



**HAL**  
open science

# Phénomènes dynamiques d'auto-assemblage et de désassemblage dans des virus icosaédriques.

Maelenn Chevreuil

► **To cite this version:**

Maelenn Chevreuil. Phénomènes dynamiques d'auto-assemblage et de désassemblage dans des virus icosaédriques.. Matière Molle [cond-mat.soft]. Université Paris-Saclay, 2020. Français. NNT : 2020UPASS043 . tel-03362693

**HAL Id: tel-03362693**

**<https://theses.hal.science/tel-03362693v1>**

Submitted on 2 Oct 2021

**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.

# Phénomènes dynamiques d'auto- assemblage et de désassemblage dans des virus icosaédriques

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

École doctorale n°564, Physique en Ile de France  
Spécialité de doctorat : Physique  
Unité de recherche : Université Paris-Saclay, CNRS, Laboratoire de Physique des Solides,  
91405, Orsay, France.  
Réfèrent : Faculté des sciences

Thèse présentée et soutenue à Orsay, le 24/02/20, par

**Maelenn CHEVREUIL**

### Composition du Jury

<b>Frank GABEL</b> Directeur de recherche, Institut de Biologie Structurale	Président & Rapporteur
<b>Cendrine MOSKALENKO</b> Maître de conférences, Laboratoire de Physique (UMR 5672)	Rapportrice
<b>Lauriane LECOQ</b> Chargée de recherche, Molecular Microbiology and Structural Biochemistry	Examinatrice
<b>Thomas BIZIEN</b> Scientifique de ligue, Synchrotron SOLEIL	Examineur
<b>Guillaume TRESSET</b> Directeur de recherche, Laboratoire de Physique des Solides	Directeur de thèse
<b>Stéphane BRESSABELI</b> Directeur de recherche, Institut de Biologie Intégrative de la Cellule	Co-Directeur de thèse
<b>Karen PERRONET</b> Chargée de recherche, Institut d'Optique Graduate School	Co-encadrante & Invitée



*Phénomènes dynamiques d'auto-assemblage et de désassemblage dans des virus icosaédriques.*

Maelenn Chevreuil © 24 février 2020

Thèse de doctorat en Physique

Rapporteurs : Frank Gabel et Cendrine Moskalenko

Directeur de thèse : Dr. Guillaume Tresset

Co-Directeur de thèse : Dr. Stéphane Bressanelli

**Université Paris Saclay**

École doctorale Physique en Ile de France

**Équipe SOBIO**

LPS - CNRS UMR8502

1 rue Nicolas Appert

91405 Orsay



« 石の上にも三年 »

– Proverbe Japonais

*Traduction littérale : **Trois années assise sur une pierre.***

*Autrement dit, persévérance est gage de réussite...*



# Table des matières

Table des matières	v
Table des figures	xii
Liste des tableaux	xviii
Abbréviations	xix
<b>A Communications scientifiques</b>	<b>1</b>
1 Articles	2
2 Présentations orales	3
3 Posters	4
<b>I Introduction générale</b>	<b>6</b>
<b>II Les virus</b>	<b>9</b>
1 Introduction sur les virus	10
1.1 Description des virus . . . . .	10
1.2 Classification des virus . . . . .	11
1.2.1 Classification de Baltimore . . . . .	11
1.2.2 Les différentes formes de capsides . . . . .	12
2 Le CCMV	16



2.1	Composition et structure du CCMV . . . . .	16
2.1.1	Le génome . . . . .	16
2.1.2	La capside . . . . .	17
2.1.3	Diagramme de phase de la protéine . . . . .	18
2.2	Production et purification . . . . .	19
2.2.1	Les solutions tampons usuelles . . . . .	19
2.2.2	Production et purification du virion . . . . .	20
2.2.3	Plantation et Infection . . . . .	20
2.2.4	Purification . . . . .	21
2.3	Purification des protéines . . . . .	23
<b>3</b>	<b>Le virus de l'hépatite B</b>	<b>25</b>
3.1	Epidémiologie . . . . .	25
3.1.1	Prévalence . . . . .	25
3.1.2	Transmission . . . . .	26
3.2	Cycle viral . . . . .	26
3.3	Composition et structure du VHB . . . . .	27
3.3.1	La protéine native et sa capside . . . . .	28
3.3.2	Dispersité de la protéine tronquée . . . . .	29
<b>4</b>	<b>Applications des virus à divers domaines</b>	<b>30</b>
4.1	Intérêt médical . . . . .	30
4.1.1	Transfert génétique . . . . .	30
4.1.2	Vecteurs de médicaments . . . . .	31
4.2	Biotechnologie . . . . .	31
4.2.1	Nano-réacteurs . . . . .	31
4.2.2	Matériaux nano-structurés . . . . .	32
<b>III</b>	<b>Diffusion de rayons X aux petits angles</b>	<b>33</b>
<b>1</b>	<b>Principe de base de la diffusion aux petits angles</b>	<b>34</b>
1.1	Interaction des rayons X avec les échantillons . . . . .	34
1.2	Intensité et contraste . . . . .	36
1.2.1	Cas d'une molécule seule . . . . .	36
1.2.2	Cas de N molécules identiques . . . . .	37
1.3	Facteur de structure . . . . .	38

1.4	Facteur de forme et fonction de distribution de distance . . . . .	40
1.5	Approximation de Guinier . . . . .	42
1.6	Diffusion de neutrons aux petits angles . . . . .	43
<b>2</b>	<b>La diffusion de rayons X aux petits angles résolue en temps</b>	<b>45</b>
2.1	Les synchrotrons . . . . .	45
2.2	Le stopped-flow . . . . .	47
2.2.1	Le principe . . . . .	47
2.2.2	Les équipements . . . . .	48
2.2.3	Les précautions à prendre . . . . .	49
<b>3</b>	<b>Outils d'analyse des données SAXS</b>	<b>51</b>
3.1	Outils mathématiques pour l'assemblage des capsides pleines . . .	51
3.1.1	Détermination de $\langle N \rangle_{up}$ et $\langle N \rangle(t \rightarrow +\infty)$ . . . . .	51
3.1.2	Modèle d'assemblage . . . . .	53
3.2	Décomposition en valeurs singulières . . . . .	56
3.2.1	Traitement mathématique . . . . .	56
3.2.2	Procédure de rotation de matrice . . . . .	59
<b>IV</b>	<b>Cinétique d'auto-assemblage du CCMV</b>	<b>64</b>
<b>1</b>	<b>Étude bibliographique de l'assemblage et du désassemblage des capsides vides du CCMV</b>	<b>65</b>
1.1	Des simulations de dynamique moléculaire... . . . .	65
1.2	...aux observations expérimentales. . . . .	66
1.2.1	Assemblage via le mécanisme <i>en masse</i> ? . . . . .	66
1.2.2	Stoechiométrie requise pour l'assemblage . . . . .	67
1.2.3	Liaison réversible . . . . .	70
1.3	Conclusion et discussion . . . . .	71
<b>2</b>	<b>Analyse par Thermo-fluorescence du CCMV</b>	<b>73</b>
2.1	Le criblage de thermostabilité des protéines . . . . .	73
2.1.1	Avantages et inconvénients . . . . .	73
2.1.2	Le principe du CTPF . . . . .	74
2.2	Modèle de champ moyen associé . . . . .	75
2.3	Estimation des températures de dénaturation . . . . .	76

2.3.1	Cas des capsides vides . . . . .	77
2.3.2	Cas des capsides pleines . . . . .	79
2.4	Conclusion . . . . .	82
<b>3</b>	<b>Dynamique d'assemblage de la capside virale</b>	<b>84</b>
3.1	Auto-assemblage des intermédiaires . . . . .	84
3.1.1	Estimation du nombre de sous-unités . . . . .	84
3.1.2	Formation de l'intermédiaire . . . . .	86
3.1.3	Énergie de liaison au sein de l'intermédiaire . . . . .	88
3.1.4	Morphologie des intermédiaires . . . . .	89
3.1.5	Modèle d'assemblage . . . . .	90
3.2	Relaxation des intermédiaires en capsides . . . . .	94
3.2.1	Formation des capsides . . . . .	94
3.2.2	Énergie d'activation . . . . .	96
3.3	Auto-assemblage des capsides autour du PSS . . . . .	98
3.4	Assemblage <i>in vivo</i> . . . . .	101
3.4.1	Influence du pH? . . . . .	101
3.4.2	Influence des ions divalents? . . . . .	102
3.4.3	Influence de la concentration en protéine? . . . . .	102
3.5	Conclusion . . . . .	103
<b>4</b>	<b>Conclusion du chapitre</b>	<b>105</b>
<b>V</b>	<b>Cinétiques d'auto-assemblage et désassemblage de la capside vide du VHB</b>	<b>108</b>
<b>1</b>	<b>Étude bibliographique de l'assemblage et du désassemblage des capsides vides du VHB</b>	<b>109</b>
1.1	Auto-assemblage des capsides vides du VHB . . . . .	109
1.1.1	Mécanisme de nucléation-croissance . . . . .	109
1.1.2	Intermédiaires tardifs . . . . .	111
1.1.3	Dimorphisme de la capside . . . . .	112
1.1.4	Dynamique d'assemblage de la capside . . . . .	113
1.2	Désassemblage des capsides vides du VHB . . . . .	114
1.3	Conclusion et discussion . . . . .	115
<b>2</b>	<b>Analyse par Thermo-fluorescence</b>	<b>116</b>

2.1	Les particules virales . . . . .	116
2.2	Température de dénaturation des capsides vides . . . . .	117
2.2.1	Énergie d'interaction entre les protéines . . . . .	119
2.2.2	Charge effective des dimères de protéines au sein de la capsid	119
2.3	Conclusion . . . . .	120
<b>3</b>	<b>Désassemblage de la capsid CP149</b>	<b>121</b>
3.1	Expériences réalisées et données obtenues par TR-SAXS . . . . .	121
3.1.1	Préparation des échantillons . . . . .	121
3.1.2	Présentation des données . . . . .	122
3.2	Reconstruction des espèces . . . . .	123
3.2.1	Première phase du désassemblage . . . . .	123
3.2.2	Cinétique entière . . . . .	124
3.3	Evolution reconstruite des espèces . . . . .	126
3.3.1	Modèle de dépolymérisation . . . . .	126
3.3.2	Nature des agrégats . . . . .	128
3.4	Conclusion sur le désassemblage . . . . .	131
<b>4</b>	<b>Assemblage du VHB</b>	<b>133</b>
4.1	Données expérimentales . . . . .	133
4.1.1	Préparation des échantillons . . . . .	133
4.1.2	Les cinétiques réalisées par TR-SAXS . . . . .	134
4.1.3	Les données . . . . .	134
4.2	Premières analyses des données de TR-SAXS . . . . .	136
4.2.1	Concentration critique d'assemblage . . . . .	136
4.2.2	Analyse par SVD . . . . .	138
4.2.3	Comparaison des espèces reconstruites et expérimentales . .	139
4.2.4	Reconstruction des cinétiques . . . . .	142
4.3	Construction d'un modèle d'assemblage de la capsid vide . . . . .	144
4.3.1	Dépendance de la cinétique d'assemblage aux concentrations en protéines et NaCl . . . . .	144
4.3.2	Chemin cinétique d'auto-assemblage des capsides vides . . .	147
4.4	Conclusion . . . . .	153
<b>5</b>	<b>Conclusion du chapitre</b>	<b>154</b>

VI Stratégie de marquage pour l'étude en molécule unique de l'auto-  
assemblage de la capsid pleine du VHB **158**

<b>1 Matériels et méthodes</b>	<b>160</b>
1.1 TIRFM . . . . .	160
1.1.1 Principe du TIRFM . . . . .	160
1.1.2 Préparation de la cellule micro-fluidique . . . . .	161
1.1.3 Accroche des systèmes biologiques . . . . .	162
1.2 Les protéines mutantes de la CP183 . . . . .	163
1.2.1 Les différentes mutations . . . . .	163
1.2.2 Avantages et inconvénients . . . . .	165
1.3 La chromatographie d'exclusion de taille . . . . .	165
1.3.1 Principe de la SEC . . . . .	165
1.3.2 Calibration . . . . .	167
1.4 Protocoles détaillés de production de la protéine native . . . . .	168
1.4.1 Les solutions tampons usuelles . . . . .	168
1.4.2 Cellules compétentes . . . . .	169
1.4.3 Transformation des bactéries . . . . .	170
1.4.4 Expression des protéines . . . . .	171
1.4.5 Purification des protéines . . . . .	172
1.4.6 Désassemblage des capsides . . . . .	174
<b>2 Les différentes stratégies pour avoir des protéines fluorescentes</b>	<b>176</b>
2.1 La protéine mutante GFP-CP183 . . . . .	176
2.1.1 Impact de la GFP sur les premières étapes de production . .	176
2.1.2 Impact du protocole de désassemblage sur la fluorescence des dimères GFP-CP183 . . . . .	178
2.2 Les mutants Lysines . . . . .	185
2.2.1 Présentation du fluorophore . . . . .	185
2.2.2 Protocole de marquage . . . . .	186
2.2.3 Conditions de marquage des mutants lysines . . . . .	188
2.2.4 Marquage du CCMV . . . . .	194
<b>3 L'ARN pré-génomique</b>	<b>196</b>
3.1 Préparation de l'ARN pré-génomique . . . . .	196
3.1.1 Les mutations effectuées sur le ARN-pg . . . . .	196

3.1.2	Production de l'ARN-pg avec insertion . . . . .	197
3.2	Résultats préliminaires en TIRFM . . . . .	198
<b>4</b>	<b>Conclusion du chapitre</b>	<b>201</b>
<b>VII</b>	<b>Conclusion et perspectives</b>	<b>204</b>
	<b>Bibliographie</b>	<b>207</b>

# Table des figures

2.2	Classification de Baltimore. . . . .	11
2.3	Description des axes de symétrie dans une capsid e icosaédrique. . . . .	12
2.4	Illustration de la variété des capsides des virus. . . . .	13
2.5	Exemples de virus avec une capsid e hélicoïdale . . . . .	13
2.6	Modèle d'assemblage des protéines . . . . .	13
2.7	Exemple de calculs des paramètres $h$ et $k$ et du nombre $T$ . . . . .	14
2.8	Génome segmenté du CCMV . . . . .	17
2.9	Capsides de CCMV à $pH$ 4,8 et $pH$ 7,5 . . . . .	17
2.10	Monomères du CCMV. . . . .	18
2.11	Diagramme de phase des protéines du CCMV. . . . .	18
2.12	Photo d'une cornille infectée. . . . .	21
2.13	Prévalence du virus de l'hépatite B dans le monde. . . . .	26
2.14	Illustration schématique du cycle de vie du VHB. . . . .	27
2.15	Capsides du VHB . . . . .	28
2.16	Structure d'un dimère de VHB. . . . .	28
2.17	Analyse de la dispersité en taille des capsides vides (CP149) du VHB	29
3.1	Motif de diffusion 2D (échantillon isotrope). . . . .	35
3.2	Vecteur de diffusion $\vec{q} = \vec{k}_s - \vec{k}_i$ . . . . .	35
3.3	Exemple de courbes obtenues par SAXS . . . . .	39
3.4	Profil SAXS d'une solution polydisperse en taille . . . . .	41
3.5	Exemples d'intensités de diffusion et fonctions de distribution de distance . . . . .	42
3.6	Représentation de Guinier . . . . .	43
3.7	Représentation schématique de la variation de contraste en SANS . . . . .	44

3.8	Échelle de temps accessibles par différentes techniques lors d'études de mélanges réactionnels. . . . .	47
3.9	Photo de l'appareil SFM 4000. . . . .	48
3.10	Schéma de fonctionnement du SFM-4000 . . . . .	48
3.11	Concentration critique à l'équilibre . . . . .	54
3.12	Évolution des temps caractéristiques de la dynamique d'assemblage des NPCs pour différentes valeurs de $\rho$ . . . . .	55
3.13	Schéma de principe de l'analyse par décomposition en valeurs singulières (SVD) . . . . .	57
3.14	Présentation des critères de SVD utilisés pour l'identification du nombre de composantes significatives . . . . .	59
4.1	Mécanismes de croissance de la capsid mis en lumière par simulations de dynamique moléculaire . . . . .	65
4.2	Images Cryo-EM des réactions d'assemblage réalisées par Garmann et al. . . . .	67
4.4	Expériences de d'assemblage réalisées par Johnson et al.. . . . .	69
4.5	Expériences de compétition réalisées par Comas-Garcia et al. . . . .	70
4.6	Schéma de principe du criblage de thermo-stabilité par fluorescence. . . . .	74
4.7	Courbes dérivées des intensités de fluorescence en fonction de la température. . . . .	74
4.8	Tests de dissociation thermique en fonction de la salinité pour les capsides vides. . . . .	77
4.9	Tests de dissociation thermique en fonction du $pH$ pour les capsides vides. . . . .	78
4.10	Tests de dissociation thermique en présence de $CaCl_2$ pour les capsides vides. . . . .	79
4.11	Tests de dissociation thermique pour les virions natifs. . . . .	81
4.12	Tests de dissociation thermique pour les capsides contenant du PSS. . . . .	82
4.13	Estimation du nombre de sous-unités dans les NPCs . . . . .	85
4.14	Dynamique d'assemblage des NPCs . . . . .	87
4.15	Évolution des temps caractéristiques de la dynamique d'assemblage des NPCs pour différentes valeurs de $\rho$ . . . . .	88
4.16	Dynamique de relargage des dimères au sein de l'intermédiaire. . . . .	89



4.17	Etude des NPCs à l'équilibre ( $\rho = 6$ ) . . . . .	90
4.18	$\langle N \rangle$ obtenus proche de l'équilibre pour différents $\rho$ . . . . .	92
4.19	Dynamique d'assemblage des NPCs ( $\rho = 8$ ) . . . . .	93
4.20	Dynamique de relaxation des NPCs . . . . .	95
4.21	Étude du temps caractéristique de relaxation structurelle . . . . .	95
4.22	Illustration schématique du mécanisme <i>synchrone</i> . . . . .	96
4.23	Logarithme de $\tau_{bind}$ et de $\tau_{bind}^s$ en fonction de l'inverse de la température pour des expériences effectuées de 10 à 40°C. . . . .	97
4.24	Images cryo-TEM des virions reformés après plusieurs jours d'incubation. . . . .	98
4.25	Dynamique d'auto-assemblage des capsides pleines de PSS . . . . .	99
4.26	Images Cryo-TEM des capsides pleines de PSS . . . . .	100
4.27	Dynamique d'assemblage des NPCs en présence d'ions divalents. . . . .	102
4.28	Superposition des courbes des NPCs obtenues après dialyse des mélanges réactionnels à différents concentrations. . . . .	103
4.29	Représentation schématique du paysage énergétique. . . . .	104
5.1	Schéma de l'assemblage de la capside du VHB proposé par Zlotnick et al. . . . .	110
5.2	Schéma de l'assemblage de la capside du VHB proposé par Holmes et al. . . . .	110
5.3	Intermédiaires d'assemblage proposés par Zlotnick et al. . . . .	111
5.4	Intermédiaires d'assemblage observés par Kondylis et al. . . . .	112
5.5	Dimorphismes observés par Sun et al. . . . .	112
5.6	Observations par Zlotnick et al. de la dépendance de la dynamique de l'assemblage aux concentrations en sel et protéines . . . . .	114
5.7	Schéma de la capside du VHB et d'un dimère de protéine . . . . .	116
5.8	Courbes issues des tests de stabilité thermique par fluorescence sur des capsides vides de VHB. . . . .	118
5.9	Représentation tridimensionnelle des intensités mesurées lors du désassemblage . . . . .	122
5.10	Données TR-SAXS pour la première phase du désassemblage de la capside vide du VHB . . . . .	124
5.11	Graphique superposant les facteurs de forme des espèces expérimentales et reconstruites de la première phase de la cinétique. . . . .	124

5.12	Données TR-SAXS pour la totalité de la cinétique de désassemblage des capsides vides. . . . .	125
5.13	Evolution, en fonction du temps, des espèces reconstruites lors du désassemblage . . . . .	126
5.14	Schéma de l'évolution des espèces au cours de la cinétique de désassemblage. . . . .	128
5.15	Graphiques . . . . .	129
5.16	Illustration schématique non-exhaustive des informations qui peuvent être obtenue à partir d'une courbe SAXS. . . . .	130
5.17	Facteur de structure des agrégats. . . . .	131
5.18	Illustration schématique de l'évolution des espèces au cours de la cinétique de désassemblage. . . . .	131
5.19	Courbes superposées lors des cinétiques d'assemblages réalisées par TR-SAXS . . . . .	135
5.20	Figures illustrant l'influence de la concentration en protéines et en NaCl sur l'assemblage des capsides vides de CP149 . . . . .	137
5.21	Présentation résumée des critères de SVD utilisés pour l'identification du nombre de composantes significatives . . . . .	139
5.22	Comparaison des facteurs de forme des espèces reconstruites à partir de l'analyse par SVD avec les courbes expérimentales de formes de la capside et du dimère . . . . .	140
5.23	Comparaison des intensités diffusées des intermédiaires identifiés par Pierson et al. et de la capside $T = 4$ . . . . .	141
5.24	Reconstruction de l'évolution massique finale en capside en fonction des conditions expérimentales . . . . .	142
5.25	Exemple de cinétique d'auto-assemblage de la capside du VHB . . . . .	145
5.26	Comparaison de l'évolution, en fonction du temps, de l'intensité $I(0)$ pour différentes conditions expérimentales . . . . .	145
5.27	Rapport $\tau_{I(0)}/\tau_{R_g}$ en fonction des conditions expérimentales . . . . .	147
5.28	Dynamique d'assemblage et de structuration des dimères en capsides	148
5.29	Comparaison de l'évolution de l'intensité $I(0)$ et de la valeur du premier minimum des facteurs de formes . . . . .	150
5.30	Intensité diffusées théoriques des capsides $T = 3$ et $T = 4$ ainsi que des mélanges de $T = 3$ et $T = 4$ . . . . .	151

5.31	Figure illustrant le dimorphisme observé par Sun et al. . . . .	152
5.32	Illustration schématique de la dynamique d'assemblage . . . . .	153
6.1	Illustration schématique du principe du TIRFM . . . . .	158
6.2	Principe du TIRFM . . . . .	161
6.3	Illustration schématique d'une cellule micro-fluidique. . . . .	161
6.4	Illustration schématique des étapes de chimie de surface. . . . .	163
6.5	Structures 3D des protéines mutantes de la CP183 . . . . .	164
6.6	Illustration schématique du principe de la SEC . . . . .	166
6.7	Courbes de calibration de la colonne S200 10/300 GL. . . . .	167
6.8	Schéma du gradient de sucrose . . . . .	172
6.9	Gels SDS-PAGE 14% issus de la purification de la protéine native . . .	174
6.10	Chromatogramme classique après désassemblage de la protéine native	175
6.11	Colonies BL21*CP exprimant le gène de la protéine GFP-CP183. . . .	177
6.12	Contrôle de la qualité du culot bactérien. . . . .	177
6.13	Chromatogramme de la protéine GFP-CP183 après dialyse dans le tampon de désassemblage . . . . .	179
6.14	Gel SDS-PAGE correspondant à l'étape 8 du protocole de purification.	180
6.15	Structure 3D de monomères de GFP-CP183 tronqués . . . . .	180
6.16	Structure 3D de la eGFP et de son chromophore. . . . .	180
6.17	Fluorescence de la eGFP pour différentes concentrations de GuHCl. .	181
6.18	Chromatogramme de la protéine GFP-CP183 dans le tampon avec 0,5 M de GuHCl . . . . .	182
6.19	Chromatogrammes de la protéine GFP-CP183 pour divers <i>p</i> H et concen- trations en sel . . . . .	184
6.20	Images en TIRFM des protéines désassemblées des tampons avec 0,5 M de GuHCl . . . . .	184
6.21	Fluorescence de la eGFP en fonction du pH . . . . .	184
6.22	Réaction de marquage du dimère avec l'ATTO 647N NHS-ester. . . .	186
6.23	Gel SDS-PAGE du marquage de la protéine A80K-CP183 . . . . .	187
6.24	Chromatogramme de la protéine A80K-CP183 après désassemblage (tampon $\text{KH}_2\text{PO}_4$ ). . . . .	188
6.25	Chromatogrammes de la protéine A80K-CP183 purifiée et désassem- blée dans différents tampons . . . . .	191

6.26	Chromatogramme avec la protéine A80K-EEE-CP183 (ratio 10%) . . .	193
6.27	Chromatogramme avec la protéine du CCMV (ratio 10%) . . . . .	194
6.28	Schéma des mutations réalisées sur l'ARN-pg . . . . .	196
6.29	Schémas du processus d'insertion des mutations dans l'ARN-pg . . .	197
6.30	Schémas du processus de production de l'ARN-pg . . . . .	198
6.31	Résultats préliminaires en TIRFM . . . . .	199

# Liste des tableaux

2.1	Tableau récapitulatif des solutions tampons usuelles . . . . .	19
2.2	Tableau récapitulatif des composants des solutions tampons. . . . .	19
3.1	Tableau récapitulatif des caractéristiques et équipements des lignes de lumière SWING et ID02 . . . . .	46
4.1	Tableau récapitulatif de $\tau_{struc}$ et $\tau_{bind}$ en fonction de la température .	97
5.1	Tableau récapitulatif des conditions expérimentales d'assemblage . .	134
5.2	Tableaux des temps caractéristiques des cinétiques d'assemblage . . .	146
5.3	Tableaux des premiers minima des facteurs de forme des capsides issues des cinétiques d'assemblage . . . . .	151
6.1	Tableau récapitulatif des avantages et inconvénients de chacune des protéines mutantes . . . . .	165
6.2	Tableau récapitulatif des étalons pour calibrations . . . . .	167
6.3	Tableau récapitulatif des solutions tampons usuelles . . . . .	168
6.4	Tableau récapitulatif des composants des solutions tampons. . . . .	168
6.5	Tableau récapitulatif des solutions de désassemblage testées . . . . .	182
6.6	Tableau récapitulatif des aliquotes de contrôles du protocole de mar- quage . . . . .	187

# Abbreviations

**ADN-ccc** ADN circulaire covalamment clos.

**ADN-rc** ADN circulaire relâché.

**AFM** Atomic Force Microscope.

**ARN-pg** ARN pré-génomique.

**ARNi** ARN interférents.

**ARNm** ARN messenger.

**BMV** Brome Mosaic Virus.

**BSA** Bovine Serum Albumin.

**CCMV** Cowpea Chlorotic Mottle Virus.

**CHES** N-Cyclohexyl-2-aminoethanesulfonic acid.

**CP** Capsid Protein.

**Cryo-TEM** Cryogenic Transmission Electron Microscopy.

**CTPF** Criblage de Thermostabilité des Protéines par Fluorescence.

**DMSO** DiMethyl SulfOxide.

**DTT** Dithiothréitol.

**EDTA** Éthylène Diamine Tétra-Acétique.

**GFP** Green Fluorescent Protein.

**GuHCl** Chlorure de guanidinium.

**MWCO** Molecular Weight Cut-Off.

**NPC** NucleoProtein Complexes.

**PDB** Protein Data Bank.

**PEG** PolyEthylène-Glycol.

**PMSF** PhenylMethylSulfonyl Fluoride.

**PSS** Polystyrène Sulfonate de Sodium.

**RE** Réticulum Endoplasmique.

**SAXS** Small Angle X-ray Scattering.

**SDS-PAGE** Sodium Dodecyl Sulfate–PolyAcrylamide Gel Electrophoresis.

**SEC** Size Exclusion Chromatography.

**SFM** Stopped-Flow Mixing.

**SLS** Static Light Scattering.

**SVD** Singular Value Decomposition.

**TCV** Turnip Crinkle Virus.

**TIRFM** Total Internal Reflection Fluorescence Microscopy.

**TR-SAXS** Time-Resolved Small-Angle X-ray Scattering.

**UA** Unité Arbitraire.

**VHB** Virus de l'Hépatite B.

A

---

Communications scientifiques



- **Study of the self-assembly and disassembly of the hepatitis B virus empty capsids**, Maelenn Chevreuil, Lauriane Lecoq, Shishan Wang, Naïma Nhiri, Eric Jacquet, Thomas Zinn, Stéphane Bressanelli and Guillaume Tresset, *in preparation*
- **Nanoparticle-Templated Self-Assembly of Viral Capsids Probed by Time-Resolved Absorbance Spectroscopy and X-Ray Scattering**, Andrew Burke, Maelenn Chevreuil, Alisier Paris, Vanessa de La Grange, Claire Goldmann, Javier Pérez, Doru Constantin and Guillaume Tresset, *Physical Review Applied* 2018 <https://doi.org/10.1103/PhysRevApplied.10.054065>
- **Nonequilibrium Self-Assembly Dynamics of Icosahedral Viral Capsids Packaging Genome**, Maelenn Chevreuil, Didier Law-Hine, Jingzhi Chen, Stéphane Bressanelli, Sophie Combet, Doru Constantin, Jéril Degrouard, Johannes Möller, Mehdi Zeghal and Guillaume Tresset, *Nature Communications* 2018 <https://doi.org/10.1038/s41467-018-05426-8>
- **Investigating the thermal dissociation of viral capsid by lattice model**, Jingzhi Chen, Maelenn Chevreuil, Sophie Combet, Yves Lansac and Guillaume Tresset, *Journal of Physics Condensed Matter* 2017 <https://doi.org/10.1088/1361-648X/aa8d88>
- **Two-Dimensional Phase Transition of Viral Capsid Gives Insights into Subunit Interactions**, Guillaume Tresset, Jingzhi Chen, Maelenn Chevreuil, Naïma Nhiri, Eric Jacquet and Yves Lansac, *Physical Review Applied* 2017 <https://doi.org/10.1103/PhysRevApplied.7.014005>

- **Caractérisation du désassemblage des protéines de capsid dérivées du virus de l'hépatite B.** *21st French Days of Virology*. Lyon, France (2019)
- **Characterization of the disassembly of capsid proteins derived from hepatitis B virus.** *Biophysical Society Annual meeting*. Baltimore, USA (2019)
- **Study on the Self-assembly dynamics of viral capsids.** *ED PIF PhD student meeting*. Paris, France (2018)
- **Single-molecule monitoring of Hepatitis B Virus nucleocapsid assembly.** *18th Meeting of the National Hepatitis Network*. Paris, France (2018)

- **Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte.** *27th Young Researcher Symposium Alain Bouyssy.* Orsay, France (2019)
- **Etude de la dissociation thermique de la capside virale du CCMV.** *20st French Days of Virology.* Paris, France (2018)
- **Characterization of the capsid proteins dissociation of a plant virus, the CCMV.** *Workshop on Physical Virology.* Trieste, Italy (2017)

I

---

Introduction générale

Les virus, souvent présentés comme l'entité biologique la plus abondante sur terre en terme de nombre d'individus, occupent une position étrange en biologie. Selon la définition la plus commune, les virus ne sont pas des êtres vivants. En effet, ceux-ci ont absolument besoin d'infecter un hôte afin de pouvoir utiliser le métabolisme et les constituants de celui-ci pour se répliquer. Cependant, au cours des dernières années, les découvertes des virophages<sup>1</sup>, des mimivirus<sup>2</sup> et des pandoravirus<sup>3</sup> ont bousculé cette vision historique[2]. La définition du vivant pourrait ainsi être élargie à une entité qui diminue le niveau d'entropie et se reproduit en commettant des erreurs. De ce fait, les virus pourraient être considérés comme des êtres vivants.

D'un point de vue purement physique, les virus sont simplement des nano-particules ayant de remarquables propriétés mécaniques et structurales. Par exemple, certains virus se sont révélés capables de conserver leur intégrité structurale dans de larges gammes de *pH* ou de force ionique tandis que d'autres sont résistants à des pressions internes pouvant atteindre 100 atmosphères. Une des particularités les plus étonnantes de certaines particules virales est l'auto-assemblage des centaines de composants moléculaires qui les composent, conduisant à des structures de taille et de forme, bien définies. Celles-ci sont un parfait exemple de nano-structures auto-assemblées que l'Homme peine encore à imiter mais dont les applications possibles sont nombreuses. Parmi celles-ci, il y a la phagothérapie, qui est l'utilisation de virus bactériophages afin de traiter certaines maladies (Par ex. la Maladie de Crohn<sup>4</sup>[3]), ou la virothérapie, qui consiste à introduire du matériel génétique chez un individu, par l'intermédiaire d'un rétrovirus, pour traiter une maladie. De plus, du fait de leurs tailles monodisperses, les particules virales peuvent être envisagées comme modèles pour des nanoparticules ou nanofils.

Si l'auto-assemblage et le désassemblage des composants des virus permettent d'envisager de nombreuses applications, les mécanismes sous-jacents sont encore mal compris. En particulier, les chemins cinétiques pour l'assemblage et le désassemblage ne sont pas entièrement résolus. Si des simulations de dynamiques moléculaires ont permis de proposer un éventail de processus possibles,

- 
1. Virus qui infecte un autre virus pour se répliquer.
  2. Virus à ADN, souvent plus gros que bien des bactéries.
  3. Virus géant, possèdent entre 1 900 et 2 500 gènes et dont la complexité est telle qu'ils seraient assimilables à une nouvelle forme de vie[1].
  4. Maladie inflammatoire chronique de l'intestin (MICI) qui peut atteindre toute partie de l'appareil digestif.

les études expérimentales, souvent réalisées à l'équilibre ou ne disposant pas d'une résolution spatiale et temporelle suffisantes, ont échoué à les confirmer ou les infirmer. Cette thèse a pour objectif d'apporter de nouveaux éléments à l'étude de l'auto-assemblage et du désassemblage des virus, étapes clés dans le cycle viral.

La première partie de ce manuscrit sera consacrée à une introduction sur les virus. Il sera notamment question du Cowpea Chlorotic Mottle Virus (CCMV) et du virus de l'hépatite B (VHB) que j'ai étudiés pendant ma thèse. Cette partie se terminera par une introduction succincte aux applications possibles des particules virales dans divers domaines. La deuxième partie du manuscrit s'intéressera à la technique de diffusion de rayons X aux petits angles résolue en temps (TR-SAXS) que j'ai majoritairement utilisée pendant ma thèse. Il sera également question des outils d'analyse associés à l'utilisation de cette technique. La troisième partie du manuscrit mettra en lumière les résultats obtenus pendant la première partie de ma thèse qui portait sur l'étude de l'assemblage des protéines de la capsid du CCMV autour de son génome. Cette partie présentera également les résultats obtenus par Criblage de Thermostabilité des Protéines par Fluorescence (CTPF) pour estimer la charge effective ainsi que l'énergie d'interaction au sein des capsides vides et pleines du CCMV, paramètres clés de la compréhension des mécanismes d'assemblage de ce virus. La quatrième partie du manuscrit sera dédiée aux résultats obtenus lors de l'étude par SAXS de l'auto-assemblage et du désassemblage des capsides vides du VHB. Cette étude met en évidence les difficultés pour étudier ces systèmes *in vitro* ainsi que les méthodes d'analyse permettant de remonter à des informations structurales ou temporelles sans introduire de biais. La dernière partie du manuscrit sera la plus trans-disciplinaire puisqu'elle est consacrée aux développements de protocoles de production et de purification de protéines de capsides fluorescentes, en vue d'une étude en molécule unique de l'auto-assemblage de la capsid pleine du VHB.



# II

---

## Les virus

Cette partie du manuscrit a pour objectif de décrire simplement certains aspects généraux des virus. Dans un premier temps, je ferai une courte introduction sur les virus et présenterai plus précisément les virus icosaédriques. Je présenterai ensuite les deux virus utilisés lors de cette étude. Je terminerai en décrivant quelques applications des capsides virales aux nanotechnologies.



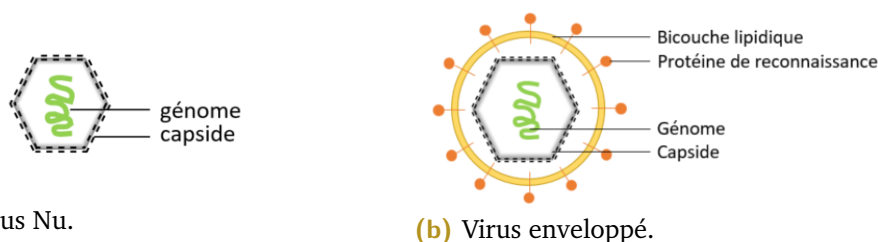
# Introduction sur les virus

La rage, le sida, Ebola, la grippe, la dengue, les hépatites, ... Depuis la découverte des virus à la fin du XIX<sup>ème</sup> siècle, la liste de ces derniers n'a fait que s'allonger. Les biologistes estiment aujourd'hui à  $10^{31}$  le nombre total de particules virales sur Terre[4]. Rien que pour l'être humain, sur les 5500 espèces connues de virus, 219 peuvent l'infecter[5].

## 1.1 Description des virus

Si les virus sont nombreux et leurs stratégies de répliquions multiples, ceux-ci possèdent tout de même des caractéristiques communes. Ainsi, la structure de ces agents infectieux peut se résumer à deux ou trois éléments, selon le type :

- Le génome, situé à l'intérieur du virus qui peut être de l'ADN ou de l'ARN (simple ou double brin).
- La capsid, assemblage de protéines, qui entoure et protège le matériel génétique.
- Une enveloppe qui peut être composée d'une bicouche lipidique ainsi que de protéines de reconnaissance.



**Figure 2.1:** Schéma simplifié des deux grandes familles de virus.

Les virus les plus simples ne sont composés que des deux premiers éléments et sont appelés virus *nus* ou *non-enveloppés* (Fig. 2.1a). C'est notamment le cas du virus de la marbrure chlorotique de la cornille (CCMV, Cowpea Chlorotic Mottle Virus en anglais). Les virus les plus complexes sont composés, au minimum des trois éléments cités précédemment, et sont appelés virus *enveloppés* (Fig. 2.1b). C'est le cas du virus de l'hépatite B (VHB). Cette description est, bien entendu,

minimaliste. En réalité, il existe de multiples types de structure chez les virus. Par exemple, les *Reoviridae* sont des virus non enveloppés avec une structure à trois capsides concentriques. Cette simple description met en évidence la nécessité d'un système de classification commun.

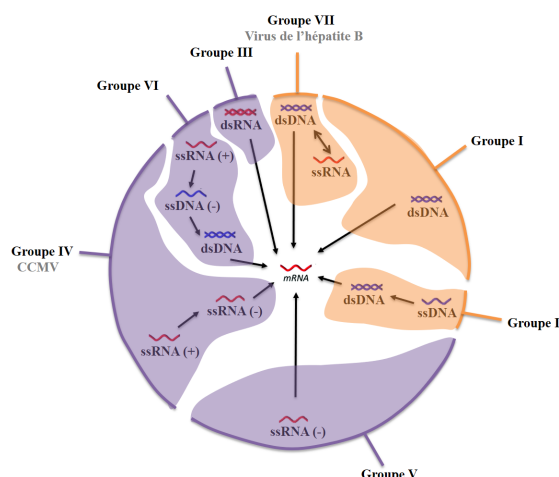
## 1.2 Classification des virus

Comme nous avons pu le voir dans l'introduction de la thèse, l'appartenance des virus au monde vivant est sujette à débat. De ce fait, la classification des virus n'est pas intégrée à celle réalisée pour les êtres vivants. Deux méthodes font autorité chez les virologues :

- La classification de Baltimore que je présenterai ci-après.
- La classification du comité international de taxonomie des virus (ICTV<sup>1</sup>), qui utilise une méthode semblable à celle utilisée pour les êtres vivants. Les virus sont rangés par ordre, famille, sous-famille, genre et espèce.

