactions in spontaneous assembly of empty versus functional single-stranded RNA viruses. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 78:051915.
  53. Erdemci-Tandogan, G., J. Wagner, ..., R. Zandi. 2014. RNA topology remodels electrostatic stabilization of viruses. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 89:032707.
  54. Wagner, J., G. Erdemci-Tandogan, and R. Zandi. 2015. Adsorption of annealed branched polymers on curved surfaces. *J. Phys. Condens. Matter*. 27:495101.
  55. Borukhov, I., D. Andelman, and H. Orland. 1995. Polyelectrolyte solutions between charged surfaces. *Europhys. Lett.* 32:499–504.
  56. Shafir, A., D. Andelman, and R. R. Netz. 2003. Adsorption and depletion of polyelectrolytes from charged surfaces. *J. Chem. Phys.* 119:2355–2362.
  57. Siber, A., A. L. Božič, and R. Podgornik. 2012. Energies and pressures in viruses: contribution of nonspecific electrostatic interactions. *Phys. Chem. Chem. Phys.* 14:3746–3765.
  58. Lubensky, T. C., and J. Isaacson. 1979. Statistics of lattice animals and dilute branched polymers. *Phys. Rev. A*. 20:2130–2146.
  59. Lee, S. I., and T. T. Nguyen. 2008. Radial distribution of RNA genomes packaged inside spherical viruses. *Phys. Rev. Lett.* 100:198102.
  60. Elleuch, K., F. Lequeux, and P. Pfeuty. 1995. Crosslink effects on equilibrium polymers. *J. Phys. I France*. 5:465–474.
  61. Troupin, C., A. Dehée, ..., A. Garbarg-Chenon. 2010. Rearranged genomic RNA segments offer a new approach to the reverse genetics of rotaviruses. *J. Virol.* 84:6711–6719.
  62. Durand, D., C. Vivès, ..., F. Fieschi. 2010. NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol.* 169:45–53.
  63. Gopal, A., Z. H. Zhou, ..., W. M. Gelbart. 2012. Visualizing large RNA molecules in solution. *RNA*. 18:284–299.
  64. Comas-Garcia, M., R. F. Garmann, ..., W. M. Gelbart. 2014. Characterization of viral capsid protein self-assembly around short single-stranded RNA. *J. Phys. Chem. B*. 118:7510–7519.
  65. Panahandeh, S., S. Li, ..., R. Zandi. 2020. How a virus circumvents energy barriers to form symmetric shells. *ACS Nano*. 14:3170–3180.
  66. Garmann, R. F., M. Comas-Garcia, ..., W. M. Gelbart. 2014. Role of electrostatics in the assembly pathway of a single-stranded RNA virus. *J. Virol.* 88:10472–10479.
  67. Maassen, S. J., J. Huskens, and J. J. L. M. Cornelissen. 2019. Elucidating the thermodynamic driving forces of polyanion-templated virus-like particle assembly. *J. Phys. Chem. B*. 123:9733–9741.
  68. Goobes, R., G. Goobes, ..., P. S. Stayton. 2006. Thermodynamics of statherin adsorption onto hydroxyapatite. *Biochemistry*. 45:5576–5586.
  69. Marichal, L., J. Degrouard, ..., S. Pin. 2020. From protein corona to colloidal self-assembly: the importance of protein size in protein-nanoparticle interactions. *Langmuir*. 36:8218–8230.
  70. Lorenz, R., S. H. Bernhart, ..., I. L. Hofacker. 2011. ViennaRNA package 2.0. *Algorithms Mol. Biol.* 6:26.
  71. Tresset, G., M. Tatou, ..., L. Porcar. 2014. Weighing polyelectrolytes packaged in viruslike particles. *Phys. Rev. Lett.* 113:128305.