### 1.2.1 Classification de Baltimore

La classification de Baltimore[6] a été proposée par David Baltimore<sup>2</sup> en 1971. Celle-ci est basée sur la nature du génome des virus ainsi que sur le nombre et la nature des étapes nécessaires à la production de l'ARN messager qui permet



**Figure 2.2:** Classification de Baltimore.

1. International Committee on Taxonomy of Viruses en anglais
2. Lauréat du prix Nobel de médecine en 1975

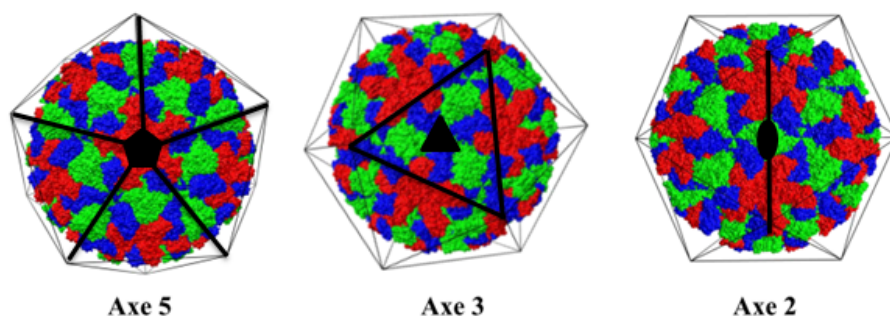
la fabrication des protéines nécessaires à l'assemblage de la capsid virale. Cette classification, visible sur la figure 2.2, montre la diversité des mécanismes généraux de réplication. Ainsi, 7 groupes sont mis en évidence. Le groupe IV concerne les virus à ARN simple brin à sens positif, i.e. le génome du virus peut être directement traduit par le ribosome de l'hôte pour créer des protéines virales. La réplication se produit généralement dans le cytoplasme. Dans cette catégorie, nous retrouvons la famille des *Bromoviridae* dont fait partie le CCMV. Le groupe VII présente un mécanisme plus complexe : la réplication de ces virus à ADN double brin passe par un intermédiaire ARN et nécessite donc une transcription inverse d'ARN simple brin à ADN double brin. La famille des *Hepadnaviridae*, dont fait partie le VHB, illustre bien ces mécanismes complexes que nous détaillerons dans la section 3.2.

## 1.2.2 Les différentes formes de capsides

### 1.2.2.1 une zoologie très variée

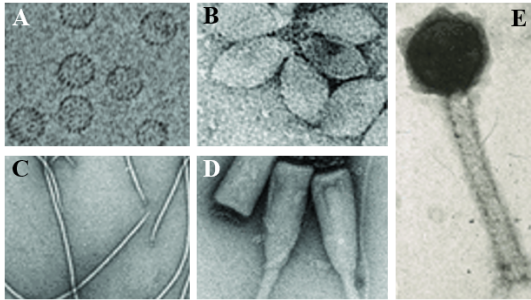
Une autre classification, reposant sur la structure de la capsid des virus, peut également être utilisée[7]. En effet, celle-ci peut être résumée en 4 catégories :

- Les virus hélicoïdaux, où les protéines et le génome s'agencent sous forme d'une hélice (Fig. 2.5 et Fig. 2.4 C).
- Les virus icosaédriques (Fig. 2.4 A). Cette forme, caractérisée par 20 faces et 12 sommets, comporte trois types d'axes de symétrie : six axes d'ordre 5, dix axes d'ordre 3 et quinze axes d'ordre 2, mis en évidence sur la figure 2.3.

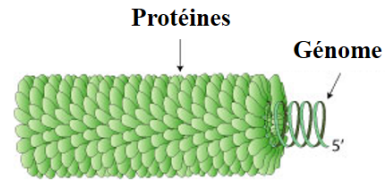


**Figure 2.3:** Description des axes de symétrie dans une capsid icosaédrique. Superposition d'une représentation schématisée, adaptée de Prasad et al.[7], (lignes noires) et d'une capsid icosaédrique, adaptée de [8].

- Les virus à structures mixtes, i.e. composés d'une partie icosaédrique ainsi que d'une partie hélicoïdale. C'est notamment le cas des bactériophages (Fig. 2.4 E).
- Les virus complexes. Dans cette catégorie, se trouvent le VIH avec une capsidie conique. Nous pouvons également retrouver des virus avec une capsidie en forme de bouteille et de citron (Fig. 2.4 B et D).



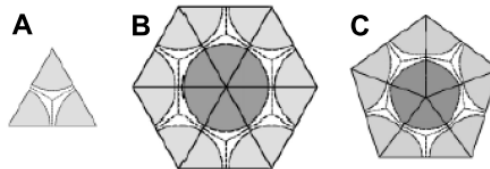
**Figure 2.4:** Exemple de forme des capsides des virus. Adapté de [9].



**Figure 2.5:** Exemple de virus avec une capsidie hélicoïdale. Adapté du site Viral-Zone d'ExpASy (<https://viralzone.expasy.org>).

### 1.2.2.2 Principe de quasi-équivalence de Caspar & Klug

Les axes de symétrie de la forme icosaédrique, présentés précédemment (Fig. 2.3), relient ensemble 60 unités asymétriques identiques. Une capsidie icosaédrique peut donc être assemblée à partir de 60 copies identiques d'une même protéine (ou d'un même complexe protéique). Au delà de cette valeur, ce n'est plus possible. Cependant, les capsides de virus icosaédriques ne sont pas limitées à 60 copies d'une même protéine. Sous certaines conditions (cf. ci-dessous),  $T \times 60$  copies dans  $T$  conformations légèrement différentes peuvent s'assembler en icosaèdre. Pour  $T > 1$ , les protéines dans l'icosaèdre forment alternativement des pentamères et des hexamères par des contacts presque identiques.



**Figure 2.6:** Modèle d'assemblage des protéines en trimère (A), hexamère (B) et pentamère (C) où chaque triangle est constitué de trois sous-unités protéiques. Adapté de [10].

De fait, lorsque l'on s'intéresse aux formes icosaédriques ou complexes, les études cristallographiques montrent que les protéines de la capsidie ont la capacité de

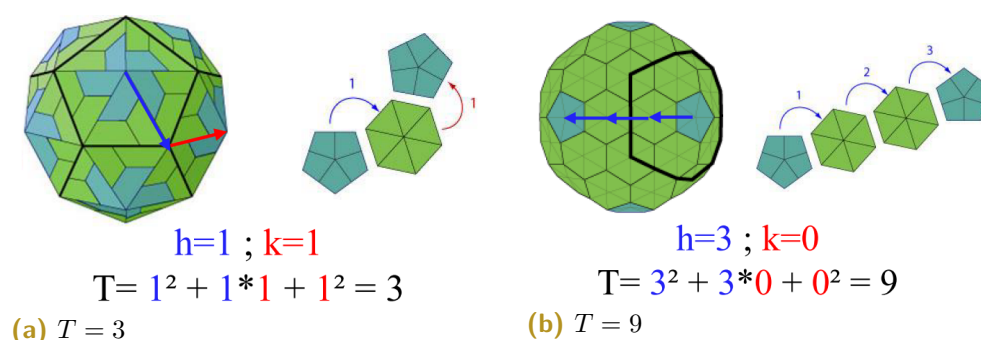
s'assembler en deux types de structures différentes[11][12]. L'assemblage de six unités protéiques est appelé hexamère tandis que l'assemblage de cinq unités protéiques est appelé pentamère. Cette aptitude des protéines s'explique par le principe de quasi-équivalence. Celui-ci, énoncé par Caspar et Klug en 1962, propose, sur la base d'arguments géométriques, que la flexibilité des liaisons au sein de la structure protéique permet aux protéines de s'assembler aussi bien par six que par cinq[10] (Fig. 2.6). Ainsi, l'environnement de la protéine est très similaire dans les deux types d'assemblages possibles (pentamères ou hexamères) et permet d'obtenir des structures composées de plus de 60 protéines.

### 1.2.2.3 Le nombre de triangulation

Il est à noter qu'au delà de 60 protéines, seuls certains multiples de 60 permettent un assemblage lors duquel des protéines de conformations légèrement différentes peuvent se positionner précisément selon la symétrie icosaédrique. Ainsi, si la capside icosaédrique est composée de  $60T$  protéines, il y aura  $T$  conformations de protéines requises. Les valeurs du nombre de triangulation  $T$  pour lesquelles les conformations peuvent être presque identiques sont données par :

$$T = h^2 + h \times k + k^2 \quad (1.1)$$

où  $h$  et  $k$  sont deux nombres entiers dont la somme représente le nombre de pas pour aller d'un pentamère au suivant. De ce fait, le nombre d'hexamères entre deux pentamères vaut  $h + k - 1$  (Fig. 2.7a).



**Figure 2.7:** Exemple de calcul des paramètres  $h$  et  $k$  ainsi que du nombre  $T$  avec un exemple de symétrie icosaédrique de type  $T=3$ (2.7a),  $T=9$  (2.7b). Les flèches bleues montrent le parcours pour  $h$ , les flèches rouges pour  $k$ . Adapté du site ViralZone d'ExPASy (<http://viralzone.expasy.org/806>).

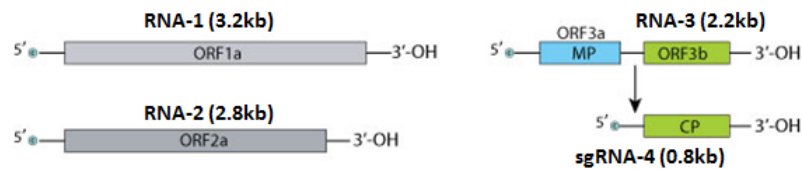
Ainsi, pour un virus icosaédrique de type  $T = 3$  ( $h=1, k=1$ ), la capside est composée de 180 copies d'une même protéine dans 3 conformations légèrement différentes appelées A, formant les pentamères autour des axes d'ordre 5, et B et C, formant les hexamères autour des axes d'ordre 3.

Le virus de la marbrure chlorotique de la cornille (CCMV) est un virus de plante qui infecte un haricot "oeil noir" d'Amérique du sud, la cornille. Le nom de ce virus, appartenant au groupe des bromovirus de la famille des *Bromoviridae*, vient du fait que les feuilles infectées développent des taches jaunâtres. Celui-ci a été isolé et caractérisé pour la première fois, en 1967, par Bancroft et al[13] et fut le premier virus icosaédrique à avoir été réassemblé *in vitro*, à partir de protéines et d'ARN purifiés, pour former des particules infectieuses[14]. De ce fait, celui-ci a fait l'objet d'investigations biochimiques et biophysiques pendant plusieurs décennies. Sa structure cristallographique a notamment été obtenue, à une résolution de 3,2 Å, en 1995[11]. Cette section est dédiée à ce virus particulier puisque c'est l'un des deux systèmes que j'ai étudiés. Je commencerai par présenter le virus, sa protéine de capsid et son génome avant de décrire les protocoles utilisés au sein du laboratoire pour le produire.

## 2.1 Composition et structure du CCMV

### 2.1.1 Le génome

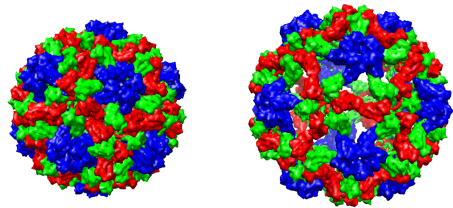
Dans une infection virale classique, chaque particule virale contient la totalité du matériel génétique du virus. Cependant, 17% des virus, comme le CCMV, ont la particularité d'être multipartites[15]. C'est l'un des modes de vie les plus étranges de la virosphère. En effet, ces virus ont des génomes segmentés en morceaux, empaquetés dans différentes capsides et transmis indépendamment. L'ensemble des segments doit donc se rencontrer dans l'hôte afin de poursuivre le cycle viral. Dans le cas du CCMV, son génome est scindé en 4 ARN simple brin de tailles différentes (Fig. 2.8). L'ARN1, l'ARN2 et l'ARN3 comptent 3 171, 2 774 et 2 173 nucléotides (nt), respectivement. L'ARN4 est plus petit et ne compte que 824 nt. Ces 4 segments sont répartis dans 3 particules virales différentes, l'ARN3 et 4 étant empaquetés ensemble. Les trois virions de CCMV contiennent, ainsi, chacun environ 3 000 nt.



**Figure 2.8:** Représentation schématique du génome segmenté du CCMV. Adapté du site ViralZone d'ExpASY (<https://viralzone.expasy.org/134>).

## 2.1.2 La capside

La capside native du CCMV a une structure icosaédrique avec une quasi-symétrie  $T = 3$ . Celle-ci est donc composée de 180 protéines de séquence identique mais ayant trois conformations différentes, classiquement désignées par A, B et C. Les monomères A forment les pentamères de la capside, et les monomères B et C donnent les hexamères (Fig. 2.9). Chaque protéine est composée de 190 acides aminés pour une masse de 20 kDa ce qui donne une masse à la capside de  $180 \times 20 \approx 3,6$  MDa. L'épaisseur de la capside est d'environ 5 nm et, à pH 4,8 le diamètre extérieur de la capside est de 28 nm. Celui-ci augmente à pH 7,5, pour atteindre 32 nm[11].

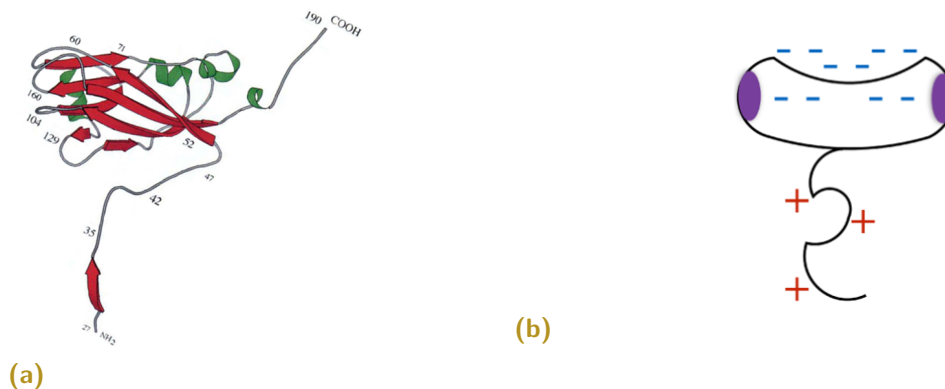


**Figure 2.9:** Capsides de CCMV à pH 4,8 (gauche) et pH 7,5 (droite). Les monomères A, B et C sont bleus, verts et rouges, respectivement. Images de [11].

Sur la figure 2.10a, nous pouvons voir la structure tertiaire de la protéine, où les 25 premiers acides aminés sont absents. Ceci vient du fait qu'ils sont désordonnés et donc invisibles dans la structure cristallographique. La partie N-terminale, i.e. les 25 premiers acides aminés du monomère, est située à l'intérieur de la capside[16]. Elle contient 10 résidus chargés positivement et correspond au domaine d'interaction avec l'ARN[17]. En effet, celui-ci, chargé négativement, va interagir électrostatiquement avec le bras chargé positivement (Fig. 2.10b). Il existe deux autres interactions qui participent à la stabilité de la capside : une attraction à courte portée, due aux forces hydrophobes, qui maintient les dimères ensemble et une répulsion électrostatique à longue distance permettant la dissociation de la capside. Il est à noter que la sous-unité de base de l'assemblage n'est pas le monomère mais le dimère. Ceci vient du fait d'interactions fortes, entre certains



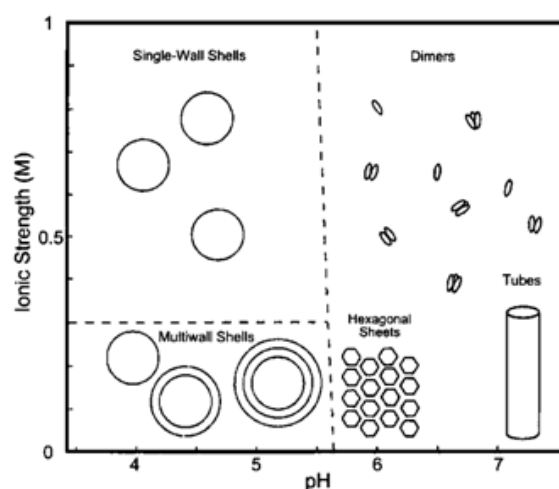
monomères[11] : les monomères A et B interagissent pour former les dimères A-B (au nombre de 60) tandis que les monomères C interagissent entre eux pour donner les dimères C-C (au nombre de 30). Ainsi, lorsque le virus est désassemblé les protéines restent sous forme dimérique.



**Figure 2.10:** Monomères du CCMV. **2.10a** Diagrammes en ruban de la structure tertiaire du monomère de CCMV. Adapté de [11]. **2.10b** Représentation schématique du monomère de CCMV. Les charges négatives des groupes carboxyles, les zones hydrophobes et les charges positives du bras flexible sont représentées en bleu, violet et rouge, respectivement. Adapté de [18].

### 2.1.3 Diagramme de phase de la protéine

Si le CCMV a autant été étudié par les physiciens, c'est pour la remarquable capacité des protéines seules à s'assembler et se désassembler, de manière réversible, au gré des conditions ioniques et de  $pH$ . En effet, comme nous pouvons le voir sur la figure 2.11, en conditions acides avec une forte salinité, les dimères de protéines



**Figure 2.11:** Diagramme de phase des protéines du CCMV. La concentration en protéines est située entre 0,5 et 1.0  $g.L^{-1}$ . La température est de 5°C. Image de [19].

s'assemblent pour former des capsides sphériques. Si la concentration saline est trop faible (<0,2 M de sel monovalent), les protéines vont s'assembler pour former de multiples coques concentriques. Dans le cas où le *pH* est neutre, les dimères peuvent se désassembler (forte salinité) ou former des structures plus exotiques comme des tubes (faible salinité).

## 2.2 Production et purification

Cette section présente les protocoles utilisés dans notre laboratoire afin de produire les plantes qui serviront à exprimer les virus avant d'être purifiés. Ceux-ci ont été adaptés de Ali et Roossinck [54] ainsi que de Aragonés et al. [55].

### 2.2.1 Les solutions tampons usuelles

**Table 2.1:** Tableau récapitulatif des solutions tampons utilisées pour la production des dimères de protéines.

NOM	COMPOSITION
Tampon d'infection	10 mM acétate de sodium <i>pH</i> 4,8 ; 10 mM MgCl <sub>2</sub>
Tampon de purification	0,15 M acétate de sodium <i>pH</i> 4,8
Tampon A	50 mM Tris <i>pH</i> 7,5 ; 0,5 M CaCl <sub>2</sub> ; 1 mM EDTA ; 1 mM DTT ; 0,5 mM PMSF
Tampon B*	50 mM acétate de sodium <i>pH</i> 4,8 ; 0,5 M NaCl ; 1 mM EDTA ; 1 mM DTT ; 0,5 mM PMSF
Tampon B	50 mM Tris <i>pH</i> 7,5 ; 0,5 M NaCl ; 1 mM EDTA ; 1 mM DTT ; 0,5 mM PMSF
Solution de sucrose	2% sucrose

**Table 2.2:** Tableau récapitulatif des fonctions des produits utilisés pour les solutions tampons.

PRODUIT	FONCTION
Tris	maintient le <i>pH</i> (gamme d'action : 6,5 à 9,7)
acétate de sodium	maintient le <i>pH</i> (gamme d'action : 3,5 à 6)
chlorure de sodium	sert à ajuster la force ionique, soit pour créer un choc osmotique afin de déstabiliser la membrane cellulaire, soit pour moduler les interactions entre les protéines.
EDTA	bloque l'activité des nucléases dépendantes des ions Mg <sup>2+</sup> et favorise le gonflement des capsides
DTT	empêche la création de pont disulfure
CaCl <sub>2</sub>	Précipite l'ARN
PMSF	c'est un inhibiteur de protéase

## 2.2.2 Production et purification du virion

## 2.2.3 Plantation et Infection

Le virus est produit à partir de la cornille (*Vigna unguiculata*). La croissance consiste en l'apparition de feuilles primaires puis de feuilles secondaires.

1. 36 pots sont remplis de terre puis, deux graines de cornille sont plantées, par pot, à environ 1 cm de profondeur.
2. Après semence, les plantes ne doivent pas être arrosées directement, sinon les graines risquent de moisir. L'eau est versée dans le réservoir pendant les premiers jours.
3. Les plantes sont arrosées quotidiennement pendant 7 à 10 jours dans une salle, éclairée 16h par jour et thermalisée à 28°C, où des lampes sont positionnées au dessus d'elles (entre 0,5 et 1 m).
4. Environ 7 jours après la plantation (avant la pousse des feuilles secondaires) et 1 jour avant l'infection, la lumière doit être éteinte pour affaiblir les plantes.
5. Le lendemain, une solution d'inoculation est préparée : 2 à 3 feuilles bien infectées sont broyées dans la solution tampon d'infection à l'aide de l'outil de trituration jusqu'à l'obtention d'une solution homogène. Le volume de solution à préparer est à calculer en fonction du nombre de feuilles à infecter<sup>1</sup>.
6. 50  $\mu$ L de la solution obtenue sont déposés sur chaque feuille avant de l'abraser doucement à l'aide du carbure de silicium (SiC) pour faire pénétrer le CCMV. Attention à ne pas déchirer les feuilles.
7. Immédiatement après infection, les pots et les réservoirs sont remplis d'eau.
8. 1 semaine après infection, des taches jaunes doivent apparaître (Fig. 2.12).
9. 10 à 14 jours après infection (avant que le tronc ne devienne violet), les plantes sont récoltées. Seules les feuilles sont gardées, qu'elles présentent ou non des taches jaunes, et conservées à -80°C.

---

1. Par exemple, pour 36 pots avec 2 graines par pot, où chaque graine donne une plante avec 2 feuilles, il y aura  $2 \times 2 \times 36 = 144$  feuilles à infecter, ce qui donne  $144 \times 0,05 \approx 7$  mL à préparer.



**Figure 2.12:** Photo d'une cornille infectée.

## 2.2.4 Purification

Le protocole est écrit pour l'extraction et la purification des virus contenus dans 70g de feuilles infectées. Il est parfaitement possible d'en faire plus, ou moins, en ajustant linéairement les volumes et masses utilisés. L'essentiel des manipulations doit être réalisé à 4°C, donc en chambre froide ou à défaut en gardant les solutions dans un bac à glace.

1. 70 g de feuilles infectées sont broyées, au mixer, en incorporant progressivement 150 mL de tampon de purification, jusqu'à obtenir un mélange homogène.
2. Le mélange obtenu est filtré, à l'aide d'un filtre en nylon, afin de ne récupérer que la phase liquide.
3. Après filtration, 150 mL de chloroforme froid sont ajoutés au filtrat. L'ensemble est recouvert et laissé à agiter en chambre froide pendant 1 h.
4. Après agitation, la solution est transvasée dans des bouteilles de centrifugation de contenance 250 mL et centrifugées à 10 000 g à 4°C pendant 10 min (rotor GSA).
5. Après centrifugation, le surnageant est délicatement récupéré et son volume est mesuré. 8 g de polyéthylène glycol (PEG) MW 8000 et 115 mg de NaCl sont pesés pour 100 mL de solution. Le PEG et le NaCl sont incorporés lentement et sous agitation pour éviter la formation de caillot insoluble. L'ensemble est laissé à agiter 1 h en chambre froide.

6. Après agitation, la solution est transvasée dans des bouteilles de centrifugation de contenance 250 mL et centrifugée à 10 000 g à 4°C pendant 10 min (rotor GSA).
7. Après centrifugation, le surnageant est rapidement éliminé afin d'éviter que le culot ne se resolubilise. Ce dernier se caractérise par une trace blanc-jaune sur le bord du contenant. Les culots sont ensuite solubilisés avec 21 mL d'une solution de 0,05 M d'acétate de sodium pH 4,8. La solution doit devenir trouble. L'ensemble est transvasé dans un seul récipient puis agité pendant 1 h en chambre froide.
8. Après agitation, la solution est transvasée dans des bouteilles de centrifugation de contenance 30 mL et centrifugée à 9 000 g à 4°C pendant 10 min (rotor SS34).
9. Après centrifugation, le surnageant est rapidement récupéré puis réparti dans 3 tubes d'ultracentrifugation. Les tubes sont complétés avec du tampon de purification. Dans le fond de chaque tube, à l'aide d'une seringue, insérer délicatement 2 mL de solution de sucrose. Les deux phases doivent être clairement distinctes. Après équilibration, l'ensemble est centrifugé à 150 000 g à 4°C pendant 2 h (rotor SW41 Ti).
10. Après centrifugation, le surnageant est rapidement éliminé puis, le culot est solubilisé avec 1 mL d'une solution de 0,05 M d'acétate de sodium pH 4,8.
11. La solution obtenue est prélevée et aliquotée dans des tubes de 1,5 mL.
12. L'absorbance de la solution est mesurée au Nanodrop<sup>2</sup> puis la concentration en capsides est calculée<sup>3</sup> et la pureté contrôlée, i.e. le rapport  $A_{260}/A_{280}$  doit se situer entre 1,5 et 1,7[20].

En général, à partir de 70 g de feuilles infectées, 5 à 10 mg de virions sont typiquement obtenus à la fin de la purification.

---

2. Spectrophotomètre UV-Visible de la marque *Thermo Scientific* qui mesure l'absorbance d'un échantillon pour une longueur d'onde donnée en suivant la loi de Beer-Lambert.

3. Usuellement, la concentration des protéines est calculée à partir de l'absorbance à 280 nm. Cependant, dans le cas des virus de CCMV, la présence de l'ARN modifie le calcul. De ce fait, l'application "nucleic acid" du Nanodrop est utilisée. La concentration est déterminée en divisant l'absorbance à 260 par 5,8.

## 2.3 Purification des protéines

Une fois le virus purifié, les acides nucléiques peuvent être retirés du virus. Le protocole est écrit pour la purification des protéines de la capsid du CCMV. L'essentiel des manipulations doit être réalisé à 4°C, donc en chambre froide ou à défaut en gardant les solutions dans un bac à glace.

1. 1 mL d'une solution de virus est déposé dans une membrane de dialyse MWCO 6 kDa. L'échantillon est complété avec de la solution tampon A, jusqu'à  $\approx 1$  cm du bord. L'ensemble est mis à dialyser dans 500 mL de tampon A, fraîchement préparé, puis agité à 4°C pendant 12 h.
2. Après incubation, la solution tampon est renouvelée puis agitée à 4°C pendant 12 h.
3. Après incubation, la solution est récupérée et transférée dans un tube d'ultra-centrifugation neuf et propre qui doit être complété jusqu'à  $\approx 1$  cm du bord avec du tampon A. Après équilibration, l'ensemble est centrifugé à 150 000 g à 4°C pendant 18 h (rotor SW41 Ti). Quelques mL de tampon A dialysé sont réservés pour l'étape 5.
4. Après centrifugation, le surnageant est rapidement prélevé, mL par mL, dans des tubes séparés, en prenant garde à bien prendre toujours la partie la plus haute de la solution. Il est important d'être le plus rapide possible car les capsides vont se réassembler autour de l'ARN. Inutile de préciser qu'il est fortement déconseillé de secouer la solution, même un peu.
5. La teneur en ARN de chaque aliquot est vérifiée au Nanodrop : le ratio  $A_{280}/A_{260}$  doit être supérieur à 1,5. Les échantillons ne respectant pas cette valeur peuvent être jetés.
6. Dans le cas où l'on souhaite obtenir des dimères libres, les échantillons restant sont placés dans une membrane de dialyse MWCO 6 kDa. L'échantillon est complété avec de la solution tampon B, jusqu'à  $\approx 1$  cm du bord. L'ensemble est mis à dialyser dans 500mL de tampon B, fraîchement préparé, puis agité à 4°C pendant 12 h.

Dans le cas où l'on souhaite obtenir des capsides vides, les échantillons restant sont placés dans des tubes de centrifugation Amicon Ultra 4 mL 3 kDa. L'échantillon est complété avec de la solution tampon B\*. L'ensemble

est centrifugé à 4 000 g à 4°C pendant 40 min (rotor A-4-44). L'opération est répétée 3 à 4 fois pour totalement changer de solution tampon.

7. La solution est récupérée puis la concentration en protéines est calculée via l'absorbance à 280 nm<sup>4</sup> et la pureté est contrôlée<sup>5</sup>.

En général, à partir de 8 mg de virus, 3 à 4 mg de protéines de capsid du CCMV sont obtenus.

---

4. Dans ce cas, il faut utiliser l'absorbance à 280, en n'oubliant pas de fixer le coefficient à 8,34.

5. Le rapport  $A_{260}/A_{280}$  doit être inférieur à 0,65[21].

# Le virus de l'hépatite B

La première épidémie, enregistrée, d'une hépatite B a été observée par Lurman, en 1885[22]. Le virus sera découvert en 1963 par B. Blumberg[23], les particules virales seront observées en 1970 au microscope électronique[24], le génome du virus sera séquencé dès 1979[25] et les premiers vaccins sont expérimentés en 1980. Et pourtant, l'infection par le virus de l'hépatite B (VHB) demeure une maladie qui représente un fardeau sanitaire mondial important. L'Agence nationale de recherches sur le sida et les hépatites virales estime, aujourd'hui, qu'environ 240 millions de personnes sont chroniquement infectées par le VHB, avec, par conséquent le risque de développer une cirrhose, une décompensation hépatique et un carcinome hépatocellulaire.

Si le VHB présente un enjeu sanitaire majeur, ce sont ses propriétés d'auto-assemblage et de désassemblage *in vitro*[26] qui ont éveillé la curiosité des biophysiciens à son sujet[27][28]. Ce virus étant le deuxième système que j'ai étudié, ce chapitre lui sera consacré. Je présenterai d'abord un rapide résumé sur l'infection virale dû au virus puis le génome du VHB et son cycle viral. Je terminerai en décrivant la capsid du VHB ainsi que sa protéine.

## 3.1 Epidémiologie

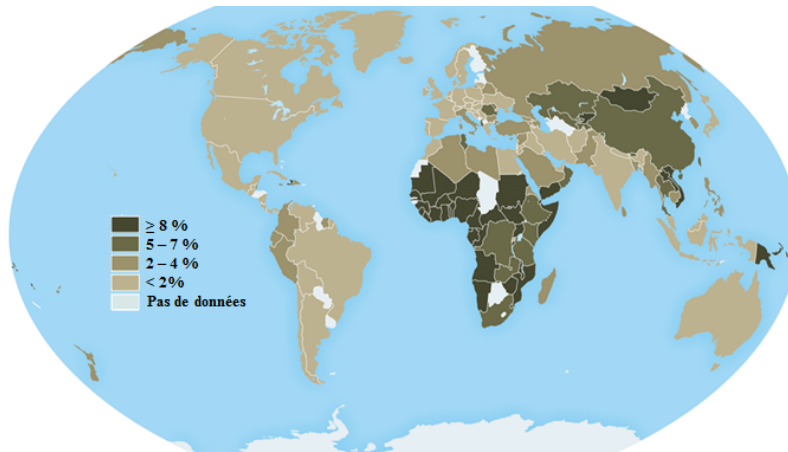
### 3.1.1 Prévalence

L'hépatite B est une infection du foie causée par le virus de l'hépatite B (VHB). Celle-ci est potentiellement mortelle. En effet, les malades chroniques ont un risque élevé de décéder des suites d'une cirrhose ou d'un cancer du foie[29]. Malgré la présence de vaccins efficaces, l'infection par le VHB reste un problème planétaire avec une incidence et une prévalence inversement proportionnelle au développement économique. En effet, nous pouvons voir sur la figure 2.13, trois grandes zones distinctes :

- une zone de faible endémie où la prévalence est inférieure à 2% (Australie, Amérique du nord ainsi que Europe de l'ouest et du nord)



- une zone médiane prévalence où la prévalence est comprise entre 2 et 8% (Europe de l'est, Russie, Amérique du sud, le proche orient ainsi que les pays méditerranéens)
- une zone de forte endémie où la prévalence est supérieure à 8% (Asie du sud-est, Chine et Afrique sub-saharienne). Pour certain pays d'Afrique, celle-ci peut même atteindre 20%[30].



**Figure 2.13:** Prévalence du virus de l'hépatite B dans le monde. Adapté du site du Département de la Santé et des Services Sociaux des États-Unis (<https://www.cdc.gov/hepatitis/hbv/>).

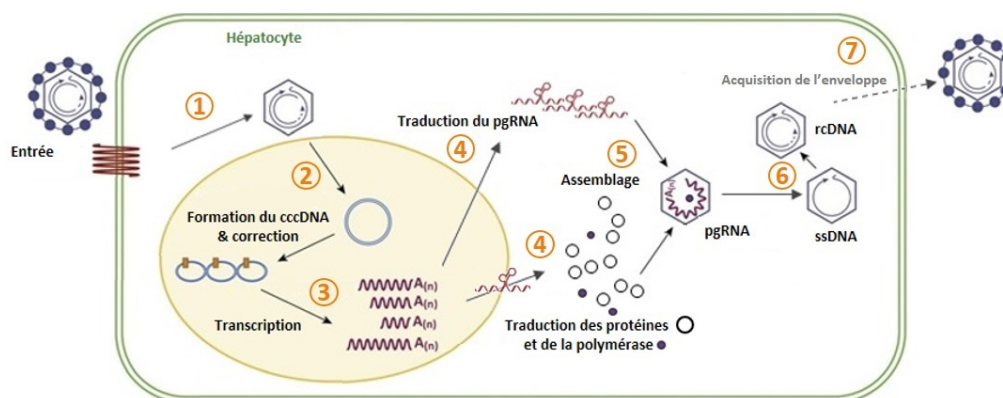
### 3.1.2 Transmission

Le virus de l'hépatite B est présent dans les fluides organiques, i.e. le sang, la cyprine, le lait maternel, etc. La transmission entre les êtres humains s'effectue de façon identique à celle du virus de l'immunodéficience humaine (VIH), i.e. via l'exposition au sang infectieux ou à des liquides organiques contenant du sang. Le virus peut survivre à l'extérieur du corps humain pendant 7 jours puis infecter une personne non vaccinée pour se répliquer[31].

## 3.2 Cycle viral

Comme nous avons pu le voir dans la section 1.2.1, les virus appartenant au groupe VII de la classification de Baltimore ont un mécanisme de réplication complexe. C'est particulièrement le cas du virus de l'hépatite B, dont une partie du cycle est encore méconnue. De manière très simplifiée, la réplication du génome du VHB peut être divisée en plusieurs phases présentées sur la figure 2.14 [32] :

1. Les virions infectieux contiennent un génome d'ADN partiellement double brin, circulaire relâché<sup>1</sup> appelé ADN-rc.
2. Après l'entrée du génome viral dans le noyau, l'ADN-rc est converti en un ADN circulaire covalentement clos appelé ADN-ccc.
3. L'ADN-ccc sert de matrice pour la transcription des ARN viraux, sous forme d'un mini-chromosome viral.
4. Les ARN transcrits sont ensuite traduits en plusieurs protéines. L'un d'eux est à la fois l'ARN messager pour la protéine de capsid et la polymérase, et l'ARN qui sert de matrice à la transcription inverse par la polymérase. C'est pourquoi il est appelé ARN pré-génomique (ARN-pg).
5. La protéine de capsid empaquète l'ARN-pg avec la polymérase dans les nucléo-capsides.
6. L'ARN-pg sert, dans la nucléocapside, de matrice à la transcription inverse en ADNrc.
7. Les nucléo-capsides matures sont ensuite dirigées vers une enveloppe via le réticulum endoplasmique (RE) où elles seront enveloppées et pourront être sécrétées pour infecter d'autres cellules.



**Figure 2.14:** Illustration schématique du cycle de vie du VHB. Adapté de [33]

### 3.3 Composition et structure du VHB

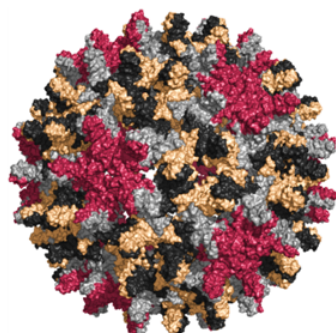
La démonstration de la production de protéines du virus de l'hépatite B en bactérie (*Escherichia coli*) a été apportée en 1979[34]. En 1988, Nassal et al. ont pu produire de cette façon, sans risques et sans nécessités d'installations sécurisées, la protéine

1. C'est un ADN double brin circulaire dont les deux brins sont ouverts et un brin est incomplet.

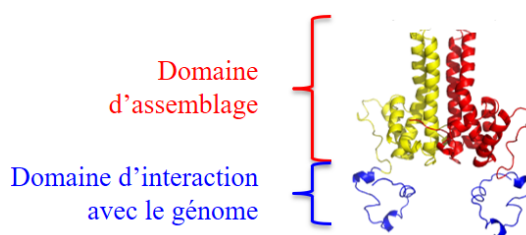
de capsid, qui s'auto-assemble spontanément en pseudo-nucléocapsides. Rapidement, les premières études sur l'auto-assemblage de la capsid ont suivi[35].

### 3.3.1 La protéine native et sa capsid

Comme la capsid native du VHB, les capsides produites en *E. coli* ont une structure icosaédrique légèrement polydisperse. En effet, les capsides ont majoritairement une quasi-symétrie  $T = 4$  ( $\approx 95\%$ ) accompagnées de quelques  $T = 3$ . Celle-ci est donc composée de 240 protéines de séquence identique mais ayant quatre conformations différentes, classiquement désignées par A, B, C et D (Fig. 2.15). Chaque protéine a une masse d'environ 21 kDa ce qui donne à la capsid une masse d'environ 5 MDa. L'épaisseur de la capsid est d'environ 6 nm et le diamètre extérieur de la capsid est de 34 nm.



**Figure 2.15:** Capsides du VHB. Les monomères A, B, C et D sont en bordeaux, gris, noir et beige, respectivement. (PDB ID : 1QGT[36])



**Figure 2.16:** Structure d'un dimère de VHB avec une coloration par monomère (jaune et rouge). Le bras flexible est coloré en bleu.

Sur la figure 2.16, nous pouvons voir la structure tertiaire du dimère de protéine<sup>2</sup>. Un monomère natif de la capsid du VHB est composé de 183 résidus comprenant un domaine d'assemblage (N-terminal, résidus 1-149) et un domaine de liaison à l'acide nucléique (C-terminal, résidus 150-183). Le domaine N-terminal est nécessaire et suffisant à l'auto-assemblage des 120 sous-unités en capsides icosaédriques où, comme pour le CCMV, la sous-unité de base correspond à un dimère de protéine[37]. Le domaine C-terminal est nécessaire pour l'interaction avec l'ARN et l'encapsidation du génome dans la capsid. Ce domaine est riche en

---

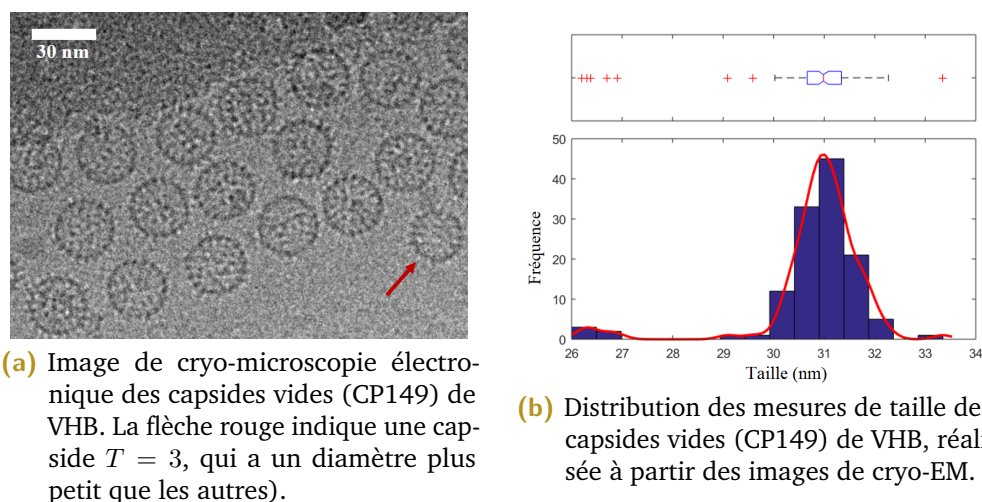
2. PDB ID : 1QGT

arginine et contient de nombreux sites de phosphorylation, portés par les sérines et thréonines[38], qui peuvent moduler les propriétés de la capside ainsi que les interactions avec le génome. Comme pour le CCMV, l'interaction génome-protéine est pilotée par l'attraction électrostatique. La répulsion électrostatique et l'attraction hydrophobe modulent les interactions entre les dimères.

### 3.3.2 Dispersité de la protéine tronquée

Comme indiqué précédemment, il a été démontré que les capsides peuvent s'assembler *in vitro* en l'absence du génome [35]. Pour l'étude de phénomènes complexes, comme l'assemblage des capsides de virus, une première approche consiste à ne considérer que les protéines et leurs interactions. C'est ce que nous avons fait à l'aide de la protéine tronquée, notée CP149 (composée uniquement du domaine d'assemblage). La monodispersité de nos échantillons est une condition requise dans le cadre de notre étude de l'auto-assemblage des capsides de virus. De ce fait, nous avons vérifié que les assemblages des protéines CP149 respectaient bien cette condition.

Nos résultats, obtenus par cryo-microscopie électronique (cryo-EM<sup>3</sup>), sont présentés sur les figures 2.17b et 2.17a. Nous pouvons observer que les capsides vides du VHB sont de taille relativement mono-disperse. En effet, nous pouvons observer deux populations dont une très majoritaire ( $\approx 95\%$ ), ce qui est cohérent avec la littérature[39].



**Figure 2.17:** Analyse de la dispersité en taille des capsides vides (CP149) du VHB

3. En anglais, cryogenic electron microscopy

Les virus ont généralement une connotation négative puisqu'ils évoquent une maladie. Cette remarque est particulièrement vraie avec les récentes épidémies d'Ebola, du Chikungunya ou du SARS-CoV-2 qui se sont propagées assez rapidement et se sont révélées difficiles à contrôler. Cependant, les virus ont également des propriétés ayant des applicabilités positives. En effet, depuis quelques années de nombreux progrès ont été réalisés en médecine ou dans le domaine des biotechnologies grâce à l'utilisation de virus non pathogènes ou de nanomatériaux à base de virus.

## 4.1 Intérêt médical

### 4.1.1 Transfert génétique

Les virus peuvent être utilisés pour transférer des ARN interférents<sup>1</sup> (ARNi), de l'ADN ainsi que des outils pour l'édition génomique, tels que les nucléases. Le transfert génétique permet le traitement d'un large éventail de troubles depuis sa conceptualisation, il y a plus de 40 ans.

Par exemple, certains virus possèdent des fibres leur permettant de se lier à des récepteurs spécifiques de leurs cellules hôtes. Dans le cas des adénovirus, ceux-ci ont une spécificité pour les cellules de mammifères qui peut être utilisée pour la thérapie génique, dans des cas tels que le glioblastome<sup>2</sup>[40].

Le transfert génétique peut également être utilisé comme thérapie antivirale à l'aide d'ARNi. Cela a été démontré pour le VHB par administration de micro-ARN<sup>3</sup> anti-VHB associé à un promoteur spécifique. Dans ce cas, le promoteur utilisé était un élément régulateur de la transcription spécifique au foie qui lui conférait

---

1. ARN simple ou double brin dont l'interférence avec un ARN messenger spécifique conduit à la dégradation dudit ARNm et à la diminution de sa traduction en protéine.

2. Cancer du cerveau très agressif.

3. ARN régulateurs traductionnels capables de l'extinction ou de l'expression d'un gène.

une spécificité élevée. De plus, l'étude montrait que l'expression du micro-ARN anti-VHB avait entraîné une suppression de la réplication du VHB jusqu'à 94% chez la souris[41].

Les virus de plantes sont également à l'étude pour le transfert génétique. Par exemple, les protéines de la capsid du CCMV ont pu être assemblées autour d'un ARN dérivé du virus Sindbis (SINV). Associé à de la Lipofectamine<sup>4</sup>, l'ensemble a pu être libéré dans le cytoplasme de cellules de mammifères[42].

### 4.1.2 Vecteurs de médicaments

En plus du transfert génétique, les virus peuvent également être utilisés comme vecteurs de médicaments pour la thérapie photothermique, la thérapie photodynamique (PDT) et la chimiothérapie.

La thérapie photodynamique est une technique qui vise à détruire des tissus par l'utilisation de substances chimiques photo-sensibilisantes qui ont notamment la propriété de devenir toxiques à la lumière. Cette technique a pu être utilisée comme approche anti-microbienne en utilisant des capsides de CCMV fonctionnalisées avec un photo-sensibilisateur à base de ruthénium et dirigée vers *Staphylococcus aureus*[43]. Afin d'obtenir une meilleure pénétration tissulaire, des colorants à base de phtalocyanines ont également pu être encapsulés[44].

## 4.2 Biotechnologie

Dans le domaine de la biotechnologie, les virus peuvent également avoir des applications variées.

### 4.2.1 Nano-réacteurs

L'intérêt pour la fonctionnalisation des particules virales avec des enzymes s'est accru à la suite d'une étude montrant le potentiel de génération de nanoparticules catalytiquement actives qui peuvent s'auto-assembler et se propager facilement[45]. Ainsi, l'encapsulation d'enzymes pour former des nano-réacteurs est devenue de plus en plus pertinente pour la biotechnologie.

---

4. Réactif de transfection utilisé en biologie moléculaire et cellulaire.

Par exemple, de la peroxydase de raifort<sup>5</sup> (HRP) a été encapsulée à l'aide des protéines de la capsid du CCMV. Environ une capsid sur 130 contenait l'enzyme HRP permettant ainsi l'étude de l'enzyme au niveau d'une seule molécule[46].

## 4.2.2 Matériaux nano-structurés

Le CCMV a pu être utilisé pour emballer divers composés organiques ou inorganiques et, notamment, des nanoparticules d'or[47][48]. Leur emballage est particulièrement étudié pour plusieurs raisons : l'or est bio-compatible et convertit efficacement la lumière en chaleur ce qui permet l'ablation thermique très spécifique de tissus malades ou infectés. De plus, les nanoparticules d'or peuvent absorber de grandes quantités de rayons X et donc être utilisées pour améliorer la radiothérapie contre le cancer, etc[49]. Comme indiqué dans l'introduction, le CCMV peut également être envisagé comme substrat pour la nano-chimie ou comme modèle pour la nano-fabrication. Il a été utilisé, par exemple, pour synthétiser et assembler des nanofils d'oxyde de cobalt à température ambiante[50] ou encore pour générer des nanoparticules de sels de polyoxométalate[51].

---

5. Oxydoréductase. En anglais, horseradish peroxidase

## Diffusion de rayons X aux petits angles

La diffusion des rayons X aux petits angles (SAXS, Small-Angle X-ray Scattering en anglais) est une technique permettant l'étude de molécules biologiques. Celle-ci permet de collecter les données structurales de protéines couvrant une large plage de poids moléculaires avec une résolution spatiale de l'ordre du nanomètre[52]. Etant donné que les données de diffusion aux petits angles (SAS, Small-Angle Scattering en anglais) peuvent être obtenues à partir d'objets biologiques en suspension dans une solution, les mesures permettent de refléter la structure globale de ces derniers en fonction de la nature de la solution (pH, salinité, etc) et les conditions peuvent être facilement modifiées pour refléter les changements physiologiques. De plus, l'absence de restrictions de taille, de nécessité de cristallisation et la rapidité de la collecte de données en font une technique puissante pour la caractérisation de systèmes macromoléculaires et nano-structurés.

Cette partie du manuscrit a pour objectif de décrire simplement les principes physiques de base de la diffusion aux petits angles ainsi que la technique de diffusion des rayons X aux petits angles résolue dans le temps (TR-SAXS, Time-Resolved Small-Angle X-ray Scattering en anglais) utilisée au cours de cette thèse.



# Principe de base de la diffusion aux petits angles

# 1

Ce chapitre couvre les principes de base des études SAXS. Celui-ci n'a pas vocation à être exhaustif puisque de nombreux ouvrages présentent déjà, en détail et de manière très didactique, les techniques de diffusion[53][54][55].

## 1.1 Interaction des rayons X avec les échantillons

Tout comme la lumière visible les rayons X, découverts en 1895 par Wilhelm Röntgen, sont des ondes électromagnétiques. La longueur d'onde des photons de rayons X peut être déterminée par la relation de Planck-Einstein :

$$\lambda = \frac{hc}{E} \quad (1.1)$$

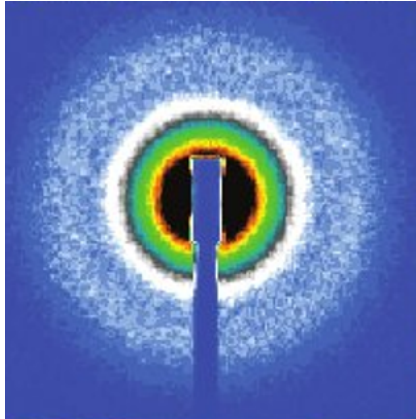
avec  $h$  la constante de Planck,  $c$  la vitesse de la lumière dans le vide et  $E$  l'énergie. Pour les études structurales, les photons de rayons X utilisés ont une énergie  $E$  d'environ 10 keV. Leur longueur d'onde associée se situe donc autour de 0,1 nm (alors que celle de la lumière visible se situe autour de 500 nm). Cette longueur d'onde est du même ordre de grandeur que les distances inter-atomiques permettant ainsi l'apparition d'un motif de diffusion.

En effet, dans le cas d'une diffusion élastique, lorsque les photons de rayons X entrent en collision avec les nuages électroniques des atomes, ceux-ci émettent un rayonnement de même fréquence que le rayonnement incident. L'ensemble des atomes produisent alors des ondes cohérentes pouvant interférer. Les interférences peuvent être constructives, destructives ou entre les deux. La projection du résultat sur un plan correspond à un motif d'interférence 2D caractéristique de la répartition spatiale des électrons au sein du matériau (Fig. 3.1). Mathématiquement, ce motif d'amplitude des rayons X diffusés par les électrons correspond à la transformée de Fourier de la distribution de la densité électronique.

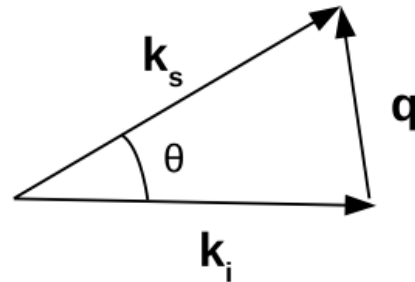
Chaque distance peut être mesurée par rapport à la longueur d'onde  $\lambda$  du rayonnement incident et les motifs de diffusion sont présentés comme des fonctions du vecteur de diffusion  $\vec{q}$ . Celui-ci est défini comme la différence entre les vecteurs de propagation des signaux incidents  $\vec{k}_i$  et diffusés  $\vec{k}_s$  (Fig. 3.2) et a pour norme :

$$q = \frac{4\pi}{\lambda} \sin \frac{\theta}{2} \quad (1.2)$$

avec  $\lambda$  la longueur d'onde du rayon incident et  $\theta$  l'angle de diffusion.



**Figure 3.1:** Motif de diffusion 2D (échantillon isotrope).



**Figure 3.2:** Vecteur de diffusion  $\vec{q} = \vec{k}_s - \vec{k}_i$

Pour accroître le rapport signal sur bruit, l'intensité est moyennée pour chaque valeur de  $\theta$  (intégration radiale). On obtient ainsi des courbes qui représentent les variations d'intensité  $I$  en fonction du vecteur de diffusion  $\vec{q}$ .

Afin de pouvoir obtenir des informations quantitatives sur les espèces en solutions, i.e. les poids moléculaires, les volumes spécifiques, etc, il est nécessaire d'avoir les intensités en unités absolues ( $\text{cm}^{-1}$ ). L'intensité est calibrée à l'aide d'un échantillon dont la diffusion en unités d'intensité absolue est connue. Cette mesure permet de déterminer un facteur d'étalonnage  $CF$  qui peut être déterminé à l'aide de l'équation ([56]) :

$$CF = \left( \frac{d\Sigma}{d\Omega} \right)_{st}(q) \frac{d_{st} T_{st}}{I_{st}(q)} \quad \text{avec} \quad \frac{d\Sigma}{d\Omega} = 0.01633 \text{ cm}^{-1} \quad \text{pour l'eau} \quad (1.3)$$

où

- $\left( \frac{d\Sigma}{d\Omega} \right)_{st}(q)$  est l'intensité de diffusion connue de l'échantillon d'étalonnage en  $\text{cm}^{-1}$ ,

- $d$  est l'épaisseur de l'échantillon,
- $T$  est le facteur de transmission,
- $I$  est l'intensité mesurée,
- $st$  est l'indice désignant la calibration standard.

Pour obtenir les données d'intensité en unités absolues et après avoir effectué les corrections nécessaires (transmission, fond, etc), il suffit de multiplier l'intensité mesurée de l'échantillon par le facteur d'étalonnage  $CF$ .

## 1.2 Intensité et contraste

### 1.2.1 Cas d'une molécule seule

Il est important de comprendre que le signal obtenu est diffusé par l'ensemble des atomes de la solution, i.e. par les molécules d'intérêt et par le solvant. Les expériences de diffusion impliquent donc d'effectuer des mesures séparées des solvants seuls afin de pouvoir soustraire leur contribution et d'étudier uniquement le signal des molécules d'intérêt. Cette soustraction permet également de réduire considérablement la contribution du bruit de fond parasite (fentes, porte-échantillons, etc.).

Afin de comprendre l'importance de ces deux mesures séparées, il est utile d'introduire la notion de densité de longueur de diffusion  $\rho(\vec{r})$ . Cette mesure représente le pouvoir diffusant d'une molécule tel que :

$$\rho(\vec{r}) = \frac{b(\vec{r})}{v} \quad (1.4)$$

où  $b(\vec{r})$  correspond à la longueur de diffusion d'une molécule, de volume moléculaire  $v$ . Nous pouvons alors définir le contraste de densité de longueur de diffusion,  $\Delta\rho$ , lorsque des particules de densité de longueur de diffusion  $\rho(\vec{r})$  sont dans une solution de densité de longueur de diffusion  $\rho_s$ , tel que :

$$\Delta\rho = \rho(\vec{r}) - \rho_s \quad (1.5)$$

Celui-ci permet de déterminer l'amplitude de diffusion  $A_{part}(\vec{q})$  qui est définie comme sa transformée de Fourier :

$$A_{part}(\vec{q}) = \int_{V_{part}} \Delta\rho(\vec{r}) e^{i\vec{q}\cdot\vec{r}} d\vec{r} \quad (1.6)$$

où l'intégration est faite sur le volume de la particule,  $V_{part}$ . De plus, le détecteur mesurant le carré de l'amplitude de l'onde, l'intensité diffusée d'une particule, pour une orientation donnée, s'écrit :

$$I_{part}(\vec{q}) = A_{part}(\vec{q}) A_{part}^*(\vec{q}) = V_{part}^2 P_{part}(\vec{q}) \quad (1.7)$$

où  $P_{part}(\vec{q})$ , qui sera détaillé dans la section 1.4, correspond au facteur de forme de la particule. Celui-ci décrit la morphologie de la particule et s'écrit :

$$P_{part}(\vec{q}) = \frac{1}{V_{part}^2} \int_{V_{part}} \int_{V_{part}} \Delta\rho(\vec{r}) \Delta\rho(\vec{r}') e^{i\vec{q}\cdot(\vec{r}-\vec{r}')} d\vec{r} d\vec{r}' \quad (1.8)$$

## 1.2.2 Cas de N molécules identiques

Ces expressions peuvent se généraliser. Ainsi, dans le cas d'une solution diluée, contenant  $N$  molécules identiques, dont les positions ne sont pas corrélées, l'intensité diffusée par chaque molécule est indépendante des autres. Autrement dit, ces intensités diffusées peuvent être sommées et l'intensité totale par unité de volume  $V$  est égale à :

$$I(\vec{q}) = \frac{N}{V} V_{part}^2 P_{part}(\vec{q}) \quad (1.9)$$

Dans le cas où la solution n'est pas diluée, et les positions des molécules corrélées, l'intensité diffusée provient de deux contributions. La première correspond aux interférences entre noyaux de diffusion situés au sein de la même particule. La seconde contribution correspond aux interférences entre noyaux de diffusion appartenant à deux particules différentes. Autrement dit, les particules diffusent en solution en fonction du temps, ce qui modifie l'intensité diffusée à chaque instant, puisque celle-ci dépend maintenant aussi de leur organisation relative. En

considérant que la probabilité que le système soit dans un état donné soit invariante au cours du temps, alors l'intensité peut être écrite sous la forme suivante :

$$I(\vec{q}) = \frac{1}{V} \left\langle \left\{ \int_V \Delta\rho(\vec{r}) e^{i\vec{q}\cdot\vec{r}} d\vec{r} \right\} \left\{ \int_V \Delta\rho(\vec{r}') e^{i\vec{q}\cdot\vec{r}'} d\vec{r}' \right\} \right\rangle_{\Omega} \quad (1.10)$$

En effectuant un changement de variable, i.e.  $\vec{r} = \vec{r}_i + \vec{u}$  et  $\vec{r}' = \vec{r}_j + \vec{v}$ , cette équation mathématique peut être réécrite pour dépendre des centres de masse,  $\vec{r}_i$  et  $\vec{r}_j$ , des particules. De plus, dans le cas de  $N$  particules homogènes, si celles-ci sont sphériques et le système isotrope, alors l'intensité diffusée peut être simplifiée pour ne dépendre que de l'amplitude de  $\vec{q}$ . Autrement dit, nous pouvons faire la moyenne des orientations et l'équation 1.10 peut alors s'écrire[57] :

$$I(q) = \frac{N}{V} \Delta\rho^2 V_{part}^2 P(q) S(q) \quad (1.11)$$

où  $P(q)$  et  $S(q)$  sont définis comme le facteur de forme des particules et le facteur de structure, respectivement.

## 1.3 Facteur de structure

Lorsque les échantillons sont concentrés ou lorsqu'il y a de l'agrégation, le motif d'interférence contiendra également des informations sur les interactions des particules entre elles. Ce motif d'interférence, appelé « facteur de structure », est une représentation de la probabilité qu'une particule particulière soit entourée par une autre particule. Pour les systèmes isotropes, deux types de facteurs de structure peuvent être distingués :

- Ceux de systèmes dits corrélés qui décrivent les interactions des particules les unes avec les autres,
- Ceux de systèmes dits polymères pouvant être décrits en terme de dimension fractale.

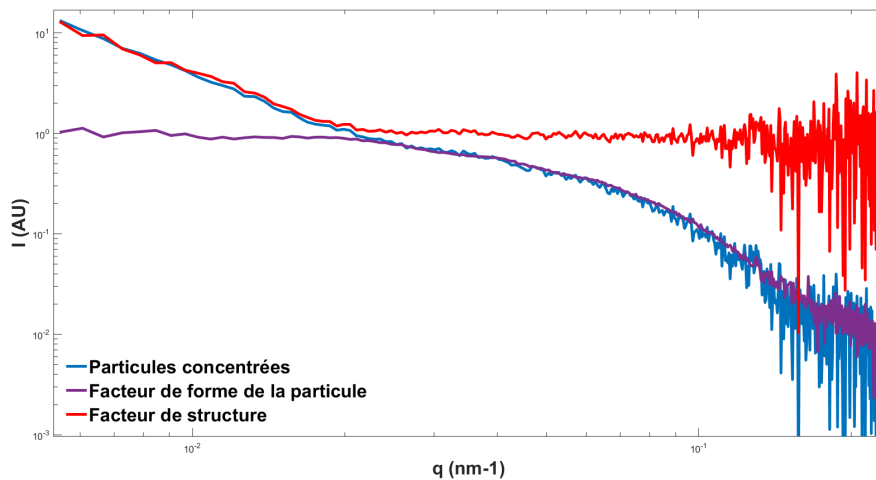
Concernant ces derniers, les petites particules peuvent s'agréger de manière aléatoire formant ainsi des structures ramifiées. Lorsque plusieurs particules de rayon  $r_0$  sont assemblées, le nombre de particules au sein de l'agrégat de rayon  $r$  est donné par l'équation suivante[58] :

$$N(r) \propto \left(\frac{r}{r_0}\right)^{D_f} \quad (1.12)$$

où  $D_f$  est une constante dont la dimension fractale reflète la compaction de l'agrégat.

Cette description n'a de sens que lorsque les particules sont compactées de sorte que le système est auto-similaire sur au moins une échelle spatiale. Le système peut être alors qualifié de fractal.

Une fois l'expression du facteur de forme obtenue (voir section suivante), le facteur de structure peut être calculé en divisant l'intensité mesurée par le facteur de forme (Fig. 3.3). Par convention, le facteur de structure est pondéré par le carré du volume des particules agrégées, de telle sorte que  $S(q \rightarrow \infty) = 1$ . Sur la figure 3.3, l'augmentation aux petites valeurs de  $q$  indique une interaction attractive, typique d'un phénomène d'agrégation.



**Figure 3.3:** Profil SAXS d'une solution concentrée en protéines de la capsid du CCMV (en bleu). C'est le produit du facteur de forme de la particule unique (en violet) et du facteur de structure (en rouge).

Par la suite, nous ne considérons que le cas d'une solution diluée, i.e. sans interactions, de  $N$  particules identiques, de sorte que le facteur de structure soit négligeable et que l'équation suivante puisse être appliquée :

$$I(q) = \frac{N}{V} \Delta\rho^2 V_{part}^2 P(q) \quad (1.13)$$

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

## 1.4 Facteur de forme et fonction de distribution de distance

Lorsqu'une particule est soumise à un faisceau de rayons X, un motif d'interférence est produit par l'ensemble des ondes envoyées au détecteur à partir du nuage électronique de la particule. Ce motif, caractéristique de l'organisation et de la forme de la particule, est appelé "facteur de forme".

Les équations 1.7, présentées précédemment, peuvent être moyennées sur toutes les orientations possibles des particules en solution. L'intégration en coordonnées sphériques de la transformée de Fourier donne alors :

$$V_{part}^2 P(q) = 4\pi \int_0^{D_{max}} p(r) \frac{\sin(qr)}{qr} dr \quad (1.14)$$

où  $p(r) = r^2 \gamma(r)$  est la fonction de distribution de distance (voir plus bas) avec  $\gamma(r)$  la fonction d'autocorrélation moyenne du contraste de densité de diffusion et  $D_{max}$  la distance maximale au sein de la particule.

Cette description succincte n'est appropriée que si la norme du vecteur d'onde  $\vec{q}$  est suffisamment petite devant l'inverse de  $1/\delta$  où  $\delta$  est une dimension structurale caractéristique des constituants du système étudié. Autrement dit,  $q\delta \ll 1$ .

Concernant ce dernier point, l'intensité moyenne normalisée de  $N$  particules homogènes identiques en solution exprimée en unités absolues est :

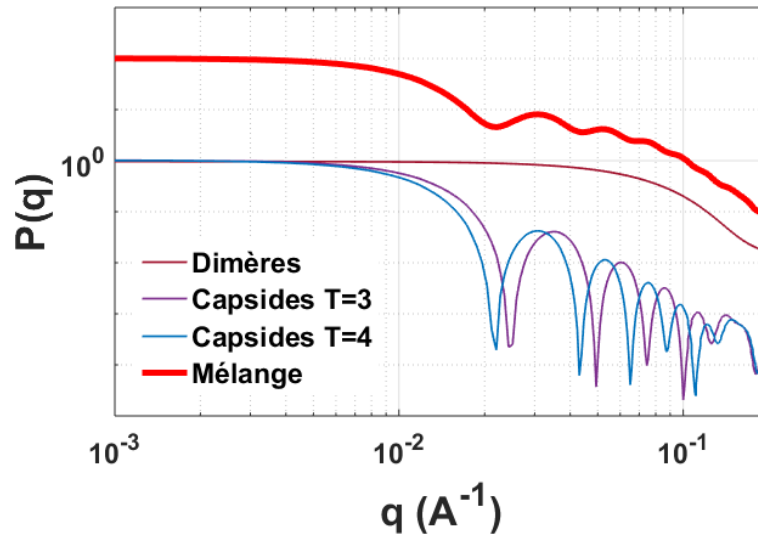
$$I(q) = \frac{N}{V} \overline{\Delta\rho}^2 V_{part}^2 P(q) = K c^{(m)} M P(q) \quad (1.15)$$

où

- $c^{(m)} = \frac{NM}{VN_A}$  est la concentration en masse ( $\text{g.cm}^{-3}$ ) avec  $N_A$  le nombre d'Avogadro,  $M$  la masse molaire de l'objet et  $V$  le volume de la solution,
- $K = \frac{(\overline{v}\Delta\rho)^2}{N_A}$  une constante ( $\text{cm}^2.\text{mol.g}^{-2}$ ),
- $\overline{v} = \frac{V_{part}N_A}{M}$  le volume spécifique ( $\text{cm}^3.\text{g}^{-1}$ ).

Dans le cas d'un échantillon polydisperse en taille ou en forme mais de contraste  $\overline{\Delta\rho}$  identique, les facteurs de forme de toutes les molécules en solution peuvent

être additionnés pour obtenir un motif de diffusion moyen de l'ensemble de l'échantillon (Fig. 3.4).



**Figure 3.4:** Profil SAXS d'une solution polydisperse. Celle-ci contient 5% de dimères de protéines ainsi que 5% de capsides vides  $T = 3$  et 90% de  $T = 4$ . La courbe rouge correspond à la somme des profils de chaque particule isolée tandis que les courbes bleue, violette et bordeaux correspondent aux capsides  $T = 4$ ,  $T = 3$  et aux dimères, respectivement. Il est à noter que la courbe rouge a été rehaussée d'un facteur 100 pour plus de clarté.

Mathématiquement, dans le cas d'un système dilué, l'intensité moyenne dudit système à un instant  $t$  peut être exprimée telle que :

$$I(q, t) = K \sum_i c_i^{(m)}(t) M_i P_i(q) \quad (1.16)$$

où  $c_i^{(m)}(t)$  est la concentration massique,  $M_i$  la masse molaire et  $P_i(q)$  le facteur de forme de chaque espèce  $i$ .

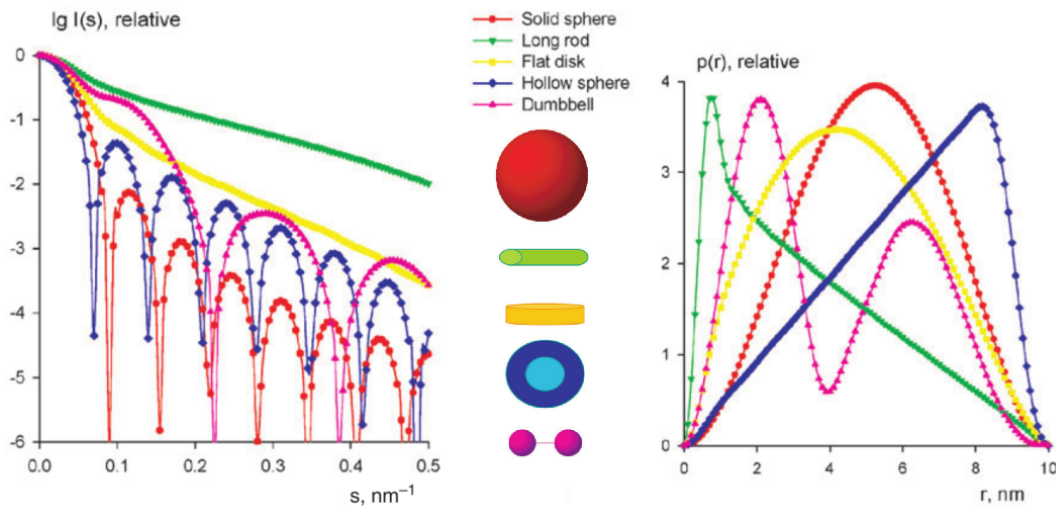
Les facteurs de forme sont des fonctions de l'espace réciproque. Leurs équivalents dans l'espace réel sont les fonctions de distribution de distance  $p(r)$  calculée par une transformée de Fourier inverse telle que :

$$p(r) = V_{part}^2 \frac{r^2}{2\pi^2} \int_0^\infty q^2 P(q) \frac{\sin(qr)}{qr} dq \quad (1.17)$$

La fonction  $p(r)$  tenant compte de l'ensemble des distances entre les composants des molécules, des objets ayant le même rayon de giration mais des formes diffé-



rentes présenteront des distributions différentes (Fig. 3.5). Ainsi, la fonction de



**Figure 3.5:** Intensités de diffusion (à gauche) et fonctions de distribution de distance (à droite) d'objets de formes différentes mais de même taille  $D_{max}$ . (Adapté de [54])

distribution de distance de particules globulaires (en rouge) présentera la forme d'une gaussienne avec un maximum au niveau de la médiane tandis que celle d'une particule allongée (en vert) présentera une asymétrie avec un maximum décalé vers les petites valeurs de  $r$ . Dans le cas de particules en forme d'haltère (en rose), deux composantes apparaissent au sein de la fonction de distribution de distance : l'une représente les distances à l'intérieur des sphères de l'haltère tandis que l'autre représente la distance séparant les sphères.

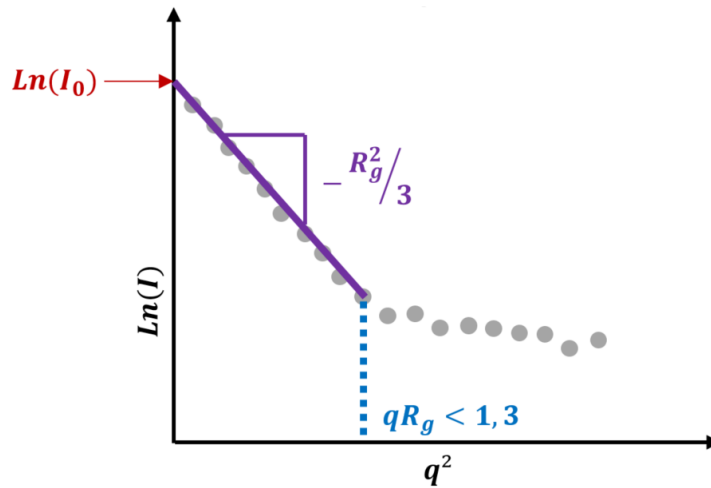
## 1.5 Approximation de Guinier

Dans la limite où  $q$  tend vers zéro, i.e.  $q$  est petit devant l'inverse de la plus grande dimension structurale caractéristique des molécules d'intérêts, le facteur de forme peut s'exprimer assez simplement. Selon Guinier [59], la forme de la courbe est due à la taille globale des particules, de sorte que :

$$I(q) \approx I(0) \exp\left(\frac{-R_g^2 q^2}{3}\right) \quad (1.18)$$

avec  $I(0)$  l'intensité à  $q = 0$  et  $R_g$ , le rayon de giration défini comme :

$$R_g^2 = \frac{1}{2} \frac{\int_{V_{part}} r^4 \gamma(r) dr}{\int_{V_{part}} r^2 \gamma(r) dr} \quad (1.19)$$



**Figure 3.6:** Représentation de Guinier afin de déterminer le rayon de giration de la particule et l'intensité  $I(0)$ .

Le paramètre  $R_g$ , souvent utilisé en première analyse des données SAXS et estimé à partir du graphique de Guinier (Fig. 3.6), ne contient pas d'information sur la forme ou la structure interne des particules mais permet de caractériser la taille. De même, le paramètre  $I(0)$  (Fig. 3.6) permet d'avoir accès aux informations telles que la masse molaire et la concentration des molécules en solution. Il est lié au contraste de densité de longueur diffusée tel que :

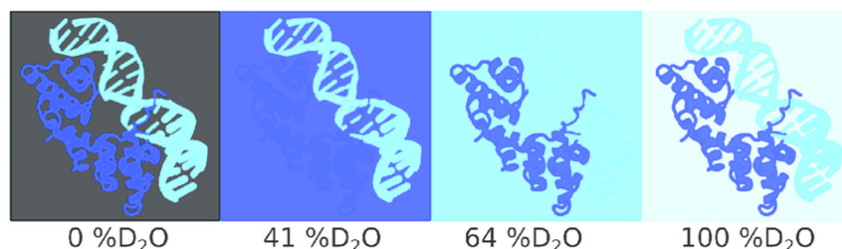
$$I(0) = \frac{N}{V} \overline{\Delta\rho^2} V_{part}^2 \quad (1.20)$$

où  $N$  est le nombre de molécules et  $V$  ainsi que  $V_{part}$  sont les volumes de la solution et de la particule, respectivement.

## 1.6 Diffusion de neutrons aux petits angles

Il est à noter que les études SAS incluent la diffusion de neutrons aux petits angles (SANS, Small-Angle Neutron Scattering en anglais). Le formalisme mathématique utilisé pour décrire la diffusion des rayons X peut également être utilisé pour les neutrons. De plus, le SANS présente quelques avantages par rapport au SAXS : sa sensibilité aux éléments légers ainsi que sa sélectivité vis-à-vis d'isotopes spécifiques[60].

Ainsi, la variation de contraste est une méthode du SANS, utilisée pour étudier la structure des macromolécules dans les systèmes multi-composants, depuis les années 70[61]. L'idée de base de cette méthode est que l'intensité diffusée dépend de  $\Delta\rho$ , i.e. de la façon dont une molécule diffuse par rapport au solvant. Dans le cas du SANS, les densités de longueur de diffusion peuvent changer de manière significative en remplaçant les atomes par leurs isotopes, en particulier en changeant  $^1\text{H}$  en  $^2\text{H}$  (D). Ainsi, en dialysant une molécule dans un solvant deutéré, les hydrogènes labiles de celle-ci vont pouvoir s'échanger avec le solvant, ce qui va modifier sa densité de longueur de diffusion et donc le contraste avec le solvant. Il existe un pourcentage de  $\text{D}_2\text{O}$  dans le solvant pour lequel  $\Delta\rho = 0$ , ce qui signifie qu'il existe une valeur pour laquelle la molécule n'est pas détectée avec des neutrons. Cette valeur dépend du type de molécule. Le principe de la variation de contraste est représenté graphiquement sur la figure 3.7. Considérons un complexe



**Figure 3.7:** Représentation schématique de la variation de contraste en SANS pour un complexe protéine-ADN utilisant différents rapports  $\text{H}_2\text{O}/\text{D}_2\text{O}$ . Adapté de [60].

composé d'une protéine et d'un ADN dans un tampon aqueux. Dans un solvant avec 0% de  $\text{D}_2\text{O}$ , la protéine et l'ADN ont un contraste net. Par conséquent, les deux molécules peuvent être distinguées du solvant. Si la teneur en  $\text{D}_2\text{O}$  est augmentée à environ 41% de  $\text{D}_2\text{O}$ ,  $\Delta\rho$  est nul pour la protéine et seul l'ADN peut être distingué du solvant. De même, si la teneur en  $\text{D}_2\text{O}$  est augmentée à 64%, les neutrons ne font pas de distinction entre l'ADN et le solvant. Dans ce cas, seule la diffusion venant de la protéine peut être observée. Ainsi, en faisant varier la quantité de  $\text{D}_2\text{O}$ , les composants individuels d'une molécule complexe peuvent être étudiés.

# La diffusion de rayons X aux petits angles résolue en temps

Au cours des dernières décennies, des progrès significatifs ont été réalisés afin de permettre la mesure en temps réel de processus cinétiques rapides. Aujourd'hui des phénomènes biologiques se produisant sur une plage temporelle extrêmement large, allant de quelques millisecondes à plusieurs heures, peuvent être observés dans des conditions quasi-physiologiques avec une résolution nanométrique. Cette technique, qui constitue une avancée majeure pour l'étude de ces phénomènes, est connue sous le nom de diffusion de rayons X aux petits angles résolue en temps (TR-SAXS) [62].

Les cinétiques en solution, observées par TR-SAXS, sont souvent initiées par un mélange entraînant une dilution, un saut de  $pH$  ou un saut de salinité. C'est ce mélange qui va générer des changements conformationnels, des formations de complexes, etc. Dans la mesure où ce type d'études renvoie à l'observation de cinétiques rapides, les principaux aspects à prendre en compte dans la conception des expériences TR-SAXS incluent :

- la nécessité de collecter des signaux interprétables dans des délais courts.
- la nécessité d'une observation immédiate après le déclenchement de la perturbation induisant les modifications d'intérêt.

Les lignes de lumière à rayons X des synchrotrons où nous avons effectué nos expériences répondent à la première problématique et seront présentées dans une première partie. Les mélangeurs « stopped-flow », appareils fréquemment utilisés pour étudier les cinétiques rapides, répondent à la seconde problématique et seront donc présentés dans une seconde partie.

## 2.1 Les synchrotrons

Lors des expériences résolues en temps, les objectifs sont d'observer des processus très rapides ou de mesurer des quantités ou des propriétés difficiles à détecter. Il est donc important d'avoir des sources permettant d'obtenir des statistiques d'intensité

suffisantes sur quelques millisecondes et des détecteurs de haute sensibilité. Les faisceaux de rayons X des synchrotrons, produits par des électrons de haute énergie circulant autour des-dits synchrotrons avant d'être émis sous forme de dizaines de faisceaux minces collimatés, répondent à ces problématiques.

Nous avons effectué nos expériences TR-SAXS dans deux synchrotrons : SOLEIL et ESRF, acronymes de « Source Optimisée de Lumière d'Énergie Intermédiaire du LURE<sup>1</sup> » et « European Synchrotron Radiation Facility », respectivement. Les lignes de lumière où nous avons effectué nos expériences, à savoir SWING et ID02, sont optimisées afin de répondre aux nombreuses questions liées à la matière molle, à la conformation de macromolécules en solution et aux matériaux composites en sciences des matériaux. Leurs caractéristiques ainsi que les équipements que nous avons utilisés durant nos expériences sont résumés succinctement dans le tableau 3.1 présenté ci-dessous.

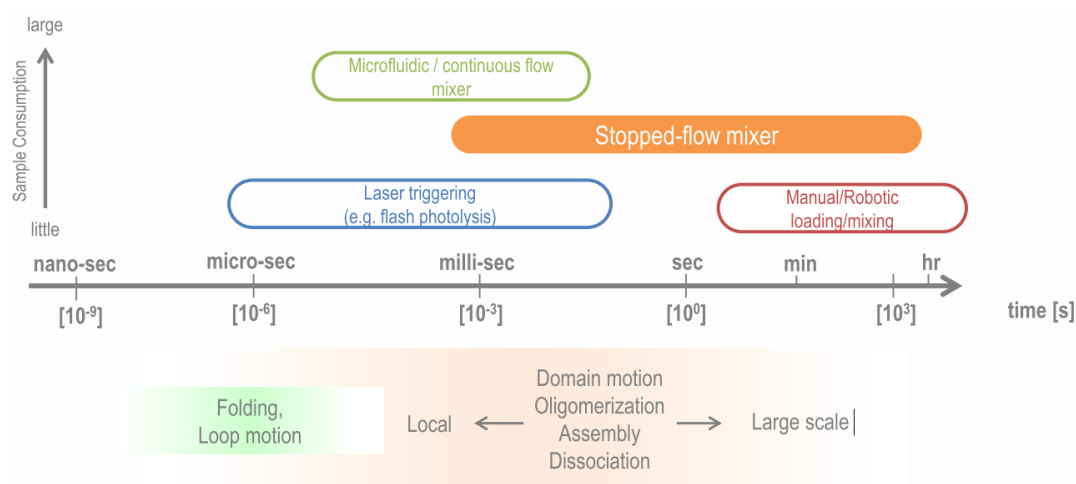
**Table 3.1:** Tableau récapitulatif des caractéristiques et équipements

	ESRF (ID02)	SOLEIL (SWING)
Monochromateur	Cristal $S_i(111)$	Cristal $S_i(111)$
Détecteur	Rayonix MX-170HS	Eiger 4M
<u>Caractéristiques des faisceaux :</u>		
Gamme d'énergie (en keV)	8 – 17	5 – 16
Résolution (en eV)	$\approx 2,5$	$\approx 2$
<u>Caractéristiques des sources :</u>		
Distance de l'échantillon au détecteur	0,8 m à 31 m	0,5 m à 6,5 m
Flux sur l'échantillon	$10^{13}$ ph/s à 12,5 keV (100 mA)	$10^{12}$ ph/s à 16 keV (500 mA)
<u>Environnement des échantillons :</u>		
	<ul style="list-style-type: none"> <li>· SFM 400</li> <li>· bain thermostaté</li> <li>· capillaires à flux continu</li> </ul>	<ul style="list-style-type: none"> <li>· SFM 400</li> <li>· bain thermostaté</li> <li>· supports capillaires faits maison</li> <li>· robot d'injection</li> <li>· échantillonneur automatique</li> </ul>

1. Laboratoire pour l'Utilisation du Rayonnement Electromagnétique

## 2.2 Le stopped-flow

Comme exposé précédemment, les objectifs des expériences résolues en temps sont l'observation de processus très rapides ou la mesure de quantités ou de propriétés difficiles à détecter. Il est donc important d'avoir un appareil permettant de réaliser des mélanges homogènes dans des délais courts. Le graphique ci-après (Fig. 3.8<sup>2</sup>), présente les résolutions temporelles de plusieurs techniques ou appareils dont le « stopped-flow » et justifie parfaitement l'usage de ce dernier.



**Figure 3.8:** Échelle de temps accessibles par différentes techniques lors d'études de mélanges réactionnels.

Pour initier les réactions, lors de nos études par TR-SAXS, nous avons donc utilisé les mélangeurs « stopped-flow » et, plus précisément, les SFM-400 et SFM-4000 de la marque BioLogic. Ces appareils permettent d'effectuer un mélange turbulent, dans un laps de temps très court (de l'ordre de quelques millisecondes), à partir de quatre seringues différentes (Fig. 3.9). Les seringues sont gérées par quatre moteurs indépendants ce qui permet d'avoir différents rapports de mélange à partir des solutions mères, un contrôle précis du débit et des volumes d'injection.

### 2.2.1 Le principe

Le principe de fonctionnement du stopped-flow est simple. Les solutions de chaque seringue sont envoyées simultanément. Celles-ci se mélangent au passage des

2. Image de Thomas M. Weiss, Stanford University, disponible à l'adresse : [https://www-ssrl.slac.stanford.edu/~saxs/download/weiss\\_practical.pdf](https://www-ssrl.slac.stanford.edu/~saxs/download/weiss_practical.pdf)

mélangeurs. De manière schématique, le mélange pourrait être décrit ainsi : les solutions des seringues S1 et S2 sont d'abord mélangées dans un premier mélangeur avant d'être incorporées à la solution S3 dans le deuxième mélangeur. L'ensemble est ensuite ajouté à la solution S4 dans le troisième mélangeur avant d'être envoyé vers la cuvette d'observation (Fig. 3.10).



Figure 3.9: Photo de l'appareil SFM 4000.

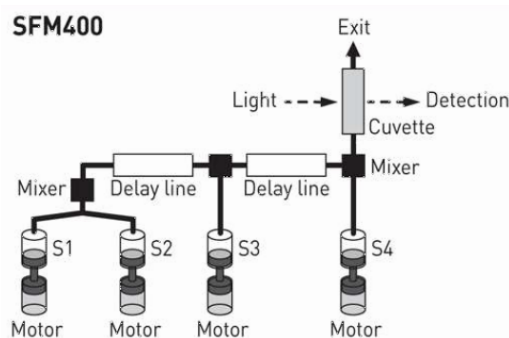


Figure 3.10: Schéma de fonctionnement du SFM-4000

Cette description met en évidence un temps mort, i.e. un temps nécessaire au mélange et transfert de la solution vers la cuvette, qu'il est important de déterminer :

$$t_{mort} = \frac{V_{mort}}{D_{tot}} \quad (2.1)$$

où  $V_{mort}$  correspond au plus grand volume entre la seringue et la cuvette et  $D_{tot}$  est le débit correspondant. Pour un SFM-4000, le volume mort est de l'ordre de quelques dizaines de  $\mu\text{L}$  et le débit total de l'ordre de quelques  $\text{mL}\cdot\text{s}^{-1}$ .

Il est à noter que le stopped-flow doit être synchronisé avec le reste de l'instrumentation (faisceau incident, détecteur, ...) afin de ne pas manquer les premières étapes de nos expériences d'assemblage ou de désassemblage. La synchronisation est effectuée en prenant en compte le temps mort,  $t_{mort}$ , présenté précédemment.

## 2.2.2 Les équipements

Plusieurs équipements ont été optimisés dans le « stopped-flow » afin d'éviter des artefacts. Cette section met en avant ceux qui ont été particulièrement utiles pour nos expériences :

- La cuvette d'observation qui est optimisée pour minimiser l'absorption des rayons X et supporter la pression due au débit. Elle est composée de quartz, a une épaisseur de 10  $\mu\text{m}$  et un chemin optique d'environ 1,5 mm.
- L'électrovalve qui permet de stopper très rapidement (quelques millisecondes) le mélange au sein de la cuvette d'observation. Ce mécanisme de fermeture de la cuvette, appelé « hard-stop », est synchronisé avec les moteurs des seringues ce qui permet d'éviter des pressions anormales pouvant générer des artéfacts. Le « hard-stop » reste également fermé durant l'ensemble de l'acquisition ce qui permet d'éviter des changements de quantités de matière au sein de la cuvette pouvant rendre impossible l'interprétation des données.
- Les mélangeurs dont il existe 2 types : Le mélangeur usuel appelé "Berger-Ball" et le mélangeur "HDS". Ce dernier est utile dans le cas d'une différence de densité importante entre les solutions stockées dans les seringues. En effet, de la convection se crée dans la cuvette lorsque la solution dans la cuvette et une des solutions d'origine à la sortie du mixeur ont une différence de densité. Cette convection génère une perturbation du signal pouvant rendre impossible l'interprétation des données.
- Le bain thermostaté, externe au stopped-flow, qui permet de maintenir l'appareil à une température située entre 10 et 70°C.

### 2.2.3 Les précautions à prendre

Malgré ces aménagements réalisés sur les « stopped-flow » pour permettre l'acquisition de données de qualité, un certain nombre de précautions sont à prendre pour avoir une bonne utilisation de l'appareil :

- L'un des artéfacts le plus courant lors d'expériences réalisées avec un « stopped-flow » est la formation de bulles pouvant rendre impossible l'interprétation des données. Il est donc important de dégazer les solutions avant toute utilisation.
- Un problème courant est également la pollution de l'expérience par un ancien échantillon. Par exemple, des protéines peuvent rester collées sur les parois de la cuvette d'observation suite à des dégâts d'irradiation et modifier le signal diffusé. Il est donc important de nettoyer régulièrement la cuvette



du « stopped-flow » et de confirmer l'état de celle-ci par des mesures sur les solutions tampons.

- Il peut arriver que les solutions stockées dans les seringues migrent et modifient le signal diffusé. Il est donc important de positionner les seringues contenant les solutions de manière à limiter ce phénomène. Par exemple, on placera les solutions les plus similaires en position S1 et S2 (Fig. 3.10).

# Outils d'analyse des données

## SAXS

Une fois obtenus, tous les jeux de données doivent être analysés. L'objectif de ce chapitre est donc de présenter certains outils qui peuvent être utiles à l'analyse des données de TR-SAXS. En premier, je présenterai les outils mathématiques nécessaires à l'étude de l'assemblage des capsides pleines. Je présenterai ensuite la décomposition en valeurs singulières (SVD, Singular Value Decomposition en anglais) qui permet d'analyser des jeux de données en mettant en évidence des caractéristiques de ces données.

### 3.1 Outils mathématiques pour l'assemblage des capsides pleines

Un des enjeux majeurs de cette thèse était de donner un aperçu quantitatif de la dynamique d'auto-assemblage des capsides pleines. Notamment, nous voulions pouvoir estimer le nombre de dimères pouvant se lier sur le génome, les énergies de liaison, ou encore, investiguer les mécanismes contrôlant la réaction. De ce fait, la section, ci-après, présentera une méthode pour déterminer le nombre de dimères liés au génome à partir des intensités diffusées  $I_0 \equiv I(q = 0)$ .

#### 3.1.1 Détermination de $\langle N \rangle_{up}$ et $\langle N \rangle(t \rightarrow +\infty)$

En effet, l'intensité diffusée  $I_0$ , d'une solution contenant des complexes nucléoprotéiques amorphes (NPC<sup>1</sup>) résultant de la liaison des dimères de protéines sur le génome du virus, est donnée par la relation suivante[54] :

$$I_0 = \Delta b_S^2 X_S + \sum_{N=0}^{+\infty} X_N (\Delta b_G + \Delta b_S N)^2 \quad (3.1)$$

1. Nucleoprotein complexes, en anglais

où  $\Delta b_S$  et  $\Delta b_G$  sont les contrastes de longueur de diffusion des dimères et du génome, respectivement.  $X_S$  est la concentration molaire en dimères de protéines libres et  $X_N$  la concentration molaire en complexes d'ordre  $N$ .

Nous pouvons noter que  $X_0$ , la concentration molaire en complexes lorsqu'il n'y a pas encore de dimères liés au génome, est équivalente à  $X_G$ , la concentration molaire en génome seul. De plus, la conservation de la masse impose deux contraintes sur les concentrations molaires, à savoir :

$$\begin{cases} X_G + \sum_{N=1}^{+\infty} X_N = c_G \\ X_S + \sum_{N=1}^{+\infty} N X_N = c_S \end{cases} \quad (3.2)$$

où  $c_S$  et  $c_G$  sont les concentrations molaires totales en dimères et génome dans la solution.

De ce fait, le nombre moyen de dimères par complexe peut s'écrire tel que :

$$\langle N \rangle = \frac{\sum_{N=0}^{+\infty} N X_N}{X_G + \sum_{N=1}^{+\infty} X_N} = \frac{\sum_{N=0}^{+\infty} N X_N}{c_G} \quad (3.3)$$

Une expression semblable peut être obtenue pour  $\langle N^2 \rangle$ , ce qui permet de réorganiser et développer l'équation 3.1, à l'aide des équations 3.2, sous la forme :

$$\Delta b_S^2 c_G \langle N^2 \rangle + (2\Delta b_G \Delta b_S - \Delta b_S^2) c_G \langle N \rangle (I_0 - I_0^*) = 0 \quad (3.4)$$

où  $I_0^* \equiv \Delta b_G^2 c_G + \Delta b_S^2 c_S$  est l'intensité diffusée produite par le mélange des dimères et du génome, avant la formation des complexes.

De plus,  $\langle N^2 \rangle$  peut être écrite sous la forme  $\langle N \rangle^2 + \sigma_N^2$  avec  $\sigma_N^2$  la variance du nombre de dimères par complexe.  $\langle N \rangle$  peut être alors calculé en résolvant une équation quadratique et devient fonction du paramètre inconnu  $\sigma_N$  :

$$\langle N \rangle = \sqrt{\lambda^2 + \left( \frac{I_0 - I_0^*}{\Delta b_S^2 c_G} - \sigma_N^2 \right)} - \lambda \quad (3.5)$$

avec  $\lambda = \Delta b_G / \Delta b_S - 1/2 \approx \Delta b_G / \Delta b_S > 0$ .

A partir de cette expression, nous pouvons déduire une limite supérieure pour  $\langle N \rangle$ , donnée par  $\sigma_N = 0$ , i.e.  $\langle N \rangle \leq \langle N \rangle (\sigma_N = 0) \equiv \langle N \rangle_{up}$ .

A l'équilibre, les expressions analytiques données par le modèle d'assemblage, développé plus bas (voir section 3.1.2), peuvent être injectées dans l'équation 3.4. Ceci donne une équation quadratique reliant le nombre moyen de dimères par complexes à l'intensité diffusée, tous deux à l'équilibre, telle que :

$$\langle N \rangle(t \rightarrow +\infty) = \sqrt{\frac{1}{4}\lambda^2 + \frac{I_0(t \rightarrow +\infty) - I_0^*}{\Delta b_S^2 c_G}} - \frac{1}{2}\lambda \quad (3.6)$$

### 3.1.2 Modèle d'assemblage

De plus, comme indiqué au début de la section, nous souhaitons obtenir un aperçu quantitatif des énergies de liaison et de la diffusion des dimères de protéines au cours de l'auto-assemblage des NPCs. De ce fait, nous avons analysé les données à l'aide d'un modèle simple d'assemblage , i.e. :



où  $S$  désigne les dimères libres,  $C_N$  les complexes d'ordre  $N$  avec  $N \in \mathbb{N}$  et  $C_0$  le génome seul.  $k^+$  et  $k^-$  sont les constantes de vitesse de la réaction.

Comme nous avons pu le voir dans la section précédente, la conservation de la masse impose des contraintes sur les concentrations molaires. De ce fait, la réaction, présentée précédemment, peut être décrite par l'évolution des concentrations molaires donnée par des équations différentielles telles que :

$$\begin{cases} \frac{dX_G}{dt} = k^- X_1 - k^+ X_S X_G \\ \vdots \\ \frac{dX_N}{dt} = k^- (X_{N+1} - X_N) + k^+ (X_{N-1} - X_N) X_S \\ \vdots \end{cases} \quad (3.8)$$

A l'équilibre, l'équation 3.8 devient  $X_N = (K X_S)^N X_G$  avec  $K = k^+/k^-$ , la constante d'équilibre. L'équation 3.2 permet alors d'obtenir une équation quadratique fonction de  $X_S$  :

$$K X_S^2 - [1 + K(c_S + c_G)] X_S + c_S = 0 \quad (3.9)$$

Les solutions de l'équation 3.9 sont présentées graphiquement sur la figure 3.11, à gauche. De plus, dans l'hypothèse où  $c_S \gg c_G$ , nous arrivons à un résultat classique dans les théories d'auto-assemblage, à savoir :

$$X_S \approx \begin{cases} c_S, & c_S < K^{-1} \\ K^{-1}, & c_S \geq K^{-1} \end{cases} \quad (3.10)$$

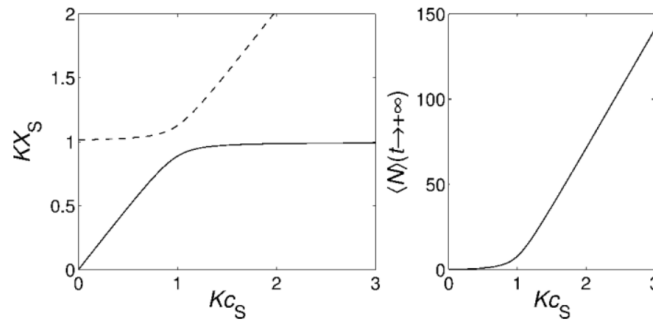
où  $K^{-1}$  devient la concentration critique en dimères au-dessus de laquelle les NPCs peuvent être formées.

Pour chaque instant, le nombre moyen de dimères de protéines par complexe est donné par :

$$\langle N \rangle = \frac{\sum_{N=0}^{+\infty} N X_N}{X_G + \sum_{N=1}^{+\infty} X_N} = \frac{c_S - X_S}{c_G} \quad (3.11)$$

La figure 3.11, à droite, montre graphiquement la solution de l'équation 3.11. De plus, dans nos expériences,  $c_S \geq K^{-1}$  (Équation 3.10). De ce fait, nous obtenons, à l'équilibre, un nombre moyen de dimères de protéines par NPCs tel que :

$$\langle N \rangle(t \rightarrow +\infty) \approx \frac{c_S - K^{-1}}{c_G} \quad (3.12)$$



**Figure 3.11:** Concentration critique à l'équilibre. Solution des équations 3.9 et 3.11 pour  $c_G = 0,2 \mu\text{M}$  et  $K^{-1} = 14 \mu\text{M}$ , les concentrations typiques pour nos expériences. La ligne en pointillé indique la branche instable.

La concentration critique  $K^{-1}$  évolue suivant une loi de type Arrhénius et peut s'écrire :

$$K^{-1} = N_A^{-1} v_0^{-1} \exp(G_{bind}/k_B T) \quad (3.13)$$

où  $N_A$  est la constante d'Avogadro,  $G_{bind}$  l'énergie libre de liaison d'un dimère libre sur un complexe,  $k_B$  la constante de Boltzmann,  $T$  la température et  $v_0$  le volume

dans l'état standard, considéré ici comme étant  $4/3\pi R_g^3$  avec  $R_g = 33 \text{ \AA}$ , le rayon de giration des dimères.

Le modèle d'assemblage peut être encore affiné en introduisant un nombre maximal de dimères par NPCs,  $N_{max}$ . La conservation de la masse est alors modifiée, i.e.  $N \in [1 ; N_{max}]$ , ce qui donne une équation reliant  $X_S$  et  $c_S$  :

$$c_s = X_S + c_G \frac{K X_S}{1 - K X_S} \frac{1 - (N_{max} + 1)(K X_S)^{N_{max}} + N_{max}(K X_S)^{N_{max}+1}}{1 - (K X_S)^{N_{max}+1}} \quad (3.14)$$

Dans l'hypothèse raisonnable où  $N_{max} \gg 1$ , nous pouvons obtenir deux solutions approchées de l'équation précédente :

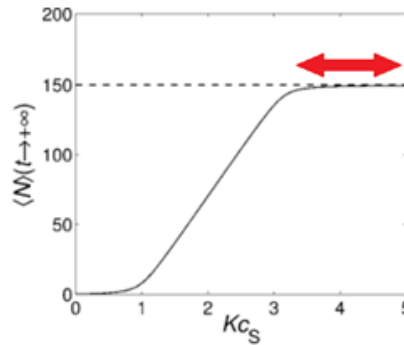
- Pour  $K X_S < 1$ , l'équation 3.9 est solution de l'équation précédente et permet de décrire la concentration critique d'assemblage.
- Pour  $K X_S > 1$ , l'équation permettant de décrire la saturation du génome, en dimères liés, peut s'écrire :

$$X_S + \frac{c_G}{1 - K X_S} = c_S - c_G N_{max} \quad (3.15)$$

Dans tous les cas, en supposant que  $c_S \gg c_G$ , à l'équilibre, le nombre moyen de dimères par complexes, donné par les équations 3.14 et 3.11, peut être approximé par la fonction suivante :

$$\langle N \rangle(t \rightarrow +\infty) \approx \begin{cases} 0, & c_S \leq K^{-1} \\ \frac{c_S - K^{-1}}{c_G}, & K^{-1} \leq c_S \leq K^{-1} + c_G N_{max} \\ N_{max}, & c_S \geq K^{-1} + c_G N_{max} \end{cases} \quad (3.16)$$

La figure 3.12 illustre graphiquement la fonction présentée précédemment (éq.



**Figure 3.12:** Concentration critique à l'équilibre. Solution des équations 3.14 et 3.11 pour  $c_G = 0,2 \mu\text{M}$ ,  $K^{-1} = 14 \mu\text{M}$  et  $N_{max} = 150$ .

3.16). Dans ce cas, il existe une limite du nombre de dimères pouvant se lier au génome. De ce fait, la courbe  $\langle N \rangle(t \rightarrow +\infty)$  présente un plateau, mis en évidence sur la figure 3.12 par une flèche rouge.

## 3.2 Décomposition en valeurs singulières

Lors de nos expériences, nous avons également sondé l'auto-assemblage ou le désassemblage de capsides vides par TR-SAXS. Autrement dit, nous avons obtenu une matrice d'intensités mesurées  $\mathbf{I}$  qui est une combinaison linéaire pondérée en masse des intensités diffusées de toutes les espèces présentes au cours de la cinétique. Elle peut être écrite comme :

$$\mathbf{I} \approx \mathbf{E}\mathbf{C} \quad (3.17)$$

où la matrice  $\mathbf{E}$  contient les données des espèces inconnues présentes en solution et la matrice  $\mathbf{C}$  contient les données inconnues de l'évolution des concentrations de ces espèces. L'objectif est de reconstruire les matrices  $\mathbf{E}$  et  $\mathbf{C}$  tout en déterminant le modèle d'association et les constantes de temps. L'analyse des données de TR-SAXS nécessite donc des approches adaptées. Une des techniques utilisées pour séparer et analyser les contributions aux signaux SAXS de chaque espèce en solution est la décomposition en valeurs singulières [62][63]. Ce chapitre a pour objectif de présenter cette analyse. Nous l'avons choisie car elle est robuste, ne dépend d'aucun modèle, permet de séparer le bruit des composantes significatives du signal ainsi qu'un contrôle non-biaisé du nombre d'espèces présentes en solution [64].

### 3.2.1 Traitement mathématique

La matrice des intensités mesurées  $\mathbf{I}$  contient les informations temporelles et structurales telle que :

$$\mathbf{I} = \begin{pmatrix} I(q_1, t_1) & I(q_1, t_2) & \cdots & I(q_1, t_N) \\ I(q_2, t_1) & I(q_2, t_2) & \cdots & I(q_2, t_N) \\ \vdots & \vdots & \ddots & \vdots \\ I(q_M, t_1) & I(q_M, t_2) & \cdots & I(q_M, t_N) \end{pmatrix} \quad (3.18)$$

où  $q$  représente le nombre d'onde, i.e. la norme du vecteur  $\vec{q}$  présenté précédemment (sous-section 1.1),  $t$  le temps à chaque prise de mesures,  $M$  et  $N$  les dimensions de la matrice telles que  $M > N$ .

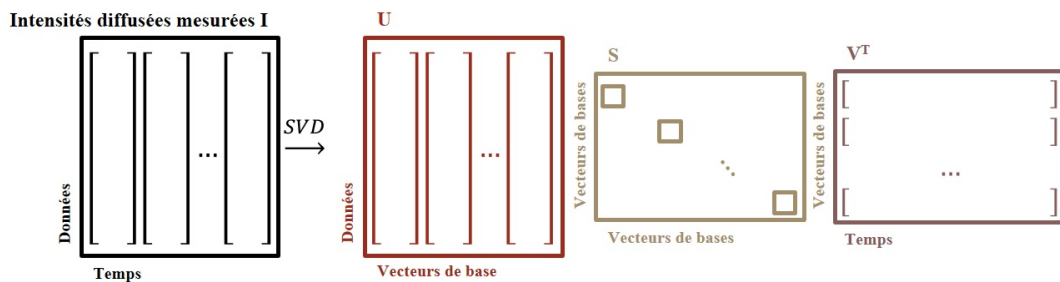
### 3.2.1.1 Le principe

La décomposition en valeurs singulières permet de réécrire cette matrice comme le produit de trois matrices (Fig. 3.13) :

$$\mathbf{I} = \mathbf{USV}^T \quad (3.19)$$

où  $\mathbf{U}$  est une matrice contenant l'information spatiale,  $\mathbf{S}$  une matrice diagonale de valeurs non-nulles et  $\mathbf{V}^T$  une matrice contenant l'information temporelle.

La figure 3.13 est une représentation schématique du processus d'obtention des matrices  $\mathbf{U}$ ,  $\mathbf{S}$  et  $\mathbf{V}^T$  par SVD. La matrice  $\mathbf{U}$  (en rouge) est une matrice  $M \times N$  dont les colonnes forment un ensemble de vecteurs de base contenant les facteurs de forme des espèces en solution. La matrice  $\mathbf{S}$  (en jaune) une matrice  $N \times N$  diagonale de valeurs non-nulles appelées valeurs singulières de  $\mathbf{S}$  et notées  $s_i$ . La matrice  $\mathbf{V}^T$  (en violet) une matrice  $N \times N$  dont les lignes forment un ensemble de vecteurs de base contenant l'évolution de la concentration des espèces.



**Figure 3.13:** Schéma de principe de l'analyse par décomposition en valeurs singulières (SVD) (Adapté de [64]).

Idéalement, le nombre de vecteurs de base significatifs doit être égal au nombre de particules différentes,  $N_{\text{espèce}}$ , présentes en solution. Lorsque l'analyse débute,  $N_{\text{espèce}}$  n'est pas connu mais cette valeur peut être estimée à partir de l'analyse par décomposition en valeurs singulières et, notamment, à l'aide de ces grandeurs :

- l'auto-corrélation  $C$  des vecteurs de base  $\vec{U}_i$  et  $\vec{V}_i$



- l'amplitude relative des valeurs singulières  $s_i$ , de la matrice  $\mathbf{S}$ .
- La comparaison des résidus, issus de l'évaluation de la matrice tronquée  $\tilde{\mathbf{I}}$  par rapport à la matrice initiale  $\mathbf{I}$ , avec le bruit expérimental.

### 3.2.1.2 Auto-corrélation des vecteurs de base

Mathématiquement, l'auto-corrélation des vecteurs de base peut être définie par les relations suivantes :

$$C(\vec{U}_i) = \sum_{j=1}^{M-1} U_{j,i} U_{j+1,i} \quad (3.20)$$

$$C(\vec{V}_i) = \sum_{j=1}^{N-1} V_{j,i} V_{j+1,i} \quad (3.21)$$

Les vecteurs propres étant normalisés, l'auto-corrélation des vecteurs de base indique la similitude entre deux signaux. Autrement dit, les auto-corrélations proches de 1 indiquent un rapport signal sur bruit élevé et donc, la présence d'informations. Les vecteurs propres correspondants peuvent être pris en compte pour un traitement ultérieur. Inversement, les valeurs faibles indiquent un rapport en faveur du bruit et les vecteurs propres peuvent être supprimés. (Fig. 3.14a)

### 3.2.1.3 Valeurs singulières

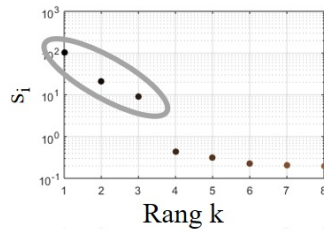
Les valeurs singulières sont analogues à des coefficients qui vont pondérer chaque vecteur de base des matrices  $\mathbf{U}$  et  $\mathbf{V}$ . Les valeurs  $s_i$  sont donc représentatives de la contribution de chaque vecteur  $\vec{U}_i$  et  $\vec{V}_i$  au signal. Classées par ordre décroissant, le bruit peut être séparé des composantes significatives du signal. Schématiquement, on peut dire que lorsque les premières composantes de la matrice  $\mathbf{S}$  ont des valeurs nettement supérieures aux autres, cela signifie que seules ces composantes sont représentatives des espèces en solution. Les autres correspondent aux bruits expérimentaux (Fig. 3.14b).

### 3.2.1.4 Matrice tronquée

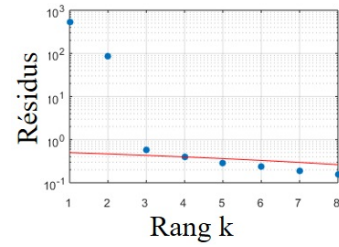
Mathématiquement, les matrices  $\mathbf{U}$ ,  $\mathbf{V}$  et  $\mathbf{S}$  peuvent être tronquées pour reconstruire une matrice  $\tilde{\mathbf{I}} = \mathbf{U}_k \mathbf{S}_k \mathbf{V}_k^T$  avec  $k$  le nombre de valeurs singulières de  $\mathbf{I}$

$i$	$C(\mathbf{U}_i)$	$C(\mathbf{V}_i)$
1	0.9016	0.8389
2	0.7244	0.5255
3	0.9323	0.7205
4	0.7725	0.4086
5	0.5090	-0.2307

(a) Auto-corrélation des vecteurs de base  $\vec{U}_i$  et  $\vec{V}_i$



(b) Valeurs singulières de la matrice  $\mathbf{S}$



(c) Résidus de la matrice tronquée  $\tilde{\mathbf{I}}$  par rapport à la matrice initiale  $\mathbf{I}$  (points bleus) comparés au bruit expérimental (ligne rouge)

**Figure 3.14:** Critères permettant de déterminer le nombre de composantes suffisantes pour décrire les données expérimentales : le tableau 3.14a présente 3 à 4 valeurs significatives, le graphique 3.14b montre 3 composantes avec des valeurs nettement supérieures aux autres et le graphique 3.14c montre 2 à 3 valeurs de résidus (points bleus) supérieures au bruit expérimental (courbe rouge). Ensemble, ces observations donnent à penser qu'il y a 3 espèces en solution dans la cinétique étudiée.

sélectionnées. La matrice  $\tilde{\mathbf{I}}$  contient un nombre suffisant de composantes pour décrire correctement la matrice  $\mathbf{I}$ . Elle fournit la meilleure approximation de rang  $k$ , déterminée par la méthode des moindres carrés, de la matrice  $\mathbf{I}$ . Cette matrice permet également d'évaluer la matrice initiale  $\mathbf{I}$  via la relation suivante :

$$\|\mathbf{I} - \tilde{\mathbf{I}}\|^2 = \sum_{i=k+1}^N s_i^2 \leq \mu\nu\sigma^2 \quad (3.22)$$

où  $\tilde{\mathbf{I}}$  est la matrice tronquée avec  $k$  le nombre de valeurs singulières de la matrice  $\mathbf{I}$  sélectionnées.  $\mu$  et  $\nu$  sont des paramètres définis tels que  $\mu = M - k$  et  $\nu = N - k$  [64].  $\sigma^2$  est la variance des données.

La comparaison des résidus issus de cette évaluation avec le bruit expérimental est le troisième critère de détermination du nombre de composantes suffisantes pour décrire les données expérimentales (Fig. 3.14c).

### 3.2.2 Procédure de rotation de matrice

Il faut garder à l'esprit que les vecteurs de base des matrices  $\mathbf{U}$  et  $\mathbf{V}$  n'ont pas de sens physique en soi. Il est nécessaire de faire une rotation de ces matrices afin de pouvoir reconstruire des signaux interprétables.

### 3.2.2.1 Les facteurs de forme des espèces en solution

La matrice des intensités peut être décrite par les  $N$  premières composantes, i.e.  $\vec{U}_1$  à  $\vec{U}_N$ ,  $\vec{V}_1$  à  $\vec{V}_N$  ainsi que  $s_1$  à  $s_N$ . Comme expliqué précédemment dans la section 3.2.1.1, la matrice  $\mathbf{U}$  forme un ensemble de vecteurs de base contenant les facteurs de forme des espèces en solution. Cependant, les vecteurs de base  $\vec{U}_1$  à  $\vec{U}_N$  ne sont pas directement représentatifs des  $N$  espèces présentes dans la solution. En réalité, le signal de ces espèces est réparti sur les  $N$  premières valeurs singulières et premiers vecteurs de base. En pratique, nous nous limitons au rang  $N = 2$ . En effet, le signal des deux espèces  $\vec{U}_{\text{espèce1}}$  et  $\vec{U}_{\text{espèce2}}$  peut être exprimé comme une combinaison linéaire des deux composantes  $\vec{U}_1$  à  $\vec{U}_2$  qui contiennent toutes les informations spectrales tirées des données. Les composantes  $\vec{U}_1$  à  $\vec{U}_2$  forment une base orthonormale ce qui permet de déterminer la matrice de rotation  $\mathbf{R}$  contenant les angles  $\theta_1$  et  $\theta_2$  de cette dernière tel que :

$$\left( \vec{U}_{\text{espèce1}} \ \vec{U}_{\text{espèce2}} \right) = \left( \vec{U}_1 \ \vec{U}_2 \right) \times \mathbf{R} \quad (3.23)$$

avec

$$\mathbf{R} = \begin{pmatrix} \cos \theta_1 & -\sin \theta_2 \\ \sin \theta_1 & \cos \theta_2 \end{pmatrix}$$

$$\vec{U}_{\text{espèce1}} = \vec{U}_1 \cdot \cos \theta_1 + \vec{U}_2 \cdot \sin \theta_1 \quad (3.24a)$$

$$\vec{U}_{\text{espèce2}} = -\vec{U}_1 \cdot \sin \theta_2 + \vec{U}_2 \cdot \cos \theta_2 \quad (3.24b)$$

où  $\vec{U}_{\text{espèce1}}$  et  $\vec{U}_{\text{espèce2}}$  sont représentatifs des espèces en solution puisqu'ils donnent les facteurs de forme des deux espèces 1 et 2 après normalisation.

### 3.2.2.2 Évolution de la concentration des espèces en solution

Une fois les facteurs de forme connus, il est possible d'estimer l'évolution de la concentration des espèces en fonction du temps, via l'analyse SVD. En effet, une fois la matrice  $\mathbf{R}$  obtenue, nous pouvons retrouver la concentration des espèces (en unité arbitraire) en solution, à chaque instant  $t$ , via l'équation suivante :

$$\left( V_{\text{espèce1}} \ V_{\text{espèce2}} \right) = \mathbf{R}^{-1} \mathbf{S}_N \mathbf{V}_N^T \quad (3.25)$$

où  $\mathbf{S}_N$  et  $\mathbf{V}_N^T$  sont les matrices tronquées au rang  $N = 2$

Dans la pratique, les angles  $\theta_1$  et  $\theta_2$  ne sont pas commutés aléatoirement. Ils doivent l'être de telle façon que la concentration ait une signification physique, i.e. des valeurs de concentration positives et une évolution dans le sens attendu. S'il est à noter que la SVD n'est qu'une approximation qui dépend d'un paramètre non physique (la matrice  $\mathbf{R}$ ), la procédure de rotation SVD permet tout de même d'avoir une bonne approximation des facteurs de forme et de l'évolution des concentrations des espèces en solution.



# IV

---

Cinétique d'auto-assemblage du  
CCMV

L'auto-assemblage est un processus omniprésent en biologie. Ceci est particulièrement vrai en ce qui concerne la formation des capsides virales, conçues pour emballer le matériel génétique du virus. Très tôt, il a été démontré que certains virus avaient des capsides pouvant s'assembler, *in vitro*, en l'absence du génome[65]. Cette propriété remarquable, associée à un certain nombre de caractéristiques intéressantes des capsides virales (monodispersité, biocompatibilité, etc), a motivé l'intérêt de la communauté scientifique pour ce sujet. Cependant, si la dynamique d'auto-assemblage des capsides vides a été largement étudiée par diverses techniques (AFM<sup>1</sup> [66], SLS<sup>2</sup> [67][68], TR-SAXS[69][70], etc), celle des capsides pleines est encore méconnue. La dynamique étant intrinsèquement plus complexe, du fait de la coexistence de composants différents, les données expérimentales sont rares et l'interprétation des résultats est souvent difficile.

Ainsi, afin de pouvoir étudier ce phénomène complexe, nous avons opté pour le virus de la marbrure chlorotique de la cornille (CCMV). Ce virus, caractérisé pour la première fois en 1967, est non seulement facile à cultiver et à isoler mais il est également possible de le désassembler et réassembler *in vitro* en jouant sur le *pH* et la force ionique. De ce fait, ce virus, devenu un modèle pour les bio-physiciens, est le candidat idéal pour une étude quantitative de la dynamique d'auto-assemblage des protéines de capsides avec leur matériel génétique.

Cette partie du manuscrit est divisée en trois chapitres. Dans un premier temps, je présenterai une rapide étude bibliographique sur l'auto-assemblage des capsides pleines du CCMV. Ces phénomènes étant dirigés par des interactions électrostatiques et hydrophobes, je présenterai ensuite une approche expérimentale visant à déterminer les énergies d'interaction existantes au sein de la capside ainsi que les charges effectives portées par les dimères de protéines dans diverses conditions ioniques et de *pH*. Je présenterai enfin les résultats ainsi que les analyses obtenus pour l'étude des cinétiques d'assemblage des capsides pleines du CCMV.

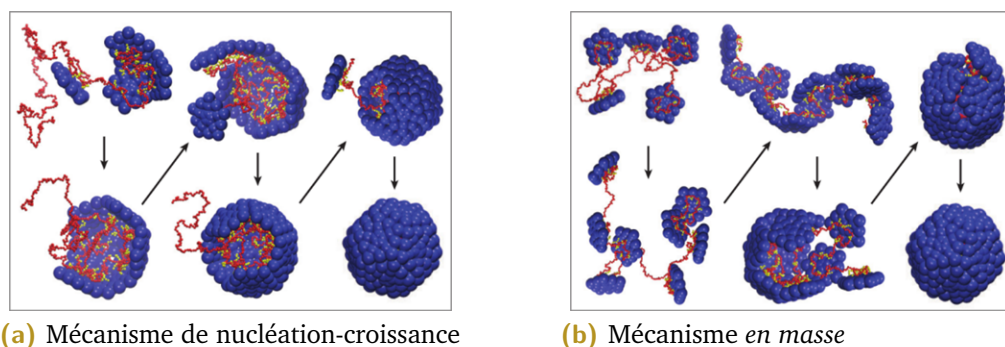
---

1. Microscope à force atomique. AFM pour atomic force microscope en anglais.  
2. Diffusion statique de la lumière. SLS pour Static light scattering en anglais.

# Étude bibliographique de l'assemblage et du désassemblage des capsides vides du CCMV

## 1.1 Des simulations de dynamique moléculaire...

Comme indiqué auparavant, l'étude expérimentale de la dynamique d'assemblage des capsides pleines est intrinsèquement complexe, du fait de la coexistence des protéines et du génome. Une approche pour surmonter cette difficulté est la réalisation de simulations numériques. Ainsi, en 2014, Perlmutter et al. ont réalisé des simulations de dynamique moléculaire ayant mis en lumière deux mécanismes d'assemblage qui dépendent des énergies d'interaction entre les différents composants (Fig. 4.1). Sur la figure 4.1a, un morceau de capside se forme sur le génome



**Figure 4.1:** Trajectoires correspondant aux deux mécanismes de croissance de la capside mis en lumière par simulations de dynamique moléculaire. (Adapté de [71])

puis croît par ajout séquentiel des protéines de capside. L'assemblage des dimères et l'empaquetage du génome se produisent de manière synchrone. Cette voie, se produisant principalement lors d'interactions attractives fortes entre les protéines, est souvent appelée mécanisme de nucléation-croissance. Sur la figure 4.1b, les



dimères s'adsorbent rapidement sur le génome et un complexe amorphe en résulte. Celui-ci relaxe ensuite en une capsidie pleine et ordonnée. Cette voie, se produisant préférentiellement pour des interactions attractives fortes entre les protéines et le génome, est appelée mécanisme *en masse*.

## 1.2 ...aux observations expérimentales.

Si l'étude expérimentale de la dynamique d'assemblage des capsides pleines est complexe, celle-ci n'est pas impossible. Ainsi, dès les années 80, le mécanisme de nucléation-croissance avait pu être observé lors de l'assemblage *in vitro* d'un virus du navet, le TCV<sup>1</sup>[72]. Dans ce cas, la formation d'un noyau de nucléation spécifique avait pu être identifié[73]. En revanche, dans le cas de l'assemblage de la capsidie du CCMV autour d'ARN, aucun site spécifique ni événement de nucléation n'ont pu être identifiés.

### 1.2.1 Assemblage via le mécanisme *en masse* ?

Ainsi, l'auto-assemblage de la capsidie pleine du CCMV a plutôt été décrit comme un processus *en masse*. C'est notamment le cas de l'équipe de Gelbart<sup>2</sup>, dans un article datant de 2014[74]. Dans celui-ci, les expériences d'assemblage ont été réalisées en mélangeant un ARN endogène aux dimères de protéines de la capsidie du CCMV. L'ARN en question était le RNA1 du virus de la mosaïque du brome (BMV<sup>3</sup>). Celui-ci a été choisi, au lieu de l'ARN endogène du CCMV, afin de minimiser les éventuelles interactions spécifiques entre ARN et protéines. Dans cet article, les expériences consistaient à dialyser le mélange réactionnel contre des solutions tampons à des forces ioniques ou des *pH* différents. Les protéines ayant été marquées préalablement à l'aide d'un fluorophore<sup>4</sup>, les mélanges ont été analysés par gradient de densité de sucrose couplé à des mesures d'intensités de fluorescence. Les résultats ont ensuite été confirmés par microscopie électronique.

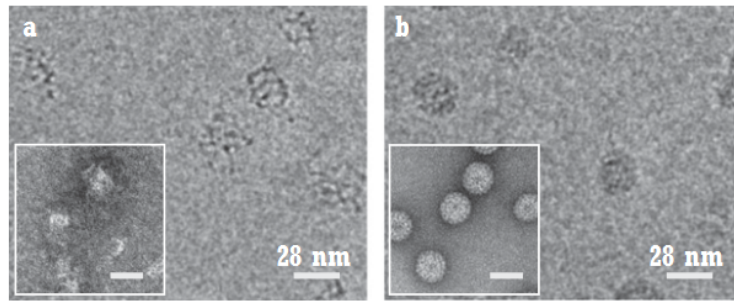
---

1. Turnip Crinkle Virus, en anglais

2. Département de chimie et biochimie de l'Université de Californie Los Angeles (UCLA)

3. Brome Mosaic Virus, en anglais

4. AF488.



**Figure 4.2:** Images Cryo-EM des réactions d'assemblage pour différents historiques de tampons (Encarts : images en coloration négative). (a) Images de complexes protéines-ARN désordonnés. Mélange dans un tampon (pH 7,2 - force ionique 1,0 M), puis dialyse contre un tampon (7,2 - 0,1 M) à 4 C, sur la nuit. (b) VLP (Virus Like Particule, en anglais) sphériques formées en dialysant les complexes en (a) contre un tampon à  $pH$  acide (4,5 - 0,1 M). (Adapté de [74])

Ainsi, sur la figure 4.2, sont présentées les images obtenues, par Cryo-EM et coloration négative, d'un mélange d'ARN et de dimères. Sur celles-ci, nous pouvons observer la formation de complexes désordonnés après une diminution de la salinité (a), suivie de l'organisation de la capsidie après diminution du  $pH$  (b). Ces observations, associées aux résultats obtenus par ultracentrifugation, motivent les auteurs de l'étude à parler d'un assemblage en deux étapes via un processus *en masse*. Il est à noter que ces résultats, obtenus avec un ARN exogène et la protéine native de la capsidie du CCMV, sont corroborés par deux autres articles de la même équipe [75][76].

## 1.2.2 Stoechiométrie requise pour l'assemblage

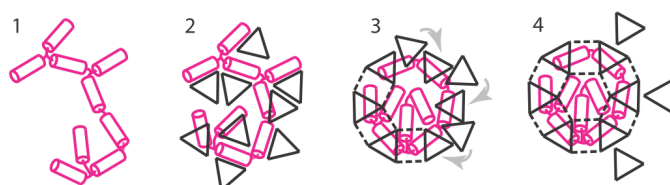
### 1.2.2.1 L'assemblage commence avec un excès de protéines...

Les deux articles, cités précédemment, ont été réalisés dans la continuité de l'article de 2012 de Cadena-Nava et al.[77]. Dans ces articles, en plus des techniques citées précédemment, les auteurs ont utilisé la mobilité électrophorétique et la spectroscopie de corrélation de fluorescence<sup>5</sup> pour étudier l'assemblage des protéines de la capsidie du CCMV autour d'ARN dont les tailles étaient variables. Dans le cas de l'article de 2012 et de 2016, les tailles d'ARN variaient de 140 à 12 000 nucléotides (nt) tandis que seul l'ARN1 du BMV (3 200 nt) était utilisé pour l'article de 2014.

5. Technique qui repose sur l'analyse des corrélations des fluctuations de l'intensité de la fluorescence. Une des applications possibles est l'analyse des fluctuations de la concentration de particules fluorescentes (souvent des molécules marquées) en solution.

De plus, dans le cas des publications de 2014 et 2016, en plus de la protéine native, des protéines mutantes portant différentes quantités de charge cationique ont été utilisées, alors que seule la protéine native a été utilisée pour l'article de 2012.

L'ensemble des travaux réalisés par l'équipe de Gelbart indiquait qu'un excédent de dimères de protéines était lié à l'ARN et, ceci, quel que soit le rapport en masse protéines-ARN. Cette surstoéchiométrie a été expliquée par les auteurs par la nécessité de neutraliser les charges négatives portées par le génome à l'aide des charges positives portées par le bras flexible de la protéine. Ainsi, la formation d'un complexe électriquement neutre serait une condition essentielle à l'auto-assemblage de la capsid pleine<sup>6</sup>. De ce fait, l'équipe de Gelbart a proposé un chemin possible d'assemblage, mis en évidence sur la figure 4.3.



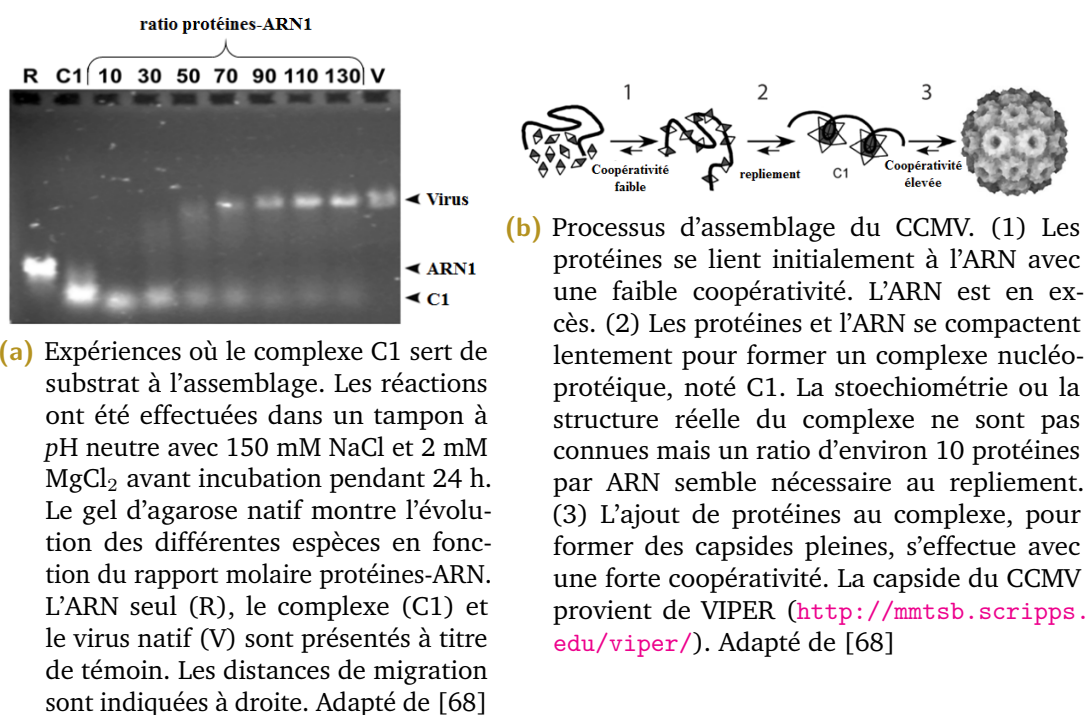
**Figure 4.3:** L'ARN et les dimères de protéines sont symbolisés par les cylindres articulés et les triangles, respectivement. Chaque cylindre représente une quantité de charges négatives égale à celles, positives, apportées par les bras de chaque protéine. 2. Les dimères saturent les sites de liaison de l'ARN lors de la première étape d'assemblage, ce qui compacte légèrement l'ARN. 3. Lors de l'acidification, les interactions latérales entre les dimères augmentent et génèrent la nucléation de la capsid (ligne en pointillée) ainsi que la réorganisation de l'ARN. 4. La croissance de la capsid déplace l'excès de protéines liées vers l'extérieur de la capsid. Adapté de [75].

Ainsi, l'empaquetage *in vitro* serait initié par la formation d'un complexe protéines-ARN désordonné. Dans le cas où la taille de l'ARN serait comprise entre 2 et 3,6 kb, celui-ci serait constitué d'une seule molécule d'ARN. Quelle que soit la taille de l'ARN, celui-ci serait décoré d'un excès de dimères de protéines, chargés positivement, compensant la charge négative du génome. A l'issue, la formation de la capsid nécessiterait une réorganisation structurelle avec un transfert de l'excès de dimères vers l'extérieur de la capsid.

6. Depuis, une erreur systématique a été découverte dans le calcul des concentrations en protéines utilisées pour les expériences présentées dans l'article de référence. De ce fait, plutôt qu'une neutralisation complète de la charge, la formation du complexe nécessiterait une compensation, d'environ 70% de la charge anionique du génome par le bras N-terminal cationique de la protéine [78].

### 1.2.2.2 ...ou un excès d'ARN ?

Cependant, dans un article datant de 2004, Johnson et al. ont observé un processus d'assemblage différent[68]. Dans cette publication, les auteurs ont étudié l'assemblage *in vitro* des protéines natives de la capsid du CCMV, soit autour de l'ARN1 du CCMV, soit autour de l'ARN Xef du kit de transcription MEGAScript T7. Leurs expériences, réalisées pour plusieurs rapports molaires et différents temps d'incubation, ont été analysées par mobilité électrophorétique et microscopie électronique. Un exemple des données obtenues est présenté ci-dessous (Fig. 4.4a).

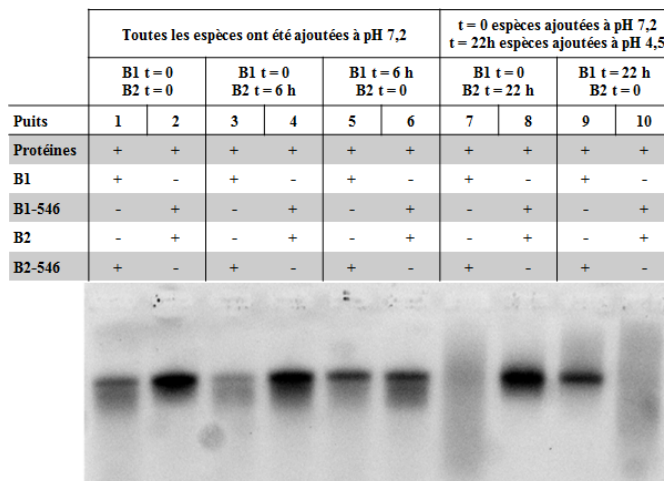


**Figure 4.4:** Expériences de d'assemblage réalisées par Johnson et al.[68].

Ainsi, les auteurs de l'étude ont pu observer indirectement la présence d'un complexe intermédiaire lors de la formation de la capsid, indiquant un processus d'assemblage en plusieurs étapes résumé sur la figure 4.4b. Dans le modèle proposé, à l'état initial une quantité limitée de protéines se lient à un excès d'ARN puis se replie pour former le complexe C1, substrat à un assemblage hautement coopératif qui donnera la capsid pleine. Pour les auteurs, ce processus est cohérent avec l'idée d'une concentration limitée en protéines, au début d'une infection, puisqu'il minimise les risques d'encapsidation d'ARN de la cellule hôte.

### 1.2.3 Liaison réversible

Cette conclusion de Johnson et al. semble plus cohérente avec les résultats obtenus, en 2012, par Comas-Garcia et al. En effet, le mécanisme proposé par l'équipe de Gelbart implique une irréversibilité des liaisons entre les protéines et l'ARN. Or, ce n'est pas ce qui a été observé par Comas-Garcia et al. lors de leurs expériences de compétition, réalisées pour comparer qualitativement et quantitativement les efficacités d'emballage d'une grande variété d'ARN par les protéines de la capsid du CCMV[79]. Les analyses qualitatives étaient simplement destinées à identifier quelle espèce d'ARN était emballé. Ceci a été réalisé en analysant le mélange réactionnel par mobilité électrophorétique (EMA) avant et après traitement par RNase<sup>7</sup>. Les analyses quantitatives ont été effectuées en marquant les ARNs avec un marqueur fluorescent (AF-546) lors de la transcription *in vitro*. Pour ces expériences, l'ARN de référence choisi est le RNA1, noté B1, du virus de la mosaïque du brome<sup>8</sup>.



**Figure 4.5:** Gel, obtenu par mobilité électrophorétique, des expériences de compétition entre les ARN B1 et B2 pour différentes conditions de tampon et à différents temps et ordre de mélange. Sur ce gel, le signal provient uniquement de l'ARN marqué par AF-546. Adapté de [79]

Ainsi, en effectuant l'assemblage *in vitro* en deux étapes, avec différents ordres de mélange, les auteurs ont constaté qu'à pH neutre, les ARN sont liés de manière

7. Une ribonucléase (ou RNase) est une nucléase qui catalyse la dégradation de l'ARN en éléments plus petits.

8. Celui-ci a été choisi, comme précédemment, afin de minimiser les éventuelles interactions spécifiques entre ARN et protéines.

réversible aux protéines, alors que la formation de nucléocapsides matures, résistantes à la RNase, se produit de manière irréversible en milieu acide. En effet, comme nous pouvons le voir sur la figure 4.5, à  $pH$  7,2, l'ordre d'ajout des ARN, ou l'intervalle de temps avant l'ajout du second ARN, influence peu la quantité finale de chacun des différents ARN encapsidés (6 premiers puits de la figure 4.5). *A contrario*, l'ordre d'ajout des ARN, à  $pH$  acide, influence fortement la quantité finale de chacun des différents ARN encapsidés (4 dernier puits de la figure 4.5). Les auteurs en concluent donc, qu'à  $pH$  neutre, la liaison entre les ARN et les protéines est réversible ce qui n'est pas le cas à  $pH$  acide.

L'ensemble des éléments de cette étude amène les auteurs à penser que l'assemblage s'effectuerait en deux étapes, avec dans un premier temps la formation de complexes ARN-protéines, à  $pH$  neutre. La formation de celui-ci se produirait conformément au scénario *en masse* dans lequel des interactions fortes, non spécifiques, conduisent à la saturation des molécules d'ARN par les protéines de capsid.

### 1.3 Conclusion et discussion

Ce chapitre met en évidence le fait que la dynamique d'auto-assemblage des capsides pleines du CCMV est encore méconnue et débattue. Des approches expérimentales et théoriques ont été utilisées afin d'apporter des éléments de réponses à ces questions.

Ainsi, deux mécanismes d'assemblage ont pu émerger par simulations de dynamique moléculaire : le mécanisme de nucléation-croissance et le mécanisme *en masse*. Si le premier a pu être observé chez certains virus, ce n'est pas le cas pour le CCMV. En effet, les études présentées dans ce chapitre semblent plutôt indiquer un mécanisme *en masse*. Cependant, l'ensemble de ces études présente une limite de taille : celles-ci ont toutes été réalisées à l'équilibre. De ce fait, un certain nombre de questions restent en suspens. Existe-t-il un noyau de nucléation, labile, n'ayant pas pu être observé, du fait de la faible résolution temporelle des études précédentes ? Si tel est le cas, le mécanisme de nucléation-croissance ne peut être exclu. De plus, dans le cas d'un mécanisme *en masse*, le nombre de dimères s'adsorbant sur le génome, afin de former un complexe amorphe, ne fait

pas consensus. En effet, l'équipe de Gelbart propose un assemblage qui commence avec un excès de protéines[76][75][77] tandis que celle de Zlotnick parle plutôt d'un excès d'ARN[68]. Enfin, la nature, réversible ou irréversible, des liaisons entre les protéines et l'ARN, lors de l'assemblage, est également sujette à discussion.

# Analyse par Thermo-fluorescence du CCMV

Le Criblage de Thermostabilité des Protéines (CTP) est une technique permettant l'étude de la stabilité de molécules biologiques. Elle permet de quantifier les changements de température de dénaturation thermique des protéines dans différentes conditions (*pH*, salinité, mutations, etc). Autrement dit, cette technique permet de mesurer la température à laquelle une protéine, ou une capsid, passe d'une structure ordonnée à désordonnée. Dans un premier temps, ce chapitre a pour objectif de décrire succinctement le principe du Criblage de Thermostabilité des Protéines par Fluorescence (CTPF) ainsi que les modèles utilisés pour analyser les données obtenues par cette technique. Dans un second temps, cette approche nous permet d'estimer expérimentalement la température de dénaturation des capsides vides et pleines du CCMV. Celle-ci étant directement liée aux interactions entre les protéines, nous pouvons estimer leurs énergies d'interaction. La comparaison des deux nous permet également de discuter des énergies d'interactions entre les protéines et le génome.

## 2.1 Le criblage de thermostabilité des protéines

### 2.1.1 Avantages et inconvénients

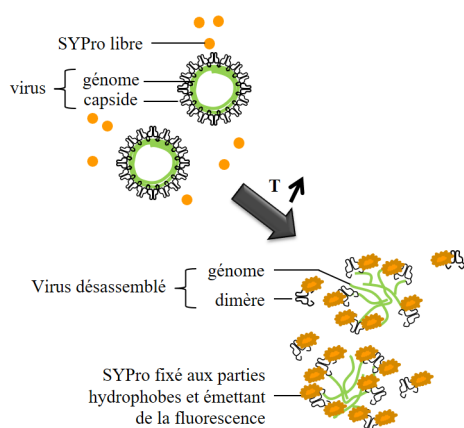
Traditionnellement, la stabilité des protéines a été étudiée par des tests biochimiques, dichroïsme circulaire ou calorimétrie à balayage différentiel. Les tests biochimiques nécessitent une activité catalytique de la protéine étudiée, la seconde et la troisième consomment de grandes quantités de protéines. Les criblages par fluorescence ont permis non seulement des essais de thermostabilité à haut débit, mais également de répondre aux limitations des précédentes techniques. Malgré ces plus-values, la technique possède néanmoins ses propres limitations :



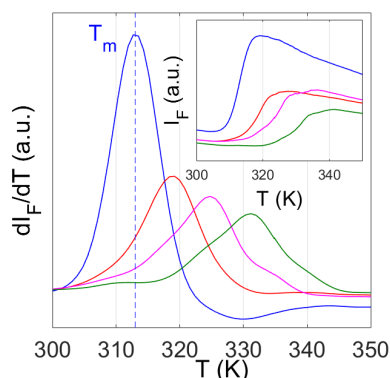
- Si l'on effectue des tests en fonction du  $pH$ , il convient de tenir compte du fait que le  $pK_a$  de chaque molécule tampon varie avec la température[80].
- Si l'on effectue des tests sur une espèce chargée, les effets des contre-ions doivent être pris en compte.

## 2.1.2 Le principe du CTPF

Le principe de la technique, décrit pour la première fois par Semisotnov et al. en 1991, est relativement simple[81]. A la solution étudiée est ajouté un fluorophore. Dans notre cas, c'est un agent intercalant, le SYPro, qui possède une affinité avec les zones hydrophobes des protéines. A l'état initial, les molécules de SYPro sont libres et émettent peu de fluorescence (cf. Fig. 4.6). Lors des essais par CTPF, la température de la solution est augmentée progressivement avec pour effet de dissocier les capsides de virus. Les parties hydrophobes des protéines sont alors accessibles et les molécules de SYPro se fixent dessus par un changement structural qui génère une excitation. Le retour à l'état fondamental de la protéine génère une



**Figure 4.6:** Schéma de principe du criblage de thermo-stabilité par fluorescence.



**Figure 4.7:** Courbes dérivées des intensités de fluorescence en fonction de la température à  $pH$  4,8. Les courbes bleue, rouge, rose et verte correspondent à 0,1 M, 0,2 M, 0,3 M et 0,5 M NaCl, respectivement. La température de dénaturation à 0,1 M est indiquée par la ligne en pointillés bleus. L'encart montre les intensités mesurées.

augmentation de la fluorescence. Celle-ci croît proportionnellement au nombre de zones hydrophobes accessibles (Fig. 4.6). L'intensité de fluorescence est mesurée

en fonction de la température et, la température de demi-dénaturation<sup>1</sup>,  $T_m$ , est choisie au maximum de la dérivée de l'intensité de fluorescence en fonction de la température (Fig. 4.7). Ce choix est courant dans les tests de thermostabilité par fluorescence car il ne nécessite pas d'hypothèse sur les mécanismes précis de dissociation.

## 2.2 Modèle de champ moyen associé

La température de dénaturation, qui permet de caractériser la transition entre les capsides assemblées sous forme icosaédrique et les protéines libres en solution, peut être décrite analytiquement par une théorie de champ moyen[18] :

$$\begin{aligned}
 F(N, T) = \frac{2\epsilon}{M} N^2 & \left\{ \begin{array}{l} \text{Terme décrivant l'énergie attractive de courte portée entre} \\ \text{les sous-unités} \end{array} \right. \\
 + \frac{Z^2 e^2 \kappa^{-1}}{16\pi \epsilon_0 \epsilon_r R^2} N^2 & \left\{ \begin{array}{l} \text{Terme décrivant la répulsion électrostatique dans le} \\ \text{cas d'un fort écrantage ionique[82]} \end{array} \right. \\
 - \mu_0 N & \left\{ \begin{array}{l} \text{Terme décrivant les échanges de sous-unités avec le réservoir} \\ \text{+ } k_B T M [\rho \ln \rho + (1 - \rho) \ln((1 - \rho))] \end{array} \right\} \left\{ \begin{array}{l} \text{Terme décrivant l'entropie de mé-} \\ \text{lange des sites libres et occupés} \end{array} \right. \\
 & \hspace{15em} (2.1)
 \end{aligned}$$

où

- $M$  est le nombre de dimères composant la capside,
- $N$  le nombre de dimères présent dans la capside à une température donnée  $T$ ,
- $\epsilon$  le potentiel de paire appliqué aux proches voisins,
- $Z$  est la charge effective portée par chaque dimère de protéines,
- $e$  la charge électrique,
- $\epsilon_0$  la permittivité du vide,
- $\epsilon_r$  la constante diélectrique (80 pour l'eau),
- $R$  le rayon de la capside (en nm)
- $\kappa^{-1}$  la longueur de Debye telle que :  $\kappa^{-1} = \sqrt{\epsilon_0 \epsilon_r k_B T / 2e^2 c_s}$

1. Il est à noter que, par abus de langage, celle-ci est souvent appelée température de dénaturation.

- $c_s$  la concentration en sel monovalent
- $\mu_0$  le potentiel chimique tel que :  $\mu_0(T) = k_B T \ln(\rho_0/1 - \rho_0)$ ,
- $\rho_0$  la densité de dimères dans le réservoir,
- $\rho$  la densité de dimères dans la capsidie tel que :  $\rho = N/M$

A l'équilibre, la dérivée première par rapport à  $N$  de l'équation 2.1 est égale à zéro. Une équation d'état relie alors  $\rho$  et  $T$ . Cette dérivée première de l'énergie libre par rapport à sa variable intensive  $N$  est discontinue ce qui est caractéristique d'une transition de phase du premier ordre. Cette propriété indique la possibilité d'une coexistence de phases, analogue à la transition liquide-gaz, à une température  $T_m$  définie comme la température de dénaturation. L'énergie libre est alors une fonction symétrique telle que  $F(\rho, T_m) = F(1 - \rho, T_m)$ . Dans l'hypothèse où les variables de l'équation 2.1 ne dépendent pas de la température, l'équation peut être ré-écrite sous la forme :

$$T_m = (2\epsilon + \frac{Z^2 e^2 M \kappa^{-1}}{16\pi\epsilon_0\epsilon_r R^2}) / k_B \ln(\frac{\rho_0}{1 - \rho_0}) \quad (2.2)$$

L'équation 2.2 met alors en évidence la dépendance de la température de dénaturation aux interactions des dimères entre eux via  $\epsilon$ ,  $Z$ ,  $\kappa$ .

Ce modèle de champ moyen qui relie la température de dénaturation aux énergies attractives, charges effectives et potentiel chimique des dimères, couplé au criblage de thermostabilité des protéines par fluorescence, donne accès aux paramètres d'interaction pour les capsides virales vides et chargées dans diverses conditions ioniques.

## 2.3 Estimation des températures de dénaturation

Dans cette section, nous nous intéresserons aux interactions entre les protéines et, entre les protéines et le génome, à travers la dissociation thermique des capsides. Les expériences ont été réalisées, à l'ICSN<sup>2</sup> par Eric Jacquet et Naïma Nhiri, sur le CCMV, présenté précédemment.

---

2. Institut de chimie des substances naturelles

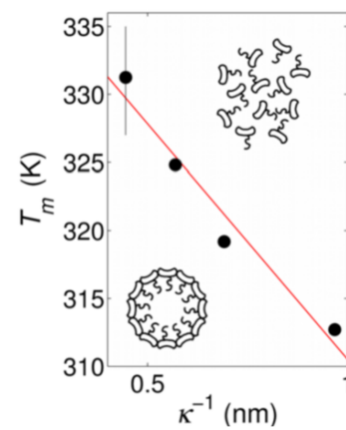
## 2.3.1 Cas des capsides vides

### 2.3.1.1 Effet de la salinité

La figure 4.7 présentée dans la section 2.1.2 montre les données de fluorescence obtenues avec des capsides vides du CCMV à pH 4,8 et une salinité variable. Les températures de dénaturation,  $T_m$ , mesurées ont été reportées sur la figure 4.8 en fonction de la longueur de Debye,  $\kappa^{-1}$ .

Dans un premier temps, nous pouvons remarquer que la température de dénaturation augmente avec la salinité, i.e.  $T_m$  diminue lorsque  $\kappa^{-1}$  augmente. Autrement dit, la cohésion entre les dimères de protéines est meilleure pour les solutions les plus salées. Cette variation, fonction de la salinité, vient de l'écrantage ionique. En effet, moins il y a de sel dans la solution, moins les groupes carboxyles chargés négativement sont écrantés. La force de répulsion électrostatique est donc plus importante. Malgré l'attraction hydrophobe, les dimères de protéines vont alors se repousser et déstabiliser la capsid. Ensuite, comme attendu par l'équation 2.1, nous observons une relation linéaire entre  $T_m$  et  $\kappa^{-1}$ . En ajustant les données à l'aide de cette équation (droite en rouge sur la figure 4.8), nous pouvons estimer une charge effective  $|Z| = 3,4e$  et une énergie d'interaction entre les dimères de protéines  $\epsilon = -4,6 k_B T_0$ , où  $k_B$  est la constante de Boltzmann et  $T_0$  la température ambiante (298 K).

Concernant la valeur estimée d'énergie d'interaction des dimères de protéines entre eux, celle-ci est relativement proche de la valeur utilisée par Johnson et al. dans leurs travaux et située à  $-6,3 k_B T_0$  [68]. Cependant, cette valeur a été choisie *ad hoc* et des valeurs similaires auraient tout autant pu correspondre aux données. Dans le cas de la charge effective mesurée, Vega-Acosta et al. ont calculé une



**Figure 4.8:** Température de dénaturation des capsides vides en fonction de la longueur de Debye, à pH 4,8. Les concentrations en sel sont identiques à celles de la figure 4.7.

charge d'environ  $-1,5e$ , ce qui est deux fois moins que ce que nous obtenons[83]. Cependant, ce résultat, obtenu par mobilité électrophorétique, correspond à la charge des résidus exposés en surface pour les virions complets. Or, comme les auteurs l'ont souligné, il est difficile d'établir un lien entre la mobilité et la charge.

### 2.3.1.2 Effet du pH

Nous avons également testé l'évolution de la température de dénaturation des capsides vides en fonction du pH. Les mesures ont été effectuées pour une concentration en chlorure de sodium de 0,5 M. Les résultats sont présentés sur la figure 4.9. En fixant la valeur de l'énergie d'interaction, telle que déterminée précédemment, nous avons pu calculer une valeur de la charge effective en fonction du pH (encart Fig. 4.9).

Sur cette figure, nous pouvons d'abord observer une valeur minimale autour de 3e pour un pH de 4,9. Pour des valeurs de pH comprises entre cette dernière et 6,5, la charge augmente, en raison de la déprotonation des groupements carboxyles portés par les protéines. Comme nous avons pu le voir dans la section 2.1.3 du premier chapitre, pour un pH et une salinité supérieurs à 6,5 et 0,3 M, respectivement, l'organisation sous forme de dimères est la plus stable. De ce fait, au-delà d'un pH à 6,5, la fraction de capsides en solution diminue drastiquement et notre modèle n'est plus valide. Les températures de dénaturation obtenues proviennent probablement de la dissociation de petits oligomères ou de dimères. Nous pouvons ensuite remarquer que la charge a un minimum non-nul. Ceci provient des termes multipolaires du développement limité utilisé pour décrire l'énergie libre de répulsion électrostatique. En effet, dans une publication datant de 2007, Siber et Podgornik ont proposé un modèle complexe de la capside tenant compte de la répartition des charges négatives et positives au sein de celle-ci[82]. Dans celui-ci, la capside est décrite

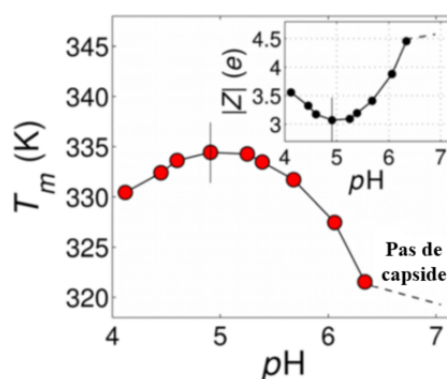


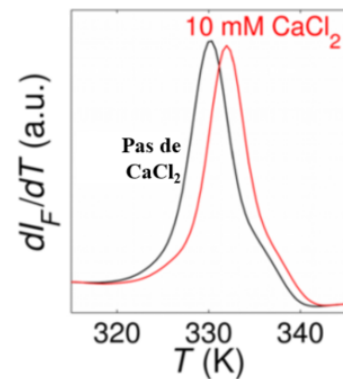
Figure 4.9: Température de dénaturation des capsides vides en fonction du pH pour une salinité à 0,5 M.

comme deux coques concentriques portant les charges  $Z_1$  et  $Z_2$  et séparées par un milieu diélectrique mince. A partir de cette description, les auteurs ont démontré que l'énergie libre conservait la même forme que l'équation 2.1, présentée précédemment, avec  $Z^2 = Z_1^2 + Z_2^2$ . Dans la mesure où les charges  $Z_1$  et  $Z_2$  ne s'annulent pas dans les mêmes conditions, et donc en même temps,  $Z$  est toujours non-nul.

### 2.3.1.3 Effet des ions divalents

Nous avons enfin testé l'effet des cations divalents  $\text{Ca}^{2+}$  sur la température de dénaturation des capsides vides. Les mesures ont été réalisées avec 10 mM de  $\text{CaCl}_2$  et la comparaison, avec et sans calcium est présentée sur la figure 4.10.

Sur celle-ci, nous pouvons observer que le maximum des dérivées de l'intensité de fluorescence en fonction de la température en présence ou en l'absence de  $\text{CaCl}_2$  sont différents. En effet, celui-ci se déplace vers une plus grande valeur de température de dénaturation en présence du calcium ( $\Delta T_m = 2 \text{ K}$ ). Cette augmentation génère une modification de la charge effective qui diminue de  $0,23e$ . Ces résultats indiquent une stabilisation des capsides vides par capture de cations divalents, ce qui est cohérent avec ce qui a pu être décrit précédemment pour les virions[11]. Ce modèle de champ moyen permet de quantifier cet effet via la charge effective. En effet, nous avons pu estimer qu'un dimère sur huit, environ, capture un ion  $\text{Ca}^{2+}$ .



**Figure 4.10:** Courbes dérivées de l'intensité de fluorescence en fonction de la température à  $\text{pH } 4,8$  et  $0,5 \text{ M NaCl}$ , en présence ou en l'absence de  $\text{CaCl}_2$  (courbes rouge et noire, respectivement).

## 2.3.2 Cas des capsides pleines

Le même type d'expérience, puis d'analyse, peut être effectué sur les capsides pleines. Cependant, dans ce cas, le potentiel chimique ne découle plus uniquement de la densité en dimères de protéines.

### 2.3.2.1 Adaptation du modèle de champ moyen

Dans un article datant de 2016, Bruinsma et al. décrivent l'assemblage de petits virus à ARN simple brin comme une transition de phase bidimensionnelle de type pelote-globule. Lors de cette transition, lorsque la pelote protéines-ARN entre en phase condensée, celle-ci n'est pas uniforme et s'apparente à un agrégat. Au sein de celui-ci, il y a une ségrégation des protéines vers la surface. Les auteurs expliquent cette séparation par la tension superficielle  $\gamma_0$ . En effet, le transfert de protéines vers la surface des globules protéines-ARN permettrait une diminution de l'énergie interfaciale libre de  $\gamma_0 D^2$ , que les auteurs ont estimé à  $k_B T (D/l)^2$ . Dans cette estimation,  $D^2$  correspond à la zone exclue des dimères à la surface de la capsid et  $l$  à la longueur de corrélation du globule protéines-ARN. A l'équilibre thermodynamique, la tension superficielle peut être reliée au potentiel chimique par l'isotherme de Gibbs<sup>3</sup>, ce qui permet de décrire :  $\mu_0 = -k_B T (D/l)^2$ . Le terme  $\ln(\rho_0/1 - \rho_0)$ , de l'équation 2.1 du modèle de champ moyen, peut alors être remplacé par  $-(D/l)^2$  afin d'obtenir l'équation reliant la température de dénaturation aux variables  $\epsilon$ ,  $Z$  et  $\kappa$ , pour les capsides pleines.

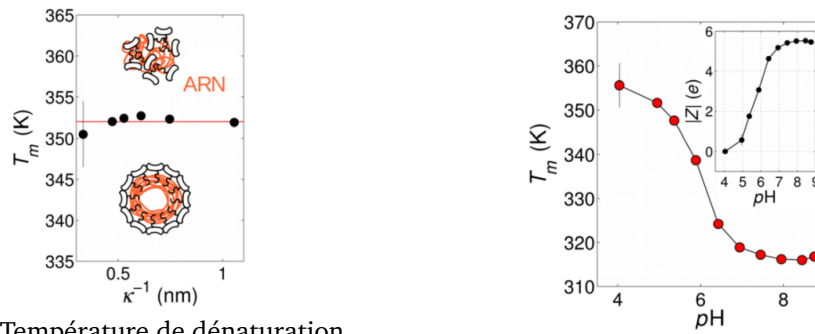
### 2.3.2.2 Influence du génome

La température de dénaturation a donc également été mesurée pour les virions natifs.

Dans un premier temps, nous avons fixé le pH à 4,8 et fait varier la salinité de 0 à 0,5 M NaCl par pas de 0,1 M. Les températures de dénaturation,  $T_m$ , mesurées ont été reportées sur la figure 4.11a en fonction de la longueur de Debye,  $\kappa^{-1}$ . Nous pouvons constater que, contrairement aux capsides vides, la température de dénaturation ne varie pas avec la longueur de Debye. En effet, celle-ci reste constante, autour de 352 K ce qui indique que, pour la capsid pleine, la charge effective  $|Z|$  disparaît à ce pH. Il est à noter que dans ce cas, la charge  $|Z|$  inclut la contribution du génome, i.e. les charges positives des bras flexibles des dimères de protéines sont compensées par l'ARN chargé négativement. De plus, à ce pH,

---

3. Pour des systèmes constitués de plusieurs éléments, l'isotherme d'adsorption de Gibbs est une équation permettant de relier les changements de concentration d'un élément, en contact avec une surface, avec les changements de tension de surface. Pour un système binaire, l'équation est la suivante :  $-d\gamma = \Gamma_1 d\mu_1 + \Gamma_2 d\mu_2$ , où  $\Gamma_i$  correspond à l'excès de surface des éléments  $i$ .



(a) Température de dénaturation des capsides vides en fonction de la longueur de Debye, à  $pH$  4,8.

(b) Température de dénaturation en fonction du  $pH$  pour une salinité à 0,5 M.

**Figure 4.11:** Tests de dissociation thermique pour les virions natifs.

les groupes carboxyles des dimères sont fortement protonés et peu chargés. Nous pouvons ensuite remarquer que la température de dénaturation des capsides pleines est plus élevée que celle des capsides vides (jusqu'à 30 K). Cette différence, peu étonnante, vient de la stabilisation de la capside, via l'attraction électrostatique entre les bras flexibles des dimères et le génome.

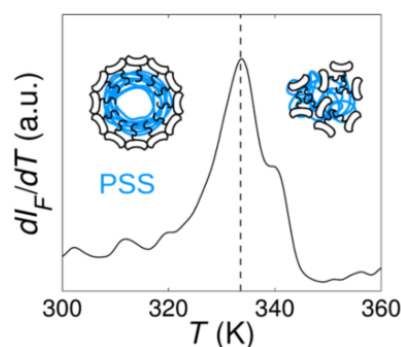
Dans un second temps, la température de dénaturation a été mesurée à une salinité fixée (0,5 M) et un  $pH$  variable. Les  $T_m$  mesurées ont été reportées sur la figure 4.11b. Comme précédemment, nous avons calculé la valeur absolue de la charge effective  $|Z|$ . Les valeurs obtenues sont visibles dans l'encart de la figure 4.11b. Il est à noter que le diamètre des virions du CCMV hausse de 28 à 32 nm lorsque le  $pH$  augmente. Ceci vient du fait que lorsque le  $pH$  augmente, les groupements carboxyles à la surface des dimères se chargent négativement ce qui amplifie la répulsion électrostatique entre les dimères. De ce fait, nous avons utilisé différentes valeurs de rayons pour remplacer  $R$  dans l'équation 2.1. Les valeurs utilisées correspondent à des rayons mesurés à divers  $pH$  par microscopie à force atomique[84]. Sur la figure 4.11b, nous pouvons observer que la température de dénaturation est constante, autour de 352 K, pour des  $pH$  inférieurs à 5, diminue rapidement au-delà, puis, se stabilise autour de 320 K pour des  $pH$  supérieurs à 7,5. Ces observations mettent en évidence l'influence du  $pH$  sur les interactions au sein de la capside. En effet, les groupes carboxyles des protéines sont plus protonés à  $pH$  4,8 qu'à  $pH$  7,5. De fait, lorsque le  $pH$  devient suffisamment basique, la répulsion électrostatique entre les dimères augmente ce qui déstabilise la structure de la capside et favorise le désassemblage des dimères. Dans l'encart de la figure, nous pouvons remarquer que l'évolution de la charge effective  $|Z|$  est cohérente avec



cette description : celle-ci est nulle pour des  $pH$  inférieurs à 5, augmente au-delà, puis, est constante autour de 5,5e pour ceux supérieurs à 7,5.

### 2.3.2.3 Influence d'un polyélectrolyte, le PSS

Dans la dernière expérience, nous avons empaqueté du poly(styrène sulfonate) de sodium (PSS) de 120 kDa au sein des capsides virales. Le PSS est un poly-électrolyte flexible, chargé négativement, dont la densité de charge linéique est d'environ  $-0,33e\text{\AA}^{-1}$  soit une valeur proche de celle de l'ARN natif. Pour cette expérience, la température de dénaturation a été mesurée à une salinité de 0,5 M et un  $pH$  de 4,8. L'évolution de la dérivée de l'intensité en fonction de la température est présentée sur la figure 4.12. Dans ce cas, nous observons un maximum de cette dérivée autour de 334 K, soit une température de dénaturation de 18 K de moins que pour les virions natifs et à peine plus que les capsides vides, dans les mêmes conditions. En dehors du fait que la longueur de corrélation du globule protéines-PSS,  $l$ , peut être plus faible que celle du globule protéines-ARN,  $|Z|$  est également susceptible d'être modifiée. En effet, il peut y avoir une surabondance de PSS empaqueté à l'intérieur de la capside[85]. Le PSS surchargeait alors beaucoup plus le bras flexible des protéines que l'ARN viral. De ce fait, nous pouvons nous attendre une valeur plus élevée de  $|Z|$ , ce qui est compatible avec la diminution observée de  $T_m$ .



**Figure 4.12:** Courbe dérivée de l'intensité de fluorescence en fonction de la température pour des capsides contenant du PSS.

## 2.4 Conclusion

En résumé, une série de mesures a été réalisée afin de déterminer la charge effective ainsi que l'énergie d'interaction au sein des capsides vides et pleines du CCMV.

Lors de nos expériences sur les capsides vides, à  $pH$  4,8 et une salinité variable, nous avons pu quantifier une énergie d'interaction  $\epsilon = -4,6k_B T_0$  qui est raisonnable et proche des valeurs expérimentales estimées précédemment. Dans le cas de la charge effective, nous avons pu estimer des valeurs absolues allant de 3 à  $4,5e$  pour des  $pH$  de 4,8 à 6,5 respectivement. Si  $3e$  correspond à la valeur minimale en présence d'ions monovalents, celle-ci peut diminuer en présence d'ions divalents. En effet, nous avons pu estimer qu'environ un dimère sur huit capture un ion  $Ca^{2+}$ , avec pour effet de stabiliser la capside.

Dans le cas du virion natif, nos expériences ont montré que la charge effective  $|Z|$  disparaissait à  $pH$  4,8, du fait de la contribution du génome. En effet, à ce  $pH$ , les charges positives des bras flexibles des dimères de protéines sont compensées par l'ARN chargé négativement et, les groupes carboxyles sont protonés. Nous avons ensuite pu observer que la température de dénaturation diminue lorsque le  $pH$  augmente. Dans ce cas, c'est la déprotonation des groupes carboxyles des dimères qui génère une augmentation de la répulsion électrostatique avec, pour effet, de diminuer la cohésion dans la capside. Nous avons enfin testé l'influence de l'empaquetage du PSS, au sein de capsides virales, sur  $T_m$ . Dans ce cas, celle-ci était plus basse que pour le virion natif, ce qui pouvait s'expliquer par une surabondance de PSS en interaction avec le bras flexible.

Plus généralement, cette partie met en avant les nombreux paramètres pouvant altérer ou stabiliser les interactions au sein de la capside. La connaissance des interactions au sein de la capside du CCMV est fondamentale pour la compréhension de l'assemblage de ce virus.

# Dynamique d'assemblage de la capsid virale

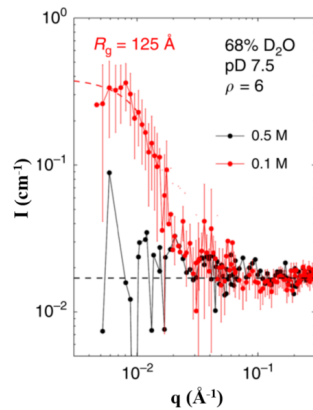
Comme nous avons pu le voir dans le chapitre précédent, les interactions entre les protéines de la capsid sont soumises à un équilibre délicat entre une attraction hydrophobe à courte portée et une répulsion électrostatique à longue portée. L'interaction entre les dimères de protéines et le génome est, quant à elle, dirigée par l'attraction électrostatique. Dans cette étude, nous étudions la dynamique d'auto-assemblage, hors équilibre, des capsides du CCMV, avec leur matériel génétique. Les expériences ont été réalisées par diffusion de rayon X aux petits angles résolue en temps (TR-SAXS) à l'ESRF sur la ligne ID02 et des expériences complémentaires ont été effectuées par diffusion de neutrons aux petits angles (SANS) au Laboratoire Léon Brillouin (LLB). En utilisant ces deux techniques, nous avons pu observer, de manière quantitative, la dynamique d'auto-assemblage de la capsid. De ce fait, je présenterai, dans un premier temps, la dynamique d'auto-assemblage des intermédiaires observés au cours de nos expériences. Je discuterai ensuite de la relaxation de ces intermédiaires en capsides. Je présenterai ensuite succinctement les résultats obtenus lors d'un assemblage, non plus avec le génome mais avec un polyélectrolyte. Enfin, je discuterai des facteurs possibles de l'assemblage *in vivo*.

## 3.1 Auto-assemblage des intermédiaires

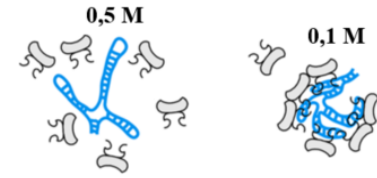
### 3.1.1 Estimation du nombre de sous-unités

Dans un premier temps, nous avons effectué des mesures statiques par diffusion de neutrons aux petits angles (SANS), d'un mélange de dimères et de génome dans 68% de D<sub>2</sub>O à pD 7,5. Dans cette fraction d'eau lourde, le contraste entre l'ARN et le tampon disparaît. De ce fait, seule l'intensité diffusée provenant des protéines est détectée[85]. Si le rapport en masse,  $\rho$ , des dimères de protéines avec le génome, dans un virion natif, est situé autour de 4, nous avons effectué nos

expériences pour  $\rho = 6$ , car celui-ci s'est avéré être le rapport minimal permettant d'assurer l'assemblage total autour du génome *in vitro*[79]<sup>1</sup>.



(a) Courbes SANS obtenue lors des mélanges de dimères et génome dans 68% de D<sub>2</sub>O à pD 7,5 et des forces ioniques variables. Les courbes noire et rouge correspondent à 0,5 M et 0,1 M NaCl, respectivement. La ligne pointillée noire indique le niveau de bruit tandis que la ligne pointillée rouge correspond à un ajustement dans la région du Guinier (voir section 1.5).



(b) Illustration schématique de la figure 4.13a aux deux forces ioniques. Les dimères et le génome sont représentés en gris clair et en bleu, respectivement.

**Figure 4.13:** Estimation du nombre de sous-unités dans les complexes nucléoprotéiques amorphes (NPCs).

Sur la figure 4.13a, nous montrons une comparaison des courbes obtenues pour deux forces ioniques : 0,5 et 0,1 M NaCl. Sur cette figure, nous pouvons observer que, si l'intensité diffusée est proche du niveau de bruit à une force ionique de 0,5 M (courbe noire), elle augmente d'un ordre de grandeur pour 0,1 M NaCl (courbe rouge). Cette augmentation peut être attribuée à la liaison des dimères de protéines sur le génome, qui forment ainsi des complexes nucléoprotéiques amorphes notés NPCs dans l'ensemble du manuscrit (Fig. 4.13b). En supposant que la solution n'est composée que de dimères libres et de complexes, ceux-ci étant constitués de  $N$  dimères liés au génome ( $N \in \mathbb{N}$ ), nous pouvons estimer le nombre de dimères par complexe à partir des intensités diffusées  $I_0 \equiv I(q = 0)$ . En effet, l'intensité diffusée  $I_0$  d'une solution contenant des NPCs est donnée par la relation suivante[54] :

$$I_0 = \Delta b_S^2 X_S + \sum_{N=0}^{+\infty} X_N (\Delta b_G + \Delta b_S N)^2 \quad (3.1)$$

1. Depuis, une erreur systématique a été découverte dans le calcul des concentrations en protéines utilisées pour les expériences présentées dans l'article de référence. De ce fait, le rapport en masse protéines/ARN, nécessaire pour empaqueter le génome, en est modifié : il passe de 6 à 4,3[78].

où  $\Delta b_S$  et  $\Delta b_G$  sont les contrastes de longueur de diffusion des dimères et du génome, respectivement.  $X_S$  est la concentration molaire en dimères de protéines libres et  $X_N$  la concentration molaire en complexes d'ordre  $N$ .

Après un traitement mathématique, présenté dans le chapitre 3.1.1, une équation quadratique peut être obtenue. Celle-ci relie le nombre moyen de dimères par complexe à l'intensité diffusée, tous deux à l'équilibre, via la concentration molaire totale en génome dans la solution, et l'intensité diffusée, produite par le mélange des dimères et du génome, avant la formation des complexes. Pour rappel, celle-ci s'écrit :

$$\langle N \rangle(t \rightarrow +\infty) = \sqrt{\frac{1}{4}\lambda^2 + \frac{I_0(t \rightarrow +\infty) - I_0^*}{\Delta b_S^2 c_G}} - \frac{1}{2}\lambda$$

Dans le cas des expériences SANS, dans 68% de D<sub>2</sub>O,  $\Delta b_G$  est nul, puisque le contraste entre l'ARN et le tampon disparaît. De plus,  $\Delta b_S$  est estimé à 0,0141 cm<sup>-0,5</sup>.μM<sup>-0,5</sup>. De ce fait, nous pouvons estimer qu'en moyenne, au maximum, chaque complexe contient  $\langle N \rangle_{up} = 75 \pm 31^2$  dimères de protéines. Il est à noter que l'extraction de la valeur  $\langle N \rangle_{up}$  à partir des données de diffusion ne nécessite aucune connaissance préalable de la distribution de  $N$  et, par conséquent, cette quantité, indépendante d'un modèle, a été utilisée dans cette étude pour estimer la liaison des dimères sur le génome dans diverses conditions.

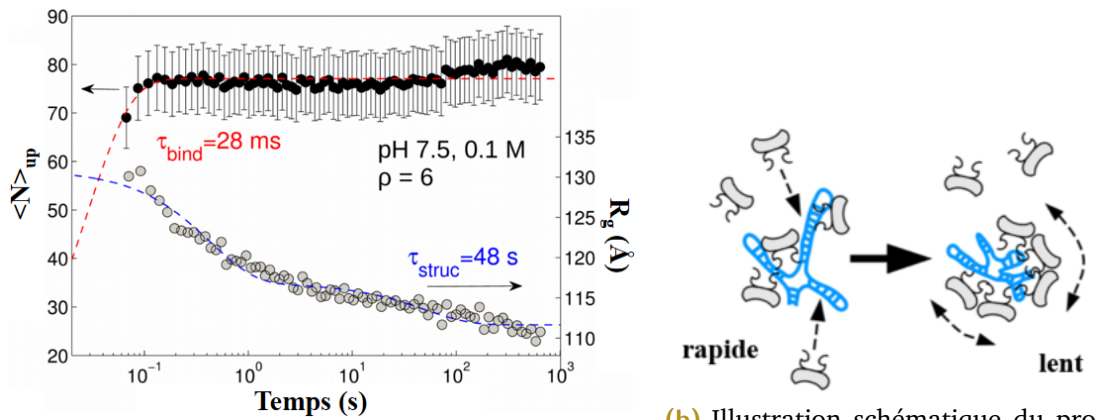
### 3.1.2 Formation de l'intermédiaire

Nous avons également réalisé des expériences par diffusion de rayon X aux petits angles résolue en temps (TR-SAXS) en mélangeant rapidement des dimères de protéines avec le génome, préalablement et séparément purifiés. L'assemblage a été réalisé pour une force ionique finale de 0,1 M et un rapport en masse  $\rho = 6$ .

La figure 4.14a montre l'évolution, en fonction du temps, de la moyenne supérieure du nombre de dimères par complexes (abscisse de gauche) et du rayon de giration (abscisse de droite) lors de assemblage des NPCs. Dans un premier temps, nous pouvons observer que les dimères de protéines se lient très rapidement au génome, puis, que le nombre moyen de dimères par complexes reste stable (autour de 77

---

2. Écart type à la moyenne



(a) Evolution de  $\langle N \rangle_{up}$  (cercles noirs) et  $R_g$  (cercles gris), après mélange des dimères libres et du génome, en fonction du temps. Les courbes rouge et bleue, en pointillées, montrent les ajustements permettant d'extraire les temps caractéristiques  $\tau_{bind}$  et  $\tau_{struc}$ , respectivement.

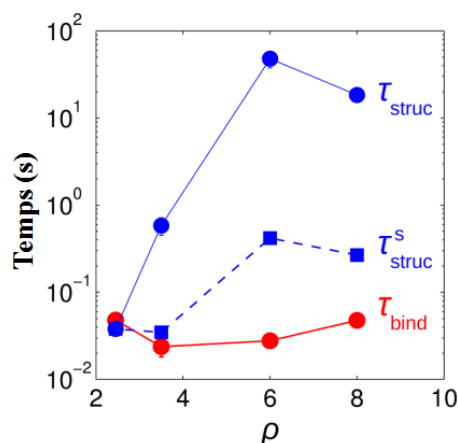
(b) Illustration schématique du processus d'assemblage *en masse* des NPC.

**Figure 4.14:** Dynamique d'assemblage du complexe nucléoprotéique amorphe.

dimères) pendant plus de 10 minutes. Nous avons ensuite réalisé un ajustement via une simple fonction exponentielle décroissante qui indique un temps de liaison au génome,  $\tau_{bind}$ , de 28 ms. Dans un second temps, nous pouvons remarquer que la variation de  $R_g$  se produit sur un temps plus long. En effet, en réalisant un ajustement à l'aide d'une double exponentielle décroissante, nous pouvons extraire deux temps de relaxation structurelle,  $\tau_{struc}^s$  et  $\tau_{struc}$ . C'est ce dernier qui nous intéresse et dure 48 s, ce qui est trois ordres de grandeur plus élevé que  $\tau_{bind}$ . La dynamique d'assemblage des complexes peut alors être décrite ainsi : la première phase semble résulter de l'adsorption rapide d'environ 75 dimères tandis que la deuxième phase semble correspondre à un compactage plus lent des NPC. Ces résultats suggèrent que les NPC ont été formés par la voie *en masse*, décrite au début de ce chapitre (Fig. 4.14b). Les dimères sont d'abord capturés par le génome en environ 100 ms, puis les NPC s'auto-organisent lentement sur plusieurs minutes.

Nous avons aussi réalisé la même expérience, à différentes valeurs de  $\rho$ . Pour ces expériences, nous avons également ajusté les courbes d'évolution, en fonction du temps, de  $\langle N \rangle_{up}$  et de  $R_g$  à l'aide d'une simple fonction exponentielle décroissante et d'une double exponentielle décroissante, respectivement. Nous obtenons les mêmes trois temps caractéristiques,  $\tau_{struc}$ ,  $\tau_{struc}^s$  et  $\tau_{bind}$ , qui correspondent aux temps de relaxation structurelle long et court ainsi qu'à celui de liaison au génome

(Fig. 4.15). Nous pouvons noter que le compactage du génome, correspondant à la phase rapide ( $\tau_{bind}$ ), peut ralentir lorsqu'un grand nombre des dimères se lie dessus, comme c'est le cas pour les valeurs élevées de  $\rho$ . De plus, comme nous pouvons le voir sur la figure 4.15,  $\tau_{struc}$  et  $\tau_{struc}^s$  augmentent également avec  $\rho$ . L'auto-organisation ultérieure des NPCs, i.e. la deuxième phase, peut être affectée en conséquence.



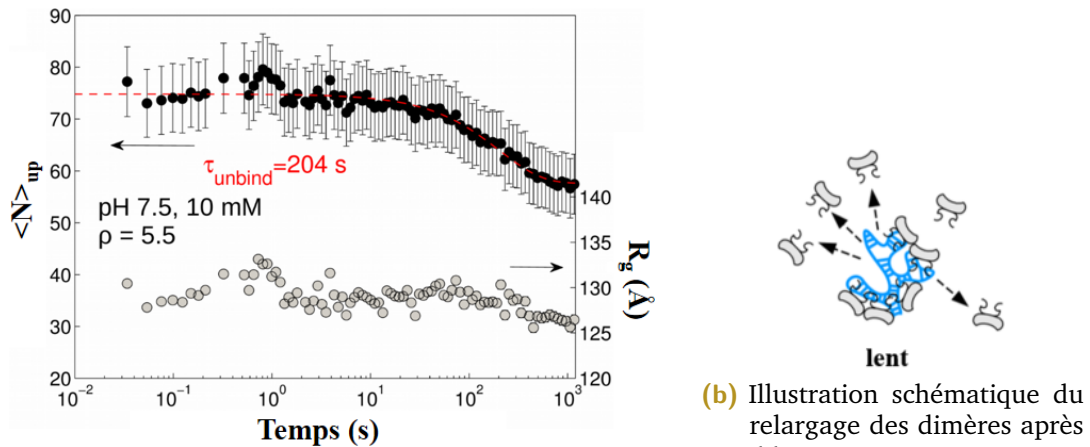
**Figure 4.15:** Temps caractéristiques de la dynamique d'assemblage des complexes nucléoprotéiques amorphes pour différentes valeurs de  $\rho$ .  $\tau_{bind}$ ,  $\tau_{struc}$  et  $\tau_{struc}^s$ , pour une solution tampon à pH 7,5 et une force ionique de 0,1 M, sont présentés en fonction du rapport massique  $\rho$  (protéines /ARN, en faveur d'un excès de protéines).

### 3.1.3 Énergie de liaison au sein de l'intermédiaire

En plus de déterminer le processus d'assemblage des complexes, nous avons également voulu tester la stabilité de la liaison des dimères avec les complexes. De ce fait, nous avons réalisé une expérience, lors de laquelle, les NPCs assemblés précédemment ont été dilués deux fois rapidement. La solution contenant les NPCs a une force ionique finale de 10 mM et un rapport en masse  $\rho = 5, 5$ .

La figure 4.14a montre l'évolution de la moyenne supérieure du nombre de dimères par complexe (ordonnée de gauche) et l'évolution du rayon de giration (ordonnée de droite) en fonction du temps, après dilution. Dans un premier temps, nous pouvons observer que  $\langle N \rangle_{up}$  diminue lentement. En effet, le nombre de dimères par NPC décroît de 75 à 58 pendant un temps de dissociation de  $\tau_{unbind} = 204$  s. Dans un second temps, nous pouvons remarquer que la variation de  $R_g$  n'est pas significative sur la période d'observation ( $\approx 1200$  s). Ceci peut être dû à des effets

antagonistes : les NPCs ont tendance à être moins compacts lors de la libération des dimères de protéines, mais dans le même temps, les dimères libérés ne contribuent plus au rayon de giration des NPCs. De ce fait,  $R_g$  reste relativement stable. Cette expérience démontre que la dilution est suffisante pour libérer des dimères des NPCs en quelques minutes, ce qui suggère une faible énergie de liaison entre les dimères de protéines et le génome.



(a) Evolution de  $\langle N \rangle_{up}$  (cercles noirs) et  $R_g$  (cercles gris), après une double dilution rapide d'un des mélanges de NPCs obtenu précédemment, en fonction du temps. La courbe rouge pointillée montre l'ajustement permettant d'extraire le temps caractéristique  $\tau_{unbind}$ .

**Figure 4.16:** Dynamique de relargage des dimères de protéines au sein de l'intermédiaire.

### 3.1.4 Morphologie des intermédiaires

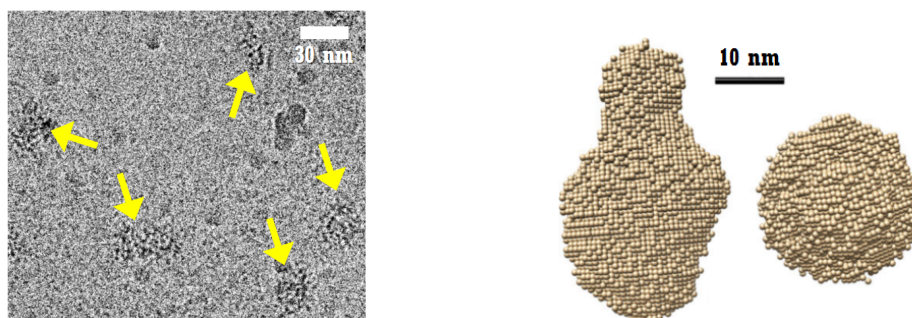
Afin de mieux comprendre ces complexes nucléoprotéiques, nous avons effectué une détermination *ab initio* de la forme des NPCs, à partir de la courbe de diffusion collectée à la fin de l'expérience d'assemblage réalisée à  $\rho = 6$ . Pour ce faire, nous avons utilisé des logiciels<sup>3</sup> de la suite ATSAS<sup>4</sup>, où le modèle utilisé, représente le complexe comme un ensemble de billes densément empilées à l'intérieur d'un volume donné. La structure obtenue, à faible résolution, est de forme allongée avec une longueur de  $\approx 380$  Å et un diamètre de  $\approx 230$  Å (Fig. 4.17b), ce qui est cohérent avec les images de cryo-microscopie électronique à transmission (cryo-TEM). En effet, sur la figure 4.17a est présentée la morphologie des NPCs

3. DAMMIN, DAMMIF et GNOM

4. Suite de programmes pour l'analyse des données de diffusion aux petits angles obtenues à partir de macromolécules biologiques[86]



à l'équilibre. Sur celle-ci, nous pouvons observer des espèces amorphes avec un diamètre d'environ  $27 \pm 4 \text{ nm}$ <sup>5</sup> et un rapport d'aspect moyen de 1,13. Ces grandeurs sont proches de celle du virion natif, i.e. environ 29 nm de diamètre.



(a) Images des NPCs réalisées par cryo-TEM. Les complexes, mis en évidence par des flèches jaunes, ont été obtenus dans les mêmes conditions que celles décrites dans la figure 4.14a. Les images ont été réalisées quelques jours après mélange. (b) Reconstruction des NPCs, calculée à partir de la dernière courbe d'intensité diffusée de l'expérience présentée sur la figure 4.14a. À gauche et à droite, sont présentées les vues latérale et supérieure, respectivement.

**Figure 4.17:** Etude des complexes nucléoprotéiques amorphes à l'équilibre ( $\rho = 6$ ).

## 3.1.5 Modèle d'assemblage

### 3.1.5.1 Résumé du modèle mathématique

Afin d'obtenir un aperçu quantitatif des énergies de liaison et de la diffusion des dimères de protéines au cours de l'auto-assemblage des NPCs, nous avons analysé les données à l'aide d'un modèle simple d'assemblage, i.e. :



où  $S$  désigne les dimères libres et  $C_N$  les complexes d'ordre  $N$  avec  $N \in \mathbb{N}$ .  $k^+$  et  $k^-$  sont les constantes de vitesse de la réaction.

Ce modèle mathématique, détaillé dans la section 3.1.2, permet d'estimer le nombre moyen de dimères par complexe,  $\langle N \rangle$ , en utilisant une expression ana-

---

5. Valeur d'écart type

lytique reliant la variation de ce dernier à  $I_0(t \rightarrow +\infty)$ . A l'équilibre, le nombre moyen de dimères est donné par l'expression suivante :

$$\langle N \rangle(t \rightarrow +\infty) \approx (c_S - K^{-1}) / c_G \quad (3.3)$$

où  $c_S$  et  $c_G$  sont respectivement les concentrations molaires totales de dimères et de génome.  $K$  est la constante d'équilibre.

De plus,  $K^{-1} = k^- / k^+$ , concentration critique en dimères au-dessus de laquelle les NPCs peuvent se former, peut être reliée à l'énergie libre de liaison d'un dimère sur un complexe via une loi de type Arrhénius. Il est à noter que, dans ce modèle, par souci de simplicité,  $k^+$  et  $k^-$  ont été assimilées à des constantes. Ceci implique que les effets des interactions des protéines entre elles, qu'elles soient libres ou liées au génome, sont moyennés sur  $N$ .

### 3.1.5.2 Énergie libre de liaison

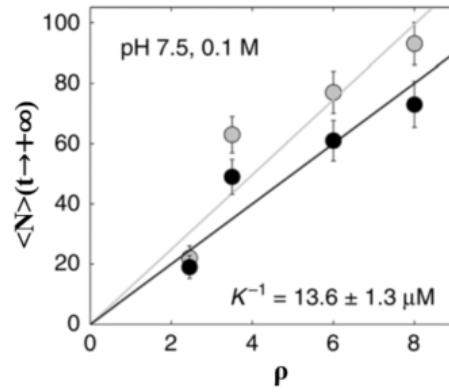
Comme précisé dans la section 3.1.2, nous avons réalisé des expériences d'assemblage à pH 7,5 et 0,1 M NaCl pour différentes valeurs de  $\rho$ . De ce fait, dans cette section, nous présentons la variation de  $\langle N \rangle$  estimé, près de l'équilibre, pour l'ensemble de ces expériences. Ces estimations ont été réalisées en utilisant le modèle d'assemblage, résumé dans la section précédente, ainsi que le calcul de  $\langle N \rangle_{up}$  présenté dans la section 3.1.1.

Sur la figure 4.18, sont présentées les estimations de  $\langle N \rangle$  obtenues pour différents rapports en masse. Les disques noirs correspondent à l'utilisation du modèle d'assemblage tandis que ceux en gris représentent le calcul de  $\langle N \rangle_{up}$ . Il est à noter que, pour toutes les expériences, la concentration totale en dimères de protéines  $c_S$  a été fixée, ce qui implique que  $\langle N \rangle(t \rightarrow +\infty) \propto \rho(1 - K^{-1} / c_S)$ .

A partir des données obtenues via le modèle d'assemblage, nous avons réalisé un ajustement linéaire des valeurs de  $\langle N \rangle$  (droite en noir sur la figure 4.18). Nous avons alors pu estimer une valeur de  $K^{-1}$  de  $13,6 \pm 1,3^6 \mu\text{M}$ . Cette valeur, associée à l'équation 3.13 du modèle d'assemblage (voir 3.1.2), permet de déterminer une énergie libre de liaison des dimères sur les NPCs de l'ordre de  $-17 \text{ kJ.mol}^{-1}$  soit

---

6. Erreur type de la moyenne



**Figure 4.18:**  $\langle N \rangle$  obtenus proche de l'équilibre pour différents rapports en masse et estimés à partir des intensités diffusées. Les disques noirs et gris correspondent à l'utilisation du modèle d'assemblage ou au calcul de  $\langle N \rangle_{up}$ , respectivement. Les lignes noires et grises correspondent, respectivement, aux ajustements linéaires de  $\langle N \rangle$  obtenus par les deux mêmes méthodes. La concentration critique  $K^{-1}$  est estimée à partir de la courbe noire. Les barres d'erreurs correspondent aux erreurs obtenues par propagation des écarts types sur le comptage des photons.

$-6,7k_B T_0$ , avec  $k_B$  la constante de Boltzmann et  $T_0$  la température ambiante (298 K). Dans le cas d'un ajustement linéaire de  $\langle N \rangle_{up}$  en fonction de  $\rho$  (ligne gris clair sur la même figure), nous obtenons une valeur de  $K^{-1}$  de  $10,9 \pm 1,8^7 \mu\text{M}$  et  $G_{bind} \simeq -6,9k_B T_0$ , soit une valeur quasi-identique. Cette valeur d'énergie est modérée, ce qui indique une interaction faible, et sans doute un comportement labile, entre les dimères de protéines et les NPCs. Il est à noter que  $G_{bind}$  comprend également la contribution des interactions des dimères entre eux, que nous avons estimée à  $\epsilon = -4,6 k_B T_0^8$  dans la sous-section 2.3.1.1.

Cependant, comme nous avons pu le voir dans le chapitre 1, précédemment, Garmann et al. ont effectué des expériences d'assemblage avec les protéines de capsid du CCMV et des ARN exogènes. Or, leurs résultats indiquaient que tous les dimères étaient liés à l'ARN quel que soit le rapport en masse dimères-ARN, et ce bien au-delà du rapport de 3,6 trouvé dans les virions natifs[75][76]. Cette constatation implique l'irréversibilité de la liaison et, de fait, une énergie de liaison élevée des dimères avec les NPCs. A contrario, l'énergie de liaison que nous avons estimée dans cette étude est modérée et ce, malgré son caractère électrostatique. Cette valeur paraît plus cohérente avec les expériences de compétition, réalisées par Comas-Garcia et al., et présentées dans le chapitre 1 [79]. En effet, une énergie

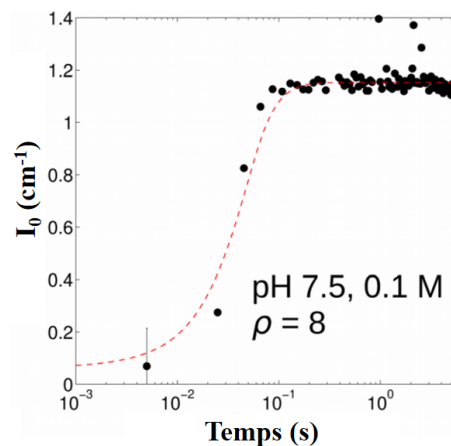
7. Erreur type de la moyenne

8. Pour rappel, cette valeur concernait les dimères associés en capsid vide.

de liaison modérée semble être avantageuse pour le virus puisqu'elle lui permettrait de sélectionner son génome de manière spécifique. De plus, cette valeur est en accord avec notre estimation de 75 dimères maximum par complexe dans des conditions où le rapport molaire dimères-génome est de 150 ( $\rho = 6$ ), puisqu'une fraction importante de dimères est restée libre.

### 3.1.5.3 Mécanismes contrôlant la réaction

Ces expériences, couplées au modèle d'assemblage, permettent non seulement d'estimer une énergie d'interaction au sein des NPCs mais également de discuter des mécanismes sous-jacents à l'assemblage. En effet, dans l'hypothèse d'une réaction limitée par la diffusion, la constante de vitesse de réaction,  $k^+$ , du modèle d'assemblage serait exprimée telle que :  $k^+ = 4\pi N_A R_C D_S$ , avec  $R_C$  un rayon typique pour les NPCs fixé à 120 Å,  $D_S$  le coefficient de diffusion apparent des dimères et  $N_A$  est la constante d'Avogadro.



**Figure 4.19:** Dynamique d'assemblage des NPCs pour  $\rho = 8$ , un pH final de 7,5 et une force ionique de 0,1 M NaCl. Les cercles noirs montrent l'évolution de  $I_0$  en fonction du temps. La courbe rouge correspond à l'ajustement réalisé à l'aide du modèle d'assemblage.

L'ajustement réalisé sur la figure 4.19, à l'aide du modèle présenté dans la section précédente, permet alors d'estimer le coefficient de diffusion apparent des dimères à  $4,9 \times 10^{-12} \text{ m}^2\text{s}^{-1}$ . Cependant, le coefficient de diffusion d'un dimère de protéines seul a été calculé, par simulations de dynamique moléculaire, à  $3,0 \times 10^{-10} \text{ m}^2\text{s}^{-1}$ , soit deux ordres de grandeur plus élevés. Cette différence indique une barrière énergétique qui peut être due aux dimères liés, au sein des NPCs. En effet, ceux-ci

pourraient empêcher l'insertion des dimères libres soit par répulsion électrostatique, soit par effet stérique. De plus, une dépendance vis-à-vis de l'orientation des dimères libres ou de la conformation des dimères, notamment au niveau du domaine de liaison à l'ARN, pourrait également participer à cette barrière. Dans tous les cas, le processus de liaison n'est pas uniquement limité par la diffusion mais également par la réaction.

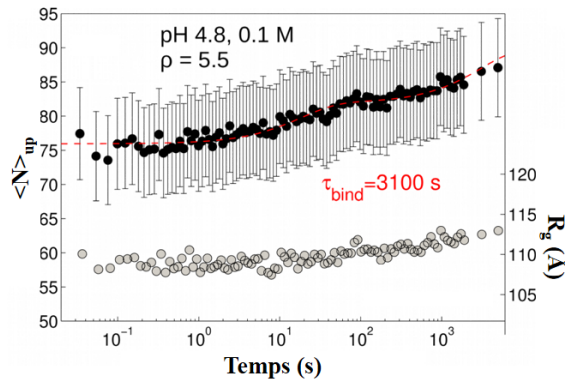
## 3.2 Relaxation des intermédiaires en capsides

Les expériences à  $pH$  7,5 n'ont pas permis d'obtenir de capsides correctement formées. Cependant, comme nous avons pu le voir dans la section 2.3.2.2, en abaissant le  $pH$  des solutions de NPCs, les groupements carboxyles à la surface des protéines peuvent être neutralisés avec, pour effet, une augmentation de l'attraction entre les dimères de protéines.

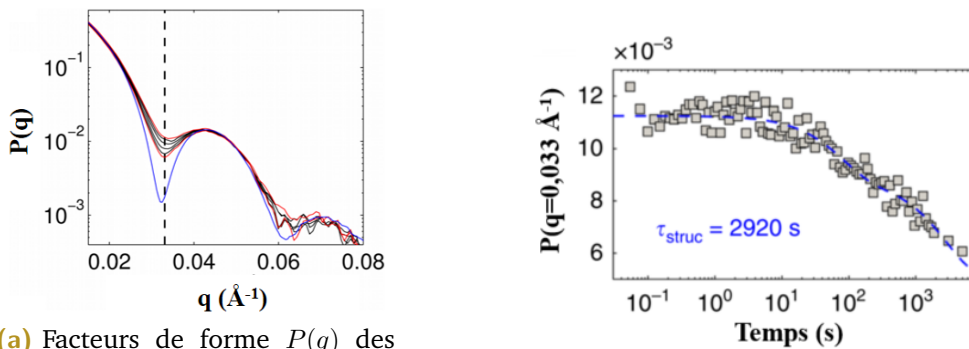
### 3.2.1 Formation des capsides

Dans cette section, nous avons réalisé un saut de  $pH$  en mélangeant une solution de NPCs pré-assemblés avec un tampon acide. La solution obtenue et, contenant des NPCs, a un  $pH$  final de 5,2, une force ionique de 55 mM et un rapport en masse,  $\rho = 5,5$ . L'expérience est présentée sur la figure 4.20. Celle-ci montre l'évolution de la moyenne supérieure du nombre de dimères par complexe (ordonnée de gauche) et, du rayon de giration (ordonnée de droite), en fonction du temps.

Dans un premier temps, nous pouvons observer une augmentation lente de  $\langle N \rangle_{up}$ , qui passe de 75 à environ 85, sur plus d'une heure. De plus, un ajustement à l'aide d'une double exponentielle croissante (ligne pointillée rouge), nous permet d'estimer deux temps de liaison  $\tau_{bind}^s$  et  $\tau_{bind}$ . C'est ce dernier qui nous intéresse et vaut environ 3100 s. Celui-ci est 2 ordres de grandeur supérieur aux temps estimés précédemment, ce qui indique une dynamique plus lente. Nous pouvons ensuite remarquer que le rayon de giration n'évolue pas de manière significative au cours de l'expérience. En effet, une variation de quelques angströms par rapport à 11 nm est inférieure à la résolution typique de la diffusion aux petits angles.



**Figure 4.20:** Évolution temporelle de  $\langle N \rangle_{up}$  (disques noirs) et  $R_g$  (disques gris clairs) après mélange des NPCs dans un tampon acide ( $T = 20^\circ\text{C}$ ). La ligne pointillée rouge correspond à l'ajustement de  $\langle N \rangle_{up}$ .



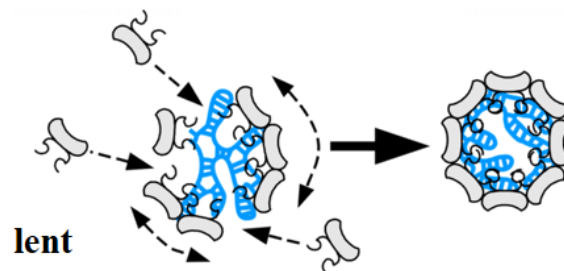
**(a)** Facteurs de forme  $P(q)$  des NPCs (courbes rouges et noires) à différents temps. La ligne en pointillé noir indique  $q = 0,033 \text{ \AA}^{-1}$ .

**(b)** Évolution de la valeur des facteurs de forme des NPCs pour  $q = 0,033 \text{ \AA}^{-1}$ , en fonction du temps.

**Figure 4.21:** Étude du temps caractéristique de relaxation structurale, après mélange des NPC préassemblés avec une solution tampon acide ( $T = 20^\circ\text{C}$ ). Conditions finales : pH 5,2, la force ionique à 55 mM,  $\rho = 5,5$ .

Cependant, si nous nous intéressons à la figure 4.21a, où sont présentés les facteurs de forme  $P(q)$  des NPCs (courbes rouges et noires) à différents instants de la cinétique, nous pouvons tout de même observer une évolution. Sur cette figure, les courbes rouges se rapportent aux première et dernière mesures collectées à 54 ms (courbe supérieure) et à 4970 s (courbe inférieure), respectivement. Une inspection minutieuse de ces facteurs de forme montre que les arches deviennent de plus en plus prononcées à mesure que la cinétique progresse, ce qui indique une auto-organisation structurale. De plus, sur cette figure également, la courbe bleue correspond au facteur de forme des virions purifiés. Une comparaison avec celui-ci montre que les NPCs relaxés ont une taille et une structure proches de celles des virions natifs. Toujours sur cette figure, le trait pointillé noir indique

le premier minimum du facteur de forme ( $q = 0,033\text{\AA}^{-1}$ ). En suivant l'évolution de celui-ci en fonction du temps, nous pouvons estimer un temps caractéristique de relaxation structurelle  $\tau_{struct}$ . C'est ce qui est présenté sur la figure 4.21b, sous la forme de carrés gris. En ajustant ces données par une double exponentielle décroissante, i.e. la courbe pointillée bleue, nous obtenons un temps  $\tau_{struct}$  à 2920 s. Cette valeur est du même ordre de grandeur que  $\tau_{bind}$  estimé précédemment. Il est à noter que nous avons utilisé une double exponentielle, ce qui indique deux phases. La première phase ( $\tau_{bind}^s$ ) est sans doute rapide du fait que plusieurs sites de liaison des complexes sont facilement accessibles aux dimères libres, tandis que la deuxième phase ( $\tau_{bind}$ ) est ralentie du fait de la restructuration lente de la capsid. De plus,  $\tau_{bind} \approx \tau_{struct}$ , ce qui indique que le système est probablement passé par un mécanisme *synchrone*, illustré sur le dessin ci-dessous, lors duquel les NPC's s'auto-organisent lentement tout en capturant des dimères libres afin de former des espèces structurellement identiques aux virions natifs (Fig. 4.22).



**Figure 4.22:** Illustration schématique du mécanisme *synchrone*.

### 3.2.2 Énergie d'activation

Les expériences présentées dans cette étude permettent non seulement d'investiguer les mécanismes d'assemblage mais également d'estimer des énergies d'activation. En effet, dans le cas d'un processus activé par l'énergie, le temps de liaison doit suivre une loi de type Arrhenius :

$$\tau_{bind} \propto \exp(E_a/k_B T) \quad (3.4)$$

où  $E_a$  est l'énergie d'activation.

De ce fait, nous avons effectué la même expérience de relaxation pour différentes températures allant de 10°C à 40°C par pas de 10°C. Pour ces expériences, les

solutions finales, contenant les NPCs préassemblés, avaient un  $pH$  final de 5,2 et une force ionique de 55 mM. L'évolution de la moyenne supérieure du nombre de dimères par complexe a également été ajustée à l'aide d'une double exponentielle décroissante afin d'estimer les temps de liaison  $\tau_{bind}^s$  et  $\tau_{bind}$ . Les valeurs de  $\tau_{bind}$  avec leurs incertitudes, calculées en tenant compte des erreurs sur  $\langle N \rangle_{up}$ , sont résumées dans le tableau 4.1.

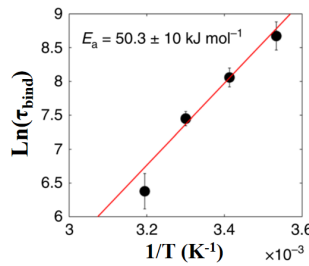
(a)

TEMPS DE LIAISON $\tau_{bind}$	
$T = 10^\circ\text{C}$	$5840 \pm 1200$ s
$T = 20^\circ\text{C}$	$3160 \pm 440$ s
$T = 30^\circ\text{C}$	$1720 \pm 180$ s
$T = 40^\circ\text{C}$	$590 \pm 160$ s

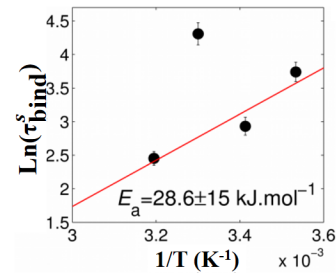
(b)

TEMPS DE STRUCTURATION $\tau_{struc}$	
$T = 10^\circ\text{C}$	$13700 \pm 2100$ s
$T = 20^\circ\text{C}$	$2920 \pm 530$ s
$T = 30^\circ\text{C}$	$1970 \pm 340$ s
$T = 40^\circ\text{C}$	$390 \pm 120$ s

**Table 4.1:** Tableau récapitulatif des temps de liaison et de structuration,  $\tau_{bind}$  et  $\tau_{struc}$ , en fonction de la température. Les erreurs correspondent aux écarts types.



(a) Logarithme de  $\tau_{bind}$  en fonction de l'inverse de la température



(b) Logarithme de  $\tau_{bind}^s$  en fonction de l'inverse de la température

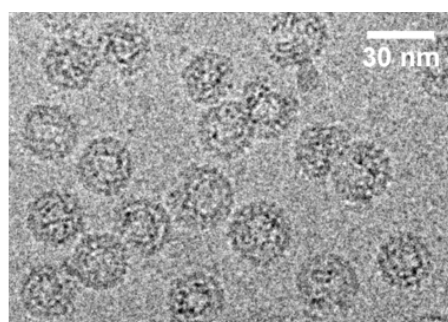
**Figure 4.23:** Logarithme de  $\tau_{bind}$  et de  $\tau_{bind}^s$  en fonction de l'inverse de la température pour des expériences effectuées de 10 à 40°C. La droite rouge est l'ajustement linéaire donnant l'énergie d'activation  $E_a$ .

De plus, en représentant le logarithme de  $\tau_{bind}$  en fonction de  $1/T$ , nous devrions trouver une relation linéaire. C'est ce que nous pouvons observer sur la figure 4.23a. En effet, les données peuvent être ajustées par une droite qui permet d'estimer une barrière d'énergie  $E_a$  de l'ordre de  $20 k_B T_0$ . Nous avons également étudié les temps caractéristiques courts,  $\tau_{bind}^s$ , lors du processus de liaison. Ainsi, la même méthode d'analyse nous a permis de découvrir que la phase rapide pouvait également être activée avec une énergie d'environ  $11 k_B T_0$  (Fig. 4.23b). Nous avons ensuite examiné l'évolution du premier minimum du facteur de forme des NPCs en fonction du temps, afin d'estimer les temps caractéristiques de relaxation



structurelle  $\tau_{struc}$  de chaque expérience. Les résultats, avec leurs incertitudes, ont été résumés dans le tableau 4.1. Nous pouvons alors remarquer que, quelle que soit la température,  $\tau_{struc}$  reste du même ordre de grandeur que  $\tau_{bind}$ , ce qui indique que le mécanisme *synchrone*, de formation des capsides, est conservé quelle que soit la température.

La morphologie de l'espèce obtenue a été vérifiée par cryo-TEM après quelques jours d'incubation (Fig. 4.24). Comme les virions natifs, ils présentaient un diamètre de  $29 \pm 1$  nm<sup>9</sup> ainsi qu'un noyau creux du fait que le génome est compacté contre la surface intérieure de la capside en raison des interactions électrostatiques[87].



**Figure 4.24:** Images Cryo-TEM des virions reformés, dans les mêmes conditions que la figure 4.21, réalisées après plusieurs jours d'incubation.

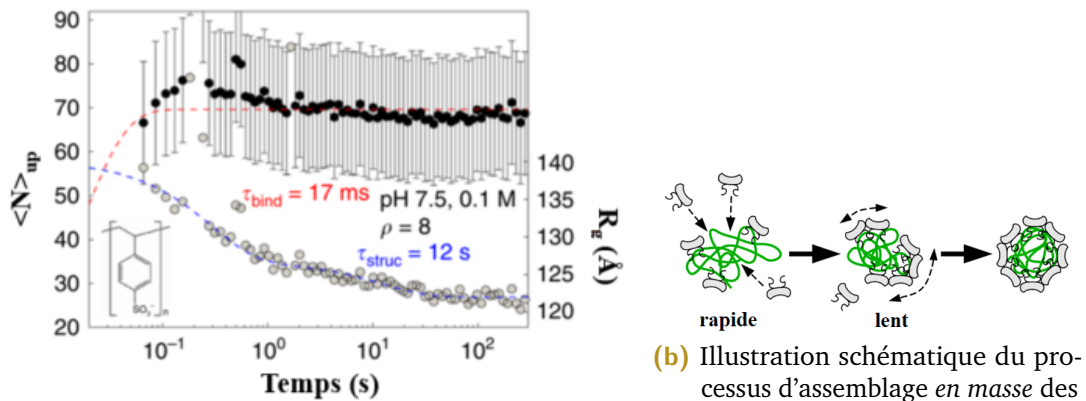
### 3.3 Auto-assemblage des capsides autour du PSS

En plus d'étudier la formation de la capside, autour du génome natif, nous avons également étudié celle-ci autour d'un polyélectrolyte. Les interactions entre celui-ci et les protéines étant non-spécifiques, cette étude permet de mieux comprendre les mécanismes sous-jacents de l'assemblage. De ce fait, dans cette section, je présenterai les expériences d'assemblage réalisées avec du poly(styrène sulfonate) (PSS), un polyélectrolyte synthétique souple chargé négativement. La densité de charge linéique de ce polyélectrolyte est d'environ  $0,33e\text{\AA}^{-1}$ , soit une valeur proche de celle de l'ARN. Lors de nos expériences, nous avons utilisé des PSS de 600 kDa qui portaient, en moyenne, environ 3100 charges négatives, soit une quantité

---

9. Ecart type

similaire à celle du génome viral. De plus, les dimères et les PSS, mélangés à un  $pH$  de 7,5, produisent des espèces sphériques similaires aux virions, ce qui en fait un candidat idéal pour étudier le phénomène d'auto-assemblage des dimères[85]. Les conditions utilisées pour déclencher l'assemblage sont les mêmes que celles utilisées dans la section 3.1.5.2.



(a) Évolution de  $\langle N \rangle_{up}$  (cercles noirs) et  $R_g$  (cercles gris), après mélange des dimères libres et du PSS, en fonction du temps. Les courbes rouge et bleue, en pointillés, montrent les ajustements permettant d'extraire les temps caractéristiques  $\tau_{bind}$  et  $\tau_{struc}$ , respectivement.

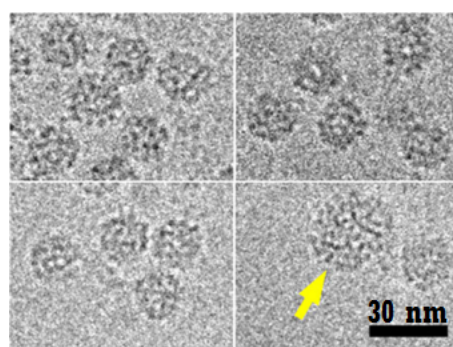
**Figure 4.25:** Dynamique d'auto-assemblage des capsides pleines de PSS de 600 kDa. La solution finale a un  $pH$  de 7,5, une force ionique de 0,1 M NaCl et un rapport en masse de 8.

La figure 4.25a montre l'évolution temporelle de  $\langle N \rangle_{up}$  (ordonnée de gauche) et du rayon de giration (ordonnée de droite) lors de la cinétique d'assemblage des dimères autour du PSS.  $\langle N \rangle_{up}$  a été estimée en supposant une moyenne de 1,5 chaînes de PSS empaquetés par capside, comme précédemment établi par SANS[85].

Dans un premier temps, nous pouvons observer que les dimères de protéines se lient très rapidement au PSS, puis que  $\langle N \rangle_{up}$  atteint un plateau autour de 70 dimères. Cette valeur indique la présence d'objets plus petits que des virions natifs (90 dimères nécessaires). Nous avons ensuite réalisé un ajustement des données, via une simple fonction exponentielle, qui indique un temps de liaison,  $\tau_{bind}$ , de 17 ms. Cette valeur est très proche de celle obtenue avec le génome dans les mêmes conditions (28 ms). Dans un second temps, nous pouvons remarquer que le rayon de giration diminue lentement sur l'ensemble de la période observée, avec une

estimation du rayon de giration, à la fin de l'expérience, comprise entre 12 et 12,5 nm. Il est à noter que cette valeur est supérieure à celle estimée pour les virions, qui était comprise entre 11 et 11,5 nm (section 3.2.1). En ajustant les données, comme précédemment avec une double exponentielle décroissante, nous obtenons un temps caractéristique  $\tau_{struc} = 12$  s, qui est du même ordre de grandeur que la valeur déterminée en présence du génome (48 s). Ainsi, l'évolution temporelle de  $R_g$  présente deux phases, probablement de même nature que celles observées avec le génome (section 3.1.2). Dans la mesure où,  $\tau_{bind} \ll \tau_{struc}$ , il est légitime de penser que les capsides pleines de PSS s'auto-assemblent via le chemin cinétique *en masse*. En effet, les dimères de protéines se lient rapidement aux polyélectrolytes et les complexes résultants s'auto-organisent lentement en espèces sphériques (Fig. 4.25b).

Les images Cryo-TEM de la figure 4.26 montrent que les capsides finales sont bien assemblées mais légèrement plus polydisperses que les virions natifs. Ceci indique la présence de défauts au sein de ces capsides. Si la grande majorité des capsides a



**Figure 4.26:** Images Cryo-TEM des capsides pleines de PSS après plusieurs jours d'incubation ( $\rho = 6$ ). La flèche jaune montre une capside de taille supérieure à 30 nm.

un diamètre de l'ordre de  $23 \pm 1$  nm<sup>10</sup>, il existe tout de même quelques espèces dont la taille est supérieure à une trentaine de nanomètres (flèche jaune sur la figure 4.26), ce qui concorde bien avec les résultats obtenus précédemment par l'équipe[85]. Les valeurs mesurées par microscopie électronique ne sont cependant pas cohérentes avec les mesures des rayons de girations des virions et des capsides pleines de PSS (voir plus haut). Cette différence s'explique par la polydispersité des capsides pleines de PSS. En effet, les quelques objets dont la taille est supérieure à 30 nm augmentent le rayon de giration moyen des objets en solution.

10. Ecart type

Ainsi, les capsides peuvent se fermer, à *pH* neutre, en présence de PSS, ce qui n'était pas possible avec le génome. Ceci peut s'expliquer par l'hydrophobicité et la flexibilité du PSS. En effet, en raison de son squelette hydrophobe, le PSS est insoluble dans l'eau lorsque sa fraction de charge tombe en dessous de 30% [88]. De plus, il est possible que les domaines de liaison à l'ARN des dimères liés neutralisent une grande partie des charges négatives portées par le PSS. Ce dernier, très souple, pourrait alors s'effondrer via une transition pelote-globule. De ce fait, malgré la barrière d'énergie liée à leur répulsion électrostatique et aux changements de conformation requis pour constituer la structure icosaédrique, les dimères liés pourraient alors interagir entre eux pour former une capside fermée. Une étude théorique a démontré que les chaînes linéaires flexibles requièrent moins d'énergie libre que les chaînes rigides et ramifiées pour être confinées dans une capside [89], ce qui est en accord qualitatif avec nos résultats. Comme les virions, les capsides remplies de PSS sont stables à la dilution, ce qui indique qu'en fin d'assemblage, il n'y a plus d'échange entre les dimères de protéines libres et en solution et ceux constituant la capside.

## 3.4 Assemblage *in vivo*

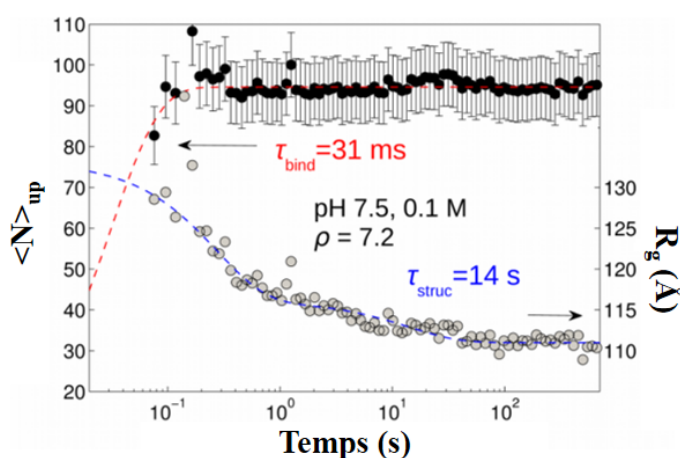
Les expériences présentées dans cette étude ont été réalisées *in vitro*. Et, si les résultats précédents suggèrent que les complexes sont formés à *pH* neutre, puis relaxent en virions à *pH* acide, ce qui peut conduire l'assemblage de la capside *in vivo* n'est toujours pas clair.

### 3.4.1 Influence du *pH* ?

Une première possibilité serait un assemblage *in vivo* en deux étapes piloté par le *pH*. Un tel processus d'assemblage est notamment utilisé par les flavivirus [90]. En effet, dans leur cas, les capsides immatures sont assemblées dans le réticulum endoplasmique (*pH* 7,2) avant d'être transportées vers l'appareil de Golgi, où la maturation se produit à la suite d'une diminution du *pH* (*pH* 5,7)

### 3.4.2 Influence des ions divalents ?

Un autre facteur possible était les cations divalents  $\text{Ca}^{2+}$ . En effet, comme nous avons pu le voir au début de ce chapitre, les mesures réalisées par CTPF, sur les capsides vides, avaient révélé un déplacement de la  $T_m$  indiquant une stabilisation de celles-ci par capture d'ions divalents. De plus, ces résultats étaient corroborés par ceux de Speir et al. sur les virions[11]. Cependant, nos expériences effectuées en présence de  $\text{Ca}^{2+}$  n'ont révélé aucun changement significatif dans la dynamique d'auto-assemblage des NPCs. Celles-ci ont été réalisées pour  $\rho = 7,2$  dans une solution tampon à pH 7,5 contenant 0,1 M de NaCl et 4 mM de  $\text{CaCl}_2$ .



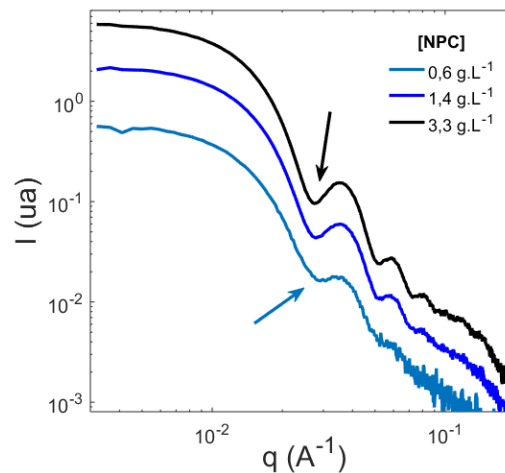
**Figure 4.27:** Dynamique d'assemblage des NPCs en présence d'ions divalents (4 mM  $\text{CaCl}_2$ ). Les disques noirs et gris montrent l'évolution temporelle de  $\langle N \rangle_{up}$  et  $R_g$ , respectivement. Les courbes pointillées rouge et bleue présentent les ajustements permettant d'extraire les temps caractéristiques  $\tau_{bind}$  et  $\tau_{struc}$ , respectivement.

En effet, comme nous pouvons le voir sur la figure 4.27, après ajustement de  $\langle N \rangle_{up}$  et de  $R_g$  par une simple exponentielle et une double exponentielle décroissante, les temps caractéristiques,  $\tau_{bind} = 31 \text{ ms}$  et  $\tau_{struc} = 14 \text{ s}$ , sont similaires à ceux obtenus précédemment (28 ms et 48 s, respectivement). Nous pouvons tout de même noter une différence, i.e.  $\langle N \rangle_{up}(t \rightarrow \infty)$  s'élève à 94, ce qui est proche de 90, la valeur attendue et déterminée par l'ajustement réalisé sur la figure 4.18.

### 3.4.3 Influence de la concentration en protéine ?

Une dernière possibilité qui n'avait pas été testée est l'influence de la concentration en protéine. En effet, il est probable que les concentrations en protéines et en ARN

ne soit pas homogènes au sein de la cellule hôte. Ainsi, *in vivo*, une concentration locale plus importante pourrait avoir un impact sur la dynamique d'assemblage.



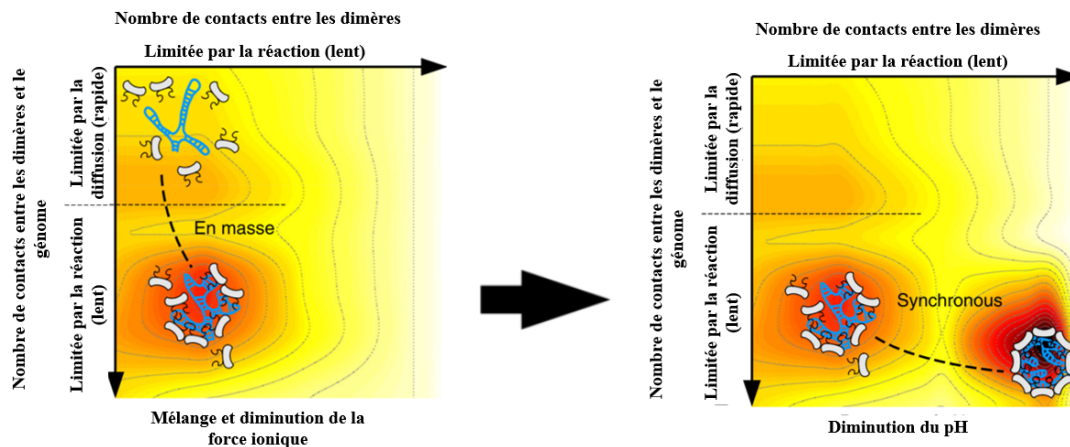
**Figure 4.28:** Graphiques superposant les courbes d'intensités obtenues, par SAXS, des NPCs formées pour différentes concentrations initiales en protéines et ARN.

C'est ce qui a été testé récemment par l'équipe SOBIO au LPS (Fig. 4.28). Dans cette expérience, les protéines et l'ARN2 du CCMV ont été mélangés dans une solution à  $pH$  neutre et forte salinité (0,5 M) puis dialysés contre une solution à plus faible salinité (0,1 M). Le ratio en masse a été conservé pour tous les mélanges (4,5). Les concentrations finales en NPCs ont été calculées de la même manière que pour les virions. Ainsi, sur la figure 4.28, nous pouvons observer que plus les concentrations en protéines et en ARN sont importantes, plus les arches des courbes sont prononcées. Ceci indique que les NPCs sont mieux structurées aux plus fortes concentrations. Nous pouvons donc imaginer qu'*in vivo*, la barrière énergétique peut être abaissée par augmentation locale de la concentration plutôt que par diminution du  $pH$ .

### 3.5 Conclusion

Dans les conditions expérimentales de l'étude, nous avons pu observer les deux mécanismes, *en masse* et *synchrone*, mis en évidence lors de simulations gros grains, réalisées précédemment par Perlmutter et al.[71].

Nous avons pu montrer, lors de nos expériences d'empaquetage du génome natif par les dimères de protéines, que ces chemins cinétiques pouvaient être sélection-



(a) Formation des complexes, via un processus d'assemblage *en masse*, après mélange des dimères et du génome à un  $pH$  fixé.

(b) Relaxation des complexes en capsides, via un processus d'assemblage *synchrones*, après un saut de  $pH$ .

**Figure 4.29:** Représentation schématique du diagramme d'énergie libre. Les couleurs sombres représentant les énergies les plus basses tandis que les couleurs claires montrent les énergies plus élevées.

nés en fonction des interactions dimère-dimère et dimère-génome et, ceci, par modification du  $pH$  et de la force ionique. La dynamique d'auto-assemblage peut alors être résumée par un diagramme d'énergie libre fonction de deux paramètres, i.e. le nombre de contacts dimères-dimères et le nombre de contacts dimères-génome, respectivement en abscisse et en ordonnée de la figure 4.29. Sur celle-ci, le long de l'axe des abscisses, la dynamique est limitée par la réaction tandis que le long de l'axe des ordonnées, la dynamique n'est limitée par la diffusion que dans la mesure où seuls quelques dimères sont liés au génome.

Ainsi, après l'assemblage des dimères autour du génome via un processus *en masse*, il existe un barrière d'énergie élevée à surmonter (Fig. 4.29a). Celle-ci peut être abaissée par diminution du  $pH$  afin de permettre la relaxation des complexes en capsides icosaédriques via un processus *synchrones* (Fig. 4.29b).

## Conclusion du chapitre

Dans ce chapitre, nous avons étudié les interactions au sein des capsides vides et pleines du CCMV par CTPF ainsi que l'auto-assemblage des capsides pleines par la technique de TR-SAXS.

Dans le premier cas, nous avons pu estimer une énergie d'interaction des dimères, entre eux, de  $-4,6 k_B T$ , ce qui est similaire à ce qui avait été rapporté précédemment. La même technique nous a également permis d'estimer une charge effective du dimère de protéine allant de 3 à  $4,5e$ , en valeur absolue, pour des  $pH$  de 4,8 à 6,5 respectivement et pouvant être abaissée en présence d'ions divalents. De plus, en présence du génome, nos expériences ont montré que la charge effective  $|Z|$  disparaissait à  $pH$  4,8, du fait de la contribution de celui-ci. Dans le deuxième cas, nos expériences d'empaquetage, du génome natif par les dimères de protéines, nous ont permis de montrer que les mécanismes, précédemment mis en évidence lors des simulations gros grains, pouvaient être sélectionnés en fonction des interactions dimère-dimère et dimère-génome. Ainsi, dans nos conditions expérimentales, l'assemblage des dimères autour du génome s'est effectué via un processus *en masse*, où les ions divalents ne jouaient pas de rôle, tandis que la relaxation des complexes en capsides icosaédriques a eu lieu via un processus *synchrone*. De plus, nous avons pu observer, de manière surprenante, qu'acidifier la solution afin d'accroître les interactions entre les dimères de protéines n'était pas nécessaire pour former des capsides remplies de PSS. Cette dernière observation montre que les dimères sont capables de créer une capsidic icosaédrique pleine, en une seule étape, à  $pH$  neutre. C'est notamment ce qui a pu être confirmé lors d'expériences préliminaires où les concentrations en protéines et ARN étaient plus élevées.

D'un point de vue biologique, une première phase de coopération entre les dimères et le génome peut être nécessaire afin d'empaqueter rapidement le génome pour le protéger. Une seconde étape, synchrone, pourrait ensuite éviter les erreurs d'assemblage et, notamment l'empaquetage d'ARN exogène ainsi que les pièges cinétiques. À cet égard, il est à noter que lorsque nous avons essayé de réaliser un assemblage en une étape, en mélangeant les dimères et le génome tout en abaissant



simultanément le  $pH$  et la force ionique, de nombreux objets mal assemblés et agrégés se sont formés.

L'ensemble des éléments précédemment présentés, nous amène à émettre l'hypothèse que l'assemblage se produit, *in vivo*, à  $pH$  neutre et que les cations divalents ont peu de rôle à jouer à ce stade. Les NPCs se forment, progressivement, à mesure que les dimères de protéines sont exprimés permettant ainsi une sélectivité efficace pour le génome. Lorsque le nombre de dimères par NPCs atteint une valeur suffisamment élevée (probablement bien supérieure à 100) les NPCs relaxent de manière irréversible en virions icosaédriques. Cette description présente l'avantage de fournir un mécanisme simple, i.e. l'expression de la protéine de capsid, permettant le passage à l'étape d'assemblage, phase cruciale du cycle viral.



---

Cinétiques d'auto-assemblage et  
désassemblage de la capsid vide du

VHB

De par le monde, et malgré l'existence d'un vaccin depuis les années 80, 240 millions de personnes sont aujourd'hui infectées par le virus de l'hépatite B (VHB) <sup>1</sup>. Or, ce dernier ne possède pas de traitement curatif et peut causer de sérieuses complications comme des cancers ou des cirrhoses. Comprendre l'assemblage et le désassemblage de la capsid du virus, phases critiques du cycle de vie du VHB, est une stratégie prometteuse pour de nouveaux traitements [91]. Au-delà des retombées thérapeutiques, l'étude de ces phénomènes permet d'améliorer les connaissances sur les mécanismes sous-jacents avec des applications possibles dans des domaines tels que la bio-ingénierie et l'administration de médicaments [92][93].

Comme nous avons pu le voir dans le chapitre 3, le virus de l'hépatite B est constitué d'une capsid entourant et protégeant le génome. Si les processus d'assemblage et de désassemblage reposent sur des interactions entre les protéines de la capsid et le génome, il a été démontré que les capsides peuvent s'assembler *in vitro* en l'absence de ce dernier [35]. Pour des phénomènes complexes, comme l'assemblage des capsides de virus, une première approche consiste à ne considérer que les protéines et leurs interactions. De ce fait, l'auto-assemblage et le désassemblage de la capsid du VHB sont étudiés à l'aide de la protéine tronquée, i.e. composée uniquement du domaine d'assemblage, CP149 (voir chapitre 3).

Cette partie du manuscrit est divisée en cinq chapitres. Dans un premier temps, je présenterai une rapide étude bibliographique sur l'auto-assemblage et le désassemblage des capsides vides du VHB. Ces phénomènes étant dirigés par des interactions électrostatiques et hydrophobes, je présenterai ensuite une approche expérimentale visant à déterminer les énergies d'interaction existantes au sein de la capsid ainsi que les charges effectives portées par les dimères de protéines. Dans une troisième partie, je présenterai l'analyse par décomposition en valeurs singulières, i.e. un outil d'analyse utilisé pour étudier l'assemblage et le désassemblage de la capsid vide du VHB. Je présenterai ensuite les résultats ainsi que les analyses obtenus pour les expériences de désassemblage. Le dernier chapitre sera consacré aux résultats et analyses obtenus pour les études cinétiques d'assemblage des capsides vides du VHB.

---

1. Source : ANRS (Agence nationale de recherches sur le sida et les hépatites virales)

# Étude bibliographique de l'assemblage et du désassemblage des capsides vides du VHB

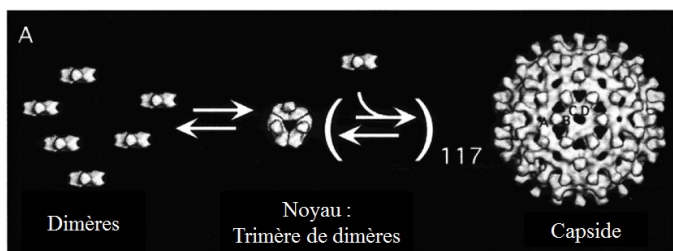
## 1.1 Auto-assemblage des capsides vides du VHB

Lors de l'assemblage, les dimères de protéines s'agencent avec précision pour aboutir à la capsid virale. En théorie, le nombre d'intermédiaires possibles croît de manière combinatoire et le processus s'apparente au paradoxe de Levinthal pour le repliement des protéines. En réalité, seuls quelques intermédiaires contribuent de manière significative à la dynamique d'assemblage. Ci-après, je présenterai tout d'abord le mécanisme le plus couramment utilisé pour décrire la dynamique d'assemblage des capsides vides. Certaines expériences ont permis l'identification d'intermédiaires possibles ainsi qu'une estimation du rôle des interactions entre les protéines dans la dynamique d'assemblage. De ce fait, celles-ci seront ensuite exposées.

### 1.1.1 Mécanisme de nucléation-croissance

L'auto-assemblage des capsides vides des virus a pu être décrit comme un mécanisme de nucléation-croissance. Dans cette description, la nucléation fait référence à la formation d'un noyau dont la probabilité de croître pour former une capsid complète, avant de se dissocier, est supérieure à 50 %. Ce noyau correspond à des structures organisées comme des pentamères de dimères ou des trimères de dimères. Le terme de croissance fait référence au temps caractéristique nécessaire au noyau pour croître afin de former une capsid complète [94]. A l'issue de cette croissance, la capsid est icosaédrique et il n'y a pas de réorganisation possible.

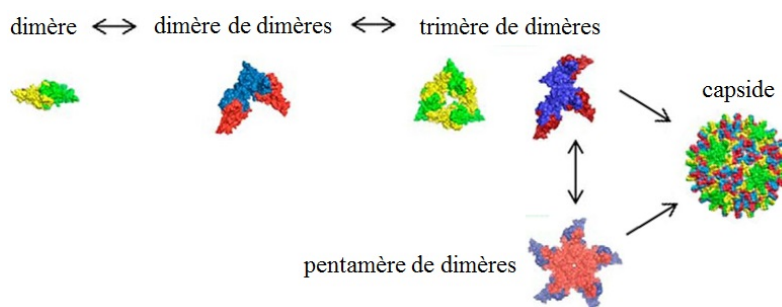
Cette description a notamment été utilisée pour décrire l'auto-assemblage du VHB par Zlotnick et al.[95]. Dans leur publication, les auteurs réalisent des expériences de diffusion de lumière couplées avec de la chromatographie d'exclusion de taille. A partir des résultats, ils proposent un modèle de croissance par polymérisation avec la formation d'un noyau sous la forme de trimères de dimères, comme nous pouvons le voir sur la figure 5.1. Ce modèle a l'avantage de donner une explication



**Figure 5.1:** L'assemblage de la capside commence par la formation d'un noyau sous la forme d'un trimère de dimères, puis la croissance continue par l'ajout séquentiel de dimères (Adapté de [95]).

pour le changement morphologique entre les formes de capside  $T = 3$  et  $T = 4$  du VHB. Les noyaux constitués de deux dimères A-B et d'un dimère C-D ou de trois dimères C-D imposerait, par la suite, une croissance aboutissant à la capside  $T = 4$ . Pour une capside  $T = 3$ , un seul noyau serait possible, composé de deux dimères A-B et d'un dimère C-C.

Cette description a également été utilisée par d'autres, tel que Holmes et al.[96]. Cependant, ces derniers ont identifié des intermédiaires tétramériques et pentamériques possibles, par spectrométrie de masse. De ce fait, cette fois-ci, un noyau trimérique ouvert est proposé afin de permettre une croissance via les intermédiaires identifiés (Fig. 5.2). Dans ce cas, les auteurs proposent que la commutation

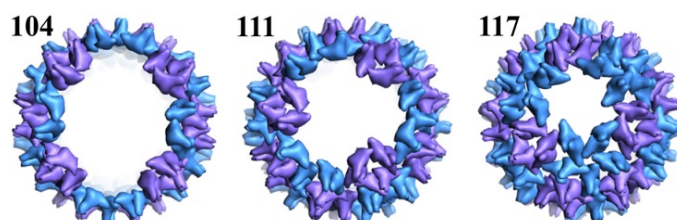


**Figure 5.2:** Schéma de la voie d'assemblage proposée, via des intermédiaires trimériques ou pentamériques (Adapté de [96]).

entre les capsides  $T = 3$  et  $T = 4$  serait régulée sur une base cinétique comme cela avait déjà été postulé [97].

### 1.1.2 Intermédiaires tardifs

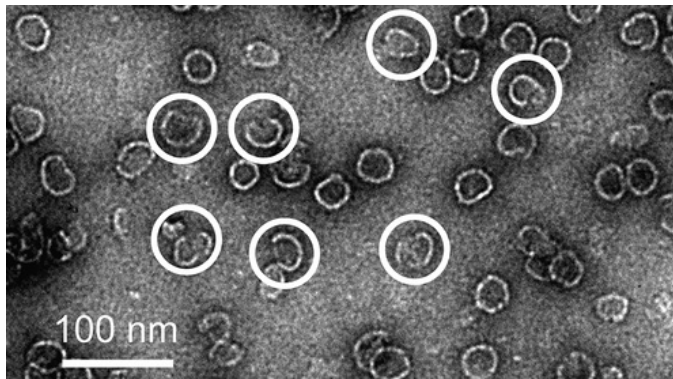
Si des intermédiaires précoces ont pu être proposés afin de décrire l'initialisation de l'assemblage de la capside vide du VHB, des intermédiaires tardifs ont également été avancés par Pierson et al [98]. Dans cette publication, de gros intermédiaires ont été identifiés à de fortes concentrations salines par spectrométrie de masse à détection de charge et cryo-microscopie électronique. L'avantage de coupler ces deux techniques est que la première donne des informations sur le nombre de dimères composant les intermédiaires tandis que la seconde permet d'obtenir



**Figure 5.3:** Modèles possibles des structures des intermédiaires identifiés par Pierson et al. (Adapté de[98]).

des informations sur la structure de ces derniers. De ce fait, les structures des intermédiaires probables peuvent être identifiées avec précision (Fig. 5.3).

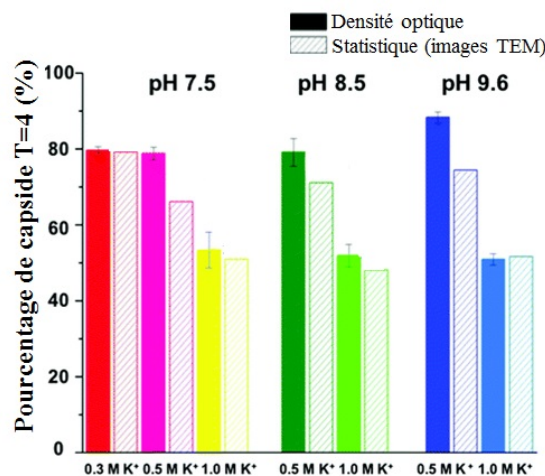
Des intermédiaires de tailles importantes ont également été observés par Kondylis et al. lors d'une étude sur l'effet des dérivés de phénylpropénamide sur l'assemblage des protéines du VHB [99]. Cette étude a été réalisée par *nanofluidic resistive pulse sensing* (NRPS) afin de caractériser les distributions de taille des produits en temps réel. Les résultats ont ensuite été confirmés par des images de microscopie électronique en coloration négative. Si les intermédiaires de Zlotnick et al. sont structurés et stables, ceux de Kondylis et al. sont plus labiles et désordonnés (Fig. 5.4). En effet, les auteurs expliquent qu'une modification de l'équilibre de la réaction, i.e. une dilution de la solution réactionnelle, provoque un réarrangement rapide. Ils observent également que l'augmentation de l'énergie d'association des dimères peut conduire à la formation de capsides incomplètes. Des observations similaires avaient été effectuées par le même groupe auparavant [100].



**Figure 5.4:** Image de microscopie électronique en coloration négative des gros intermédiaires d'assemblage observés par Kondylis et al. (Adaptée de[99]).

### 1.1.3 Dimorphisme de la capside

Comme nous avons pu le voir dans le chapitre 3, la capside du virus de l'hépatite B est polymorphique. Les protéines peuvent s'associer afin de former des capsides  $T = 3$  ou  $T = 4$ . Cependant, ces dernières sont largement majoritaire *in vivo* et *in vitro* (> 90%). Une étude récente de Sun et al met en évidence la possibilité



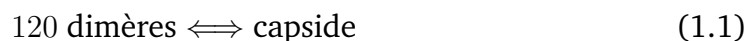
**Figure 5.5:** Distribution des capsides assemblées par titrage rapide au KCl. Les pourcentages analysés par densité optique après gradient de densité de sucrose sont présentés par les barres pleines. L'analyse statistique à partir des images de microscopie électronique est montrée par les barres rayées. (Adapté de[101]).

de déplacer cet équilibre, entre les deux formes de capside, à l'aide de cations tels que  $K^+$  ou  $Ca^{2+}$  [101]. Les expériences d'assemblage ont été réalisées par des mélanges, rapides ou lents, de solutions de dimères de protéines avec des solutions salines de différentes concentrations. Les résultats ont été caractérisés par gradient de densité de sucrose puis densité optique, et ont été confirmés par microscopie

électronique. Un exemple des données obtenues est présenté ci-dessous (Fig. 5.5). Les auteurs expliquent ce phénomène par le fait que des concentrations plus élevées en sel génèreraient une nucléation plus rapide en écrantant fortement la répulsion électrostatique entre les protéines. C'est ce phénomène qui aurait un impact sur la forme finale des capsides.

#### 1.1.4 Dynamique d'assemblage de la capside

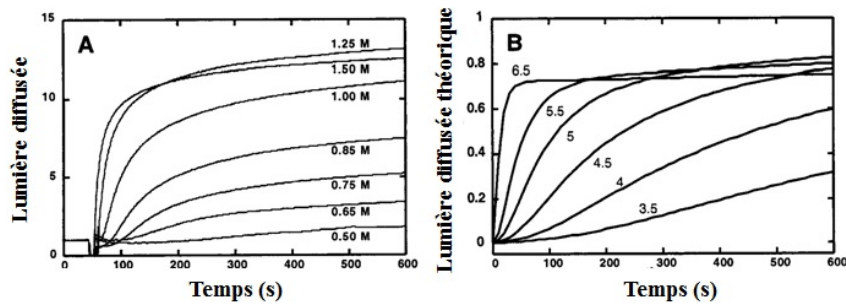
Les interactions entre les protéines jouent également un rôle dans la dynamique de l'assemblage de la capside composée des protéines CP149. En effet, dans une publication de 2002, Ceres et al. observent par chromatographie d'exclusion de taille l'absence d'assemblage pour de faibles concentrations initiales en dimères de protéines [102]. Les auteurs expliquent la présence d'une concentration pseudo-critique d'assemblage comme une conséquence naturelle de l'équation décrivant la réaction de formation de la capside :



Cette pseudo-concentration correspond à un régime où des concentrations molaires égales de dimères et de capsides coexistent. De ce fait, pour des concentrations supérieures à celle-ci, la concentration en dimères libres en solution est presque constante et la fraction massique finale en capside augmentera avec la concentration initiale en dimères. En dessous de cette concentration, l'assemblage n'est pas favorable.

De plus, le rôle des interactions entre les dimères de protéines est également mis en avant par Zlotnick et al. dans une publication en 1999 [95]. Dans celle-ci, les auteurs combinent deux techniques, la diffusion de lumière et la chromatographie d'exclusion de taille, afin d'observer la dynamique d'assemblage de la capside vide à différentes concentrations en sel et protéines. Leurs résultats mettent en avant une dépendance de la dynamique de l'assemblage aux concentrations en sel et protéines : les vitesses de réaction augmentent avec les concentrations. Les résultats sont associés à des simulations numériques faisant varier l'énergie d'association entre les dimères afin d'expliquer cette dépendance. Les auteurs décrivent des dynamiques similaires aux expériences : aux énergies d'association les plus élevées, les vitesses de réaction sont les plus élevées. Les auteurs notent





**Figure 5.6:** Comparaison de l'évolution des intensités diffusées en fonction du temps pour des expériences et des simulations sur l'assemblage de la capsid. A gauche, les courbes expérimentales : différentes concentrations en sel ont été utilisées pour induire l'assemblage de 5  $\mu\text{M}$  de VHB. A droite, les simulations numériques : différentes énergies d'association (en kilocalories par mole par contact). (Adapté de[95]).

que, pour la cinétique avec la plus forte énergie d'association, le maximum du signal est atténué et quasi-constant et attribuent ce comportement à la formation d'un piège cinétique.

## 1.2 Désassemblage des capsides vides du VHB

Si l'auto-assemblage des capsides est un sujet de recherche devenu populaire au cours des dernières décennies, l'étude du désassemblage reste marginale. Cependant, rien ne dit que la cinétique de désassemblage est équivalente à celle de l'auto-assemblage. Ceci est confirmé par des travaux expérimentaux réalisés par Singh et Zlotnick [26]. En utilisant la fluorescence intrinsèque, le dichroïsme circulaire et la chromatographie par exclusion de taille, les auteurs observent que l'énergie d'interaction apparente mesurée, entre les dimères, dans les expériences de désassemblage est beaucoup plus forte que lors de l'assemblage. Leurs auteurs remarquent également que, contrairement à l'assemblage, l'énergie d'association apparente varie systématiquement avec la concentration en réactifs. Afin d'expliquer ces observations, les auteurs ont réalisé des simulations d'assemblage de la capsid. Leurs calculs indiquent que la dissociation est entravée par des concentrations élevées d'intermédiaires ayant pour effet de favoriser le réassemblage malgré la tendance globale à la dissociation.

## 1.3 Conclusion et discussion

Ce chapitre met en évidence le fait que les dynamiques d'auto-assemblage et de désassemblage des capsides vides du VHB sont encore méconnues et débattues. Des approches expérimentales et théoriques ont été utilisées afin d'apporter des éléments de réponse à ces questions.

Concernant l'assemblage, si le mécanisme de nucléation-croissance et les intermédiaires précoces semblent faire consensus au sein de la communauté, la nature des intermédiaires tardifs reste encore incertaine. En effet, Zlotnick et al. ont observé des intermédiaires stables et structurés, tandis que Katen et al. mettent en évidence des intermédiaires labiles et désordonnés. De plus, la question de l'existence d'autres intermédiaires au cours de l'assemblage reste ouverte, les données cinétiques étant peu nombreuses.

De manière étonnante, Sun et al. mettent en évidence un impact de la salinité sur la morphologie finale des capsides qui n'avait pas été observé auparavant. Les auteurs expliquent ces observations par une modification de la vitesse de nucléation qui aurait comme conséquence de favoriser les capsides  $T = 3$ . L'hypothèse d'une modification des énergies d'interaction entre les dimères a également été utilisée pour expliquer la concentration critique d'assemblage et la dépendance de la dynamique aux concentrations en sel et protéines. Cependant, les études précédemment citées présentent plusieurs limites. Celles-ci sont soit réalisées à l'équilibre, soit ne disposent pas d'une résolution spatiale permettant de suivre ces processus, impliquant des espèces dont les tailles sont de l'ordre du nanomètre avec une résolution temporelle de l'ordre de la milliseconde.

Concernant le désassemblage, si très peu de travaux ont été réalisés, Singh et Zlotnick ont tout de même pu observer que la dissociation de la capside est entravée par des intermédiaires dont l'effet est de favoriser le réassemblage. Cependant, la nature de ces intermédiaires est encore méconnue. Ceci s'explique par le fait qu'une des difficultés majeures lors de l'étude des cinétiques de désassemblage est la coexistence de plusieurs espèces en solution sur plusieurs échelles de temps. Or, la technique de diffusion de rayons X aux petits angles résolue en temps répond parfaitement à cette problématique.

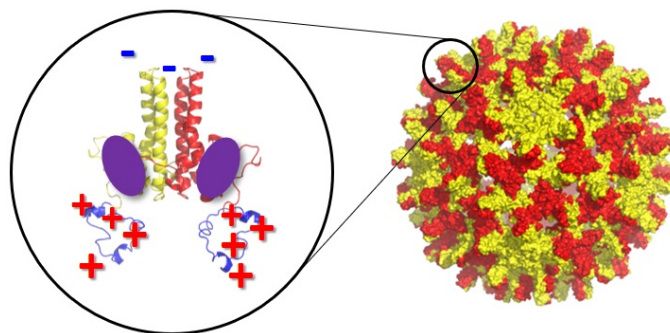
# Analyse par Thermo-fluorescence

Dans ce chapitre, nous étudions les interactions entre les protéines à travers la dissociation thermique des capsides. Pour ce faire, nous utilisons la technique de Criblage de Thermo-stabilité des Protéines par Fluorescence (CTPF) ainsi qu'un modèle théorique de champ moyen, présentés dans le chapitre 2. Cette approche nous permet de mesurer la température de dénaturation des capsides de VHB. Celle-ci étant directement liée aux interactions entre les protéines, nous pouvons estimer leurs énergies d'interaction.

## 2.1 Les particules virales

La stabilité des particules virales de VHB repose principalement sur trois interactions (voir Fig. 5.7) :

- une attraction à courte portée, due aux forces hydrophobes, qui maintient les sous-unités ensemble
- une répulsion électrostatique à longue distance permettant la dissociation de la capside
- une attraction électrostatique à longue portée entre les sous-unités et le génome, permettant un empaquetage efficace.



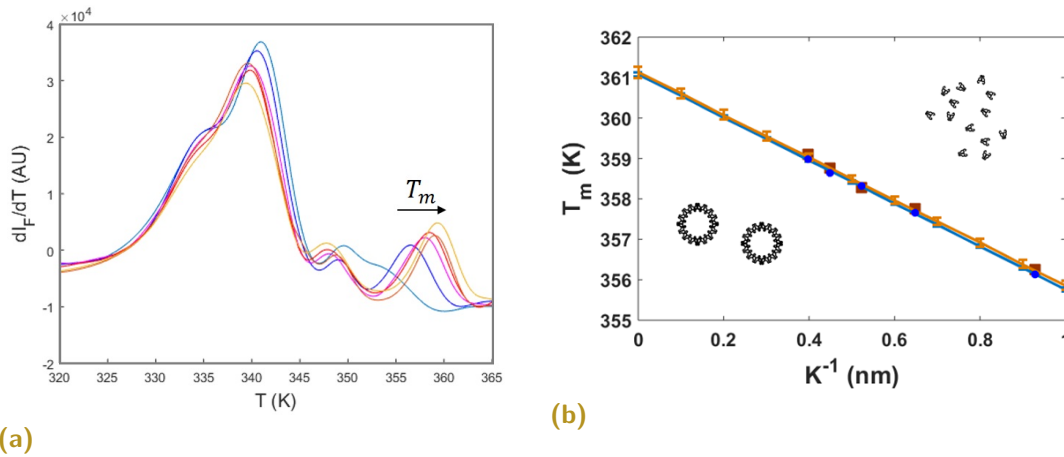
**Figure 5.7:** Capside du VHB (à droite) et un dimère de protéines (à gauche). Ce dernier possède des zones hydrophobes sur les faces latérales (taches violettes) ainsi que des charges positives (en rouge) le long des bras flexibles et des charges négatives (en bleu) au niveau du spicule.

Les énergies libres correspondantes sont méconnues. Certaines valeurs estimées ont été déduites de calculs théoriques [103] [82] ou indirectement d'expériences cinétiques notamment par diffusion statique de la lumière au cours du temps [104]. Concernant les mesures expérimentales, les capsides étaient vides et composées de la protéine tronquée (CP149) ou de mutants ponctuels de cette dernière. Dans ces cas, l'absence du bras flexible C-terminal permet de s'affranchir des interactions entre le génome et les protéines et de ne s'intéresser qu'aux interactions entre les protéines. En effet, la protéine tronquée (CP149) permet de former une capside identique à celle formée par la protéine native, sans pour autant empaqueter le génome. Une bonne connaissance des interactions entre les protéines étant essentielle à la compréhension du cycle de vie du VHB, nous avons choisi de réaliser des mesures complémentaires de celles présentées précédemment.

## 2.2 Température de dénaturation des capsides vides

Les résultats des tests de stabilité des capsides vides de VHB, i.e. composées uniquement du domaine d'assemblage, sont présentés sur la figure 5.8a. Celle-ci montre les données de fluorescence obtenues à pH 7,5 et à différentes concentrations en NaCl. Les températures de dénaturation mesurées sont rapportées sur la figure 5.8b en fonction de la longueur de Debye  $\kappa^{-1}$ .

Dans un premier temps, nous pouvons constater que les échantillons présentent des comportements relativement similaires jusqu'à 345 K. En effet, jusqu'à cette température, les dérivées de l'intensité de fluorescence  $dI_F/dT$  de chaque courbe augmentent de manière comparable, jusqu'à atteindre un premier maximum pour des valeurs de température autour de 341 K. Au-delà de cette température, les comportements divergent. La courbe d'évolution de la solution avec une concentration nulle en sel (courbe bleu clair) présente un double pic difficilement caractérisable. Les courbes d'évolution des solutions avec des concentrations en NaCl allant de 0,1 M (courbe bleu foncé) à 0,5 M (courbe jaune), par pas de 0,1 M, présentent un pic bien isolé entre 356 et 360 K. Il est également à noter que le maximum de ce pic se décale vers les plus hautes températures à mesure que la concentration en sel de la solution augmente. La présence de deux pics distincts peut s'expliquer par



**Figure 5.8:** Résultats de tests de stabilité thermique par fluorescence sur des capsides vides de VHB. **5.8a** La figure présente la dérivée de l'intensité de fluorescence  $I_F$  en fonction de la température à pH 7,5 et différentes salinités : 0 M (courbe bleu clair) à 0,5 M (courbe jaune) par pas de 0,1 M. **5.8b** La figure montre la température de dénaturation  $T_m$  en fonction de la longueur de Debye  $\kappa^{-1}$  à pH 7,5 et pour les mêmes salinités (cercles bleus et carrés marrons). Les lignes bleue et jaune sont des ajustements des données expérimentales par l'équation 2.2.

le fait que la capside est composée de deux types de dimères : les dimères A-B et les dimères C-D. Pour la capside du VHB, les dimères A-B forment 12 pentamères de dimères situés aux sommets de l'icosaèdre et sont souvent décrit comme l'un des sous-ensembles le plus stable de la capside. Les dimères C-D agissent quant à eux comme de la colle entre ces pentamères pour former une capside complète. Ainsi, les dimères C-D faiblement connectés seront d'abord dissociés lors de l'augmentation de la température tandis que la dissociation des pentamères, plus stables, aura lieu à une température plus élevée. En 2017, Chen et al. avait proposé cette hypothèse lors de leur étude, par simulation numérique, de la dissociation thermique de la capside virale du CCMV[105].

Les maximums de température, extraits des données de la figure 5.8a, sont reportés sur la figure 5.8b en fonction de la longueur de Debye  $\kappa^{-1}$ , qui dépend de la concentration en sel de la solution étudiée (voir chapitre 2). En ajustant les données à l'aide de l'équation 2.2, nous pouvons estimer une charge effective  $|Z| = 1,6e$  et une énergie d'interaction entre les sous-unités  $\epsilon = -6 k_B T$ , où  $k_B$  est la constante de Boltzmann et  $T$  la température ambiante (298 K).

## 2.2.1 Énergie d'interaction entre les protéines

Concernant la valeur estimée d'énergie d'interaction entre les dimères de protéines, celle-ci est raisonnable et proche des valeurs expérimentales précédemment estimées. En effet, l'énergie de liaison des sous-unités (par contact) a été mesurée autour de  $-3 \text{ kcal.mol}^{-1}$  soit environ  $-5 k_B T$  [104]. On retrouve des valeurs similaires,  $-2.9$  à  $-4.4 \text{ kcal.mol}^{-1}$ , soit environ  $-5$  à  $-8 k_B T$  pour des concentrations en NaCl variables chez Ceres et al. [102]. Par contre, il y a un net désaccord entre les prédictions théoriques et notre valeur estimée. En effet, nous pouvons trouver une valeur moyenne de l'estimation de l'énergie d'association publiée sur VIPER de l'ordre de  $-15 \text{ kcal.mol}^{-1}$  pour chaque sous-unité de la capsid et, Baker. et al. [106] donnent des valeurs autour de  $-9 \text{ kcal.mol}^{-1}$ . Ces valeurs équivalent à  $-25$  et  $-15 k_B T$  respectivement. La question des raisons de ces différences entre valeurs expérimentales et théoriques se pose. Bahadur et al. proposent dans leur discussion que, dans certains cas particuliers, les critères géométriques de sélection des interfaces de contact au sein de la capsid pourraient générer un paramétrage incorrect dans les calculs théoriques de l'énergie [107]. De plus, selon Horton et al. les valeurs expérimentales pourraient refléter un coût énergétique dû au changement de conformation des dimères lors de la formation de la capsid [108].

## 2.2.2 Charge effective des dimères de protéines au sein de la capsid

Dans le cas de la charge effective mesurée, il existe peu de points de comparaison dans la littérature. Sun et al. ont estimé, théoriquement, que la charge de la protéine tronquée devait être située entre  $-10e$  et  $-25e$  par dimère, pour des solutions ayant des  $pH$  allant de 6,6 à 9,6 [104]. Dans le même article, des expériences ont permis d'estimer des valeurs de la charge effective entre  $-14.5e$  et  $-28.8e$  par dimère pour des  $pH$  allant de 7,5 à 9,6. Ces valeurs sont similaires à celles obtenues par calculs théoriques. Cependant, ces mesures ont été réalisées par mobilité électrophorétique et, comme souligné par Vega-Acosta et al., il est difficile d'établir un lien entre la mobilité et la charge [83]. De plus, si la valeur de la charge obtenue par nos expériences peut sembler faible, il ne faut pas oublier

que nos mesures reflètent la charge des protéines au sein de la capside, tandis que la charge mesurée par Sun et al. correspond à celle des dimères libres.

## 2.3 Conclusion

En résumé, une série de mesures a été réalisée afin de déterminer la charge effective ainsi que l'énergie d'interaction au sein de la capside formée de la protéine CP149 du VHB. Nous avons pu estimer une charge effective de  $|Z| = 1,6e$  et une énergie d'interaction entre les dimères telle que  $\epsilon = -6 k_B T$ . La valeur estimée d'énergie d'interaction est raisonnable et proche des valeurs expérimentales estimées précédemment. La valeur estimée de la charge effective est relativement faible. Cependant, cette valeur est raisonnable. En effet, la formation de la capside est dirigée par une compétition entre la répulsion électrostatique et l'attraction hydrophobe. Une charge faible favorise l'assemblage en raison d'une répulsion électrostatique réduite. La connaissance des interactions au sein de la capside du VHB est fondamentale pour la compréhension du cycle de vie du virus.

# Désassemblage de la capsidie CP149

# 3

Dans ce chapitre, nous étudions le désassemblage de la capsidie vide, formée des protéines tronquées (CP149). Les expériences ont été réalisées par diffusion de rayons X aux petits angles résolue en temps (TR-SAXS) au synchrotron SOLEIL sur la ligne SWING et des expériences complémentaires ont été obtenues par diffusion statique de lumière au LPS. Un des objectifs de cette approche était d'obtenir des informations sur ce processus fort méconnu.

## 3.1 Expériences réalisées et données obtenues par TR-SAXS

Dans cette section, je présenterai la préparation des échantillons ainsi que les données brutes pour les données obtenues par TR-SAXS.

### 3.1.1 Préparation des échantillons

Les capsidies de CP149, utilisées pour l'ensemble des expériences, ont été exprimées et purifiées, selon les protocoles présentés en annexes. Ces étapes ont été réalisées à l'institut de Microbiologie Moléculaire et de Biochimie Structurale (MMSB) de Lyon, par Shishan Wang et Lauriane Lecoq. La préparation des échantillons, i.e. changement de solution tampon et de concentration, a été faite au LPS par moi-même. Les solutions tampons utilisées dans cette étude étaient :

- La solution tampon de stockage des capsidies, composée de : 50 mM CHES à pH 9, 1 mM DTT
- La solution tampon de désassemblage, composée de : 50 mM CHES à pH 9, 1 mM DTT + 6 M urée<sup>1</sup>

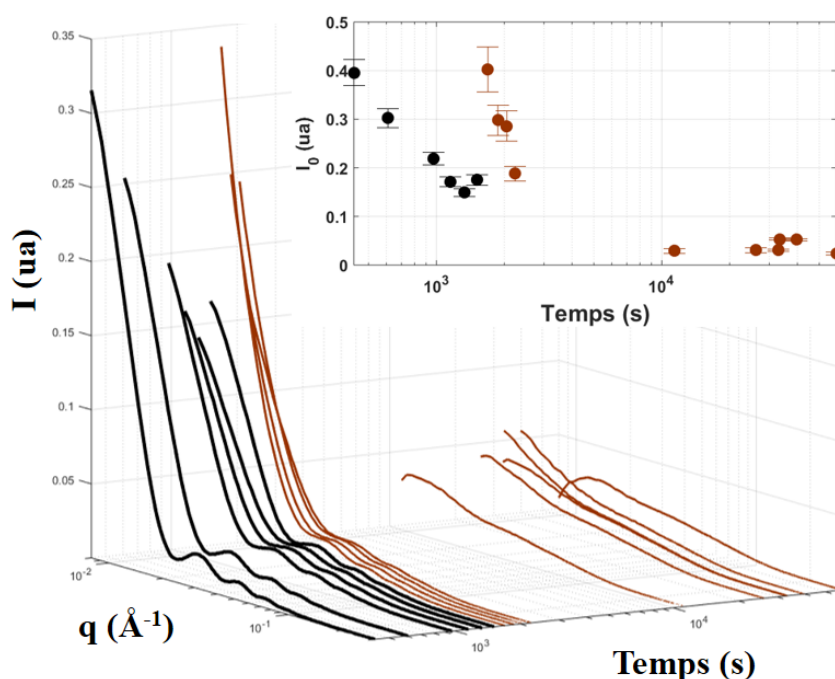
1. L'urée est un agent chaotrope, i.e. une molécule capable de casser les liaisons hydrogènes ce qui a pour effet de diminuer l'attraction hydrophobe.



La concentration en capside de la solution stock était de  $2,4 \text{ g.L}^{-1}$ . L'échantillon avait été dialysé en tampon de stockage, sur la nuit, deux jours avant les expériences à SOLEIL.

### 3.1.2 Présentation des données

Dans ce paragraphe, je présenterai les résultats obtenus pour le désassemblage des capsides vides initié par un mélange avec de l'urée. La concentration finale en protéines était de  $1,2 \text{ g.L}^{-1}$  et la concentration finale en urée était de 3 M.



**Figure 5.9:** Représentation tridimensionnelle des intensités mesurées de 7 min à 16h 30min. L'encart présente l'évolution des intensités  $I_{(q \rightarrow 0)}$  au cours de la cinétique.

Dans un premier temps, nous pouvons constater que la cinétique de désassemblage présentée sur la figure 5.9 met en évidence deux phases :

- Une première phase courte (courbes noires) qui dure environ 25 min. Cette phase montre une diminution lente des intensités sur l'ensemble de la période concernée.
- Une seconde phase plus longue (courbes rouges) qui dure plusieurs heures (environ 16h). Cette phase présente d'abord une remontée soudaine des intensités avant de montrer également une diminution lente des intensités.

Ces premières observations nous amènent à un découpage naturel de la cinétique afin de pouvoir retrouver les espèces présentes en solution durant la cinétique.

## 3.2 Reconstruction des espèces

Dans cette section, je présenterai l'analyse réalisée par SVD (voir section 3.19) sur la matrice des intensités de diffusion pour la première phase puis l'ensemble des données.

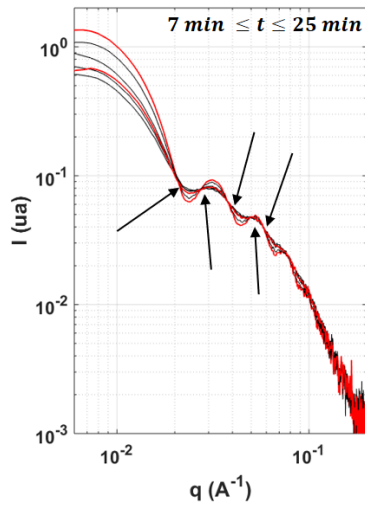
### 3.2.1 Première phase du désassemblage

La figure 5.10a montre l'évolution des courbes d'intensités au cours des 25 premières minutes de la cinétique. Sur celle-ci, les courbes superposées mettent en évidence des points isobestiques, i.e. des valeurs des nombres d'onde pour lesquels l'intensité de diffusion ne varie pas lors du désassemblage. Ces points suggèrent la présence de seulement deux espèces en solution, contribuant au signal, pour l'ensemble de la période concernée. Ces observations sont confirmées par les critères fournis par l'analyse par SVD présentés précédemment (voir sous-section 3.2.1.1). Nous pouvons voir seulement deux valeurs significatives, aussi bien sur le tableau d'auto-corrélations des vecteurs de base  $\vec{U}_i$  et  $\vec{V}_i$  que sur le graphique de comparaison des résidus<sup>2</sup> avec le bruit expérimental (Fig. 5.10b).

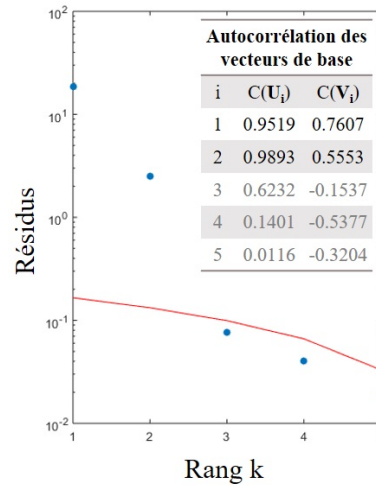
L'ensemble de ces éléments nous permet de conclure que deux espèces, uniquement, contribuent au signal pour la première phase. La figure 5.11 présente les courbes superposées des facteurs de forme des espèces en solution. Les courbes en traits pleins représentent les données reconstruites à partir de l'analyse par SVD. Les courbes en pointillés montrent les courbes expérimentales : la courbe rouge et la courbe bleue représentent les capsides et les dimères, respectivement. Nous pouvons noter un bon accord entre les facteurs de forme, ce qui indique que les deux espèces contribuant à la première phase sont la capside et le dimère.

---

2. Les résidus sont issus de l'évaluation de la matrice tronquée  $\tilde{I}$  par rapport à la matrice initiale  $I$



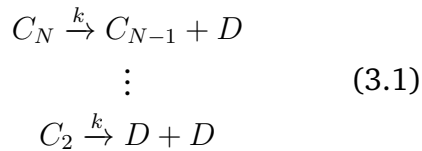
(a) Courbes superposées



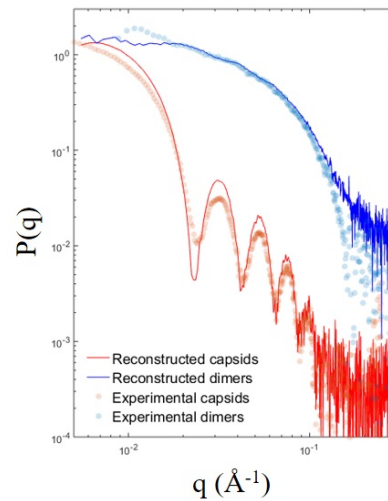
(b) Analyse par SVD

**Figure 5.10:** Données TR-SAXS pour la première phase du désassemblage de la capside vide du VHB. La figure 5.10a présente les courbes superposées des 25 premières minutes de la cinétique. Les courbes en rouge indiquent les premiers et derniers temps de la phase. Les flèches indiquent les points isobestiques. Les intensités sont en unités absolues. La figure 5.10b montre l'analyse par SVD des 25 premières minutes de la cinétique

Ces conclusions nous amènent à proposer un modèle de dépolymérisation pour décrire la réaction ayant lieu durant les 25 premières minutes tel que :



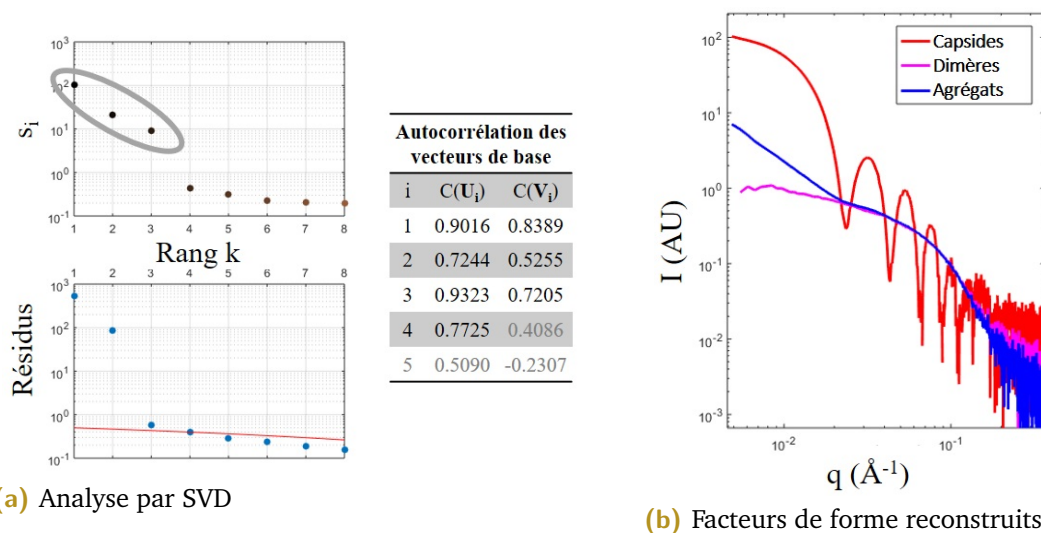
où  $C$  et  $D$  symbolisent la capside et le dimère, respectivement.  $N$  est le nombre de dimères composant la capside et  $k$  est la constante de vitesse de la réaction.



**Figure 5.11:** Graphique superposant les facteurs de forme des espèces expérimentales et reconstruites de la première phase de la cinétique.

### 3.2.2 Cinétique entière

Dans le but de pouvoir reconstruire l'ensemble de la cinétique, nous avons réalisé le même type d'analyse par SVD sur l'ensemble du jeu de données (Fig. 5.12a).



**Figure 5.12:** Données TR-SAXS pour la totalité de la cinétique de désassemblage des capsides vides. La figure 5.12a présente l'analyse par SVD de l'ensemble de la cinétique. La figure 5.12b montre les facteurs de forme reconstruits des espèces présentes en solution lors de la cinétique

La matrice originelle des intensités diffusées  $I$  pour la cinétique de désassemblage n'est composée que de 16 temps caractéristiques. Ce nombre, relativement faible pour des données de TR-SAXS, diminue grandement le rapport signal-sur-bruit et rend plus complexe la séparation du bruit des composantes significatives du signal. Sur la figure 5.12a, nous retrouvons le tableau d'auto-corrélation des vecteurs de base  $\vec{U}_i$  et  $\vec{V}_i$  (Fig. 5.12a, à droite), le graphique de comparaison des résidus avec le bruit expérimental (Fig. 5.12a, en bas à gauche) et le graphique des valeurs singulières de la matrice  $S$  (Fig. 5.12a, en haut à gauche). Si, pris séparément, chacun de ces critères pourrait indiquer un nombre d'espèces différentes (2 à 4), l'ensemble indique la présence de trois espèces au sein de la solution durant l'ensemble de la cinétique.

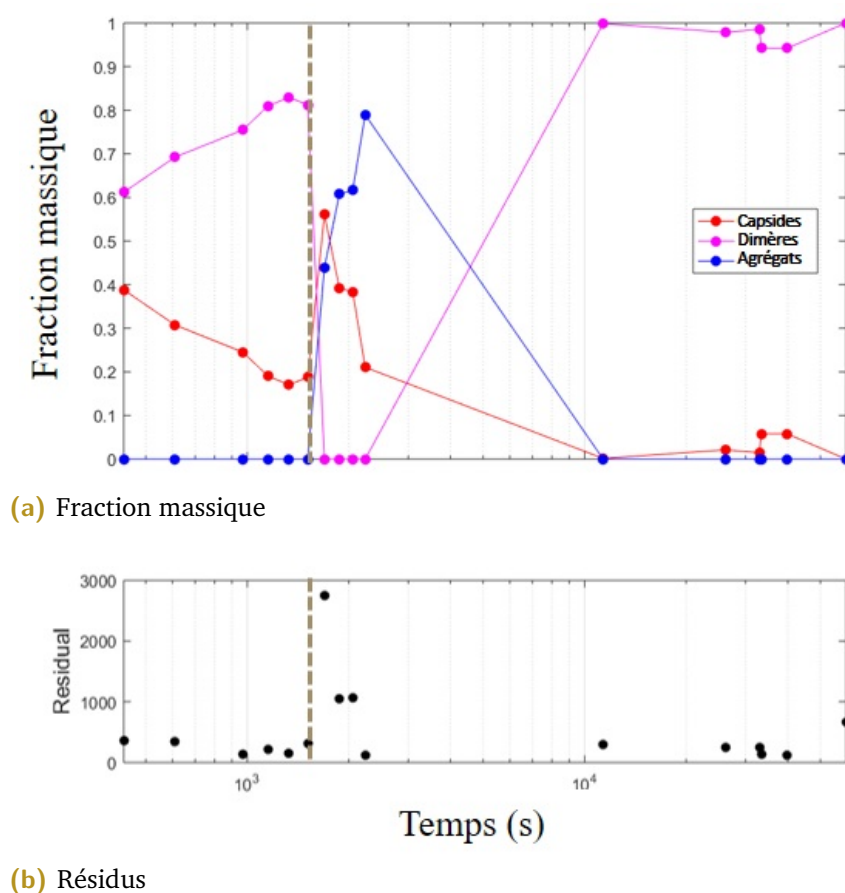
Du fait d'avoir pu identifier que les deux espèces de la première phase de la cinétique étaient la capside et le dimère, nous avons pu déterminer par SVD le facteur de forme de la troisième espèce, visible sur la figure 5.12b. Sur cette figure, nous pouvons voir le facteur de forme de la capside (en rouge), du dimère (en magenta) et de l'intermédiaire de désassemblage. La nature de cet intermédiaire sera discutée plus loin (section 3.3.2).

## 3.3 Evolution reconstruite des espèces

Dans cette section, je présenterai la reconstruction de l'évolution de la fraction massique des espèces en solution en fonction du temps à l'aide des facteurs de forme déterminés précédemment. Je la comparerai à la matrice initiale des intensités diffusées  $I$  et discuterai ensuite de l'intermédiaire de désassemblage.

### 3.3.1 Modèle de dépolymérisation

A l'aide de l'équation 3.25, présentée dans la section 3.2.2.2 et des facteurs de forme précédemment déterminés, nous avons reconstruit l'évolution de la concentration des espèces en solution au cours de la cinétique. Sur la figure 5.13a, nous pouvons voir l'évolution de la fraction massique des dimères (en magenta), de la capsidite (en rouge) et de l'intermédiaire (en bleu).



**Figure 5.13:** Évolution, en fonction du temps, de la fraction massique des espèces reconstruites par SVD (5.13a) et des résidus (5.13b).

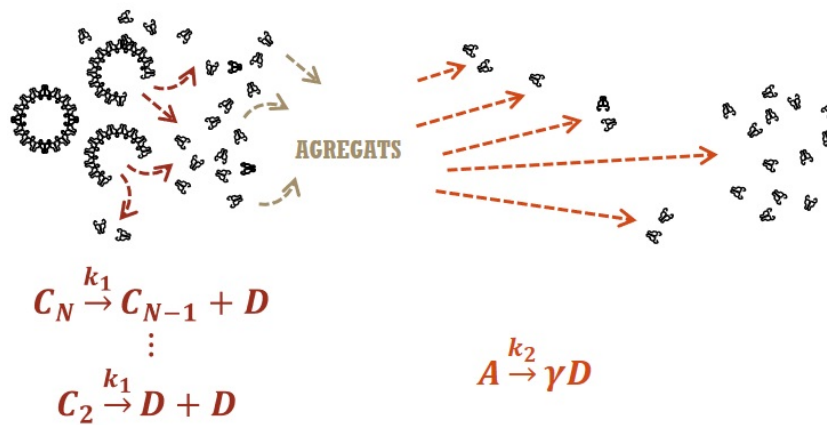
Dans un premier temps, nous pouvons constater la présence importante de dimères (environ 60 %) dès le début de la cinétique reconstruite. Ceci peut s'expliquer de deux manières :

- Le tampon de stockage des capsides (pH 9) était favorable aux dimères. Une partie des capsides s'est très certainement déjà dissociée avant le mélange avec de l'urée.
- Le premier point de notre cinétique expérimentale (et donc de notre reconstruction) se situe 7 minutes après l'initialisation du désassemblage par mélange avec de l'urée du fait de contraintes techniques. De ce fait, il est tout à fait possible que des capsides aient déjà entamé leur dissociation.

Ensuite, sur le graphique d'évolution de la fraction massique en fonction du temps, nous pouvons observer la même cinétique en deux temps que celle décrite dans la section 3.1. En effet, il existe une première phase dans cette cinétique où le pourcentage de dimères en solution, initialement à 60%, augmente jusqu'à atteindre 80% tandis que le pourcentage de capsides en solution évolue de manière exactement opposée. Les capsides se dissocient lentement en dimères. A l'issue de cette phase, une seconde commence par une augmentation soudaine de la fraction massique de l'intermédiaire de désassemblage associé à une diminution de celle des dimères. Cet événement est analogue à un phénomène d'agrégation rapide suivi d'une dissociation lente des agrégats en dimères.

En dessous, nous pouvons voir l'évolution des résidus, issus de l'évaluation de la matrice des intensités reconstruites  $I_{th}$  par rapport à la matrice initiale  $I$ , en fonction du temps (Fig. 5.13b). Les valeurs relativement faibles des résidus indiquent un assez bon accord entre les cinétiques reconstruite et expérimentale sauf au début du processus d'agrégation.

Nous pouvons alors résumer la cinétique à l'aide du schéma présenté en figure 5.14. Sur celui-ci, la première phase peut être décrite par un modèle de dépolymérisation où les capsides se dissocient lentement en dimères (flèches et équations en rouge). La phase d'agrégation, en beige, n'est pas schématisée puisque la nature de l'intermédiaire est encore inconnue et sera discutée dans la section suivante. La dernière phase correspond à une lente dissociation des agrégats en dimères (flèches et équations en orange).



**Figure 5.14:** Illustration schématique partielle de l'évolution des espèces au cours de la cinétique de désassemblage.

### 3.3.2 Nature des agrégats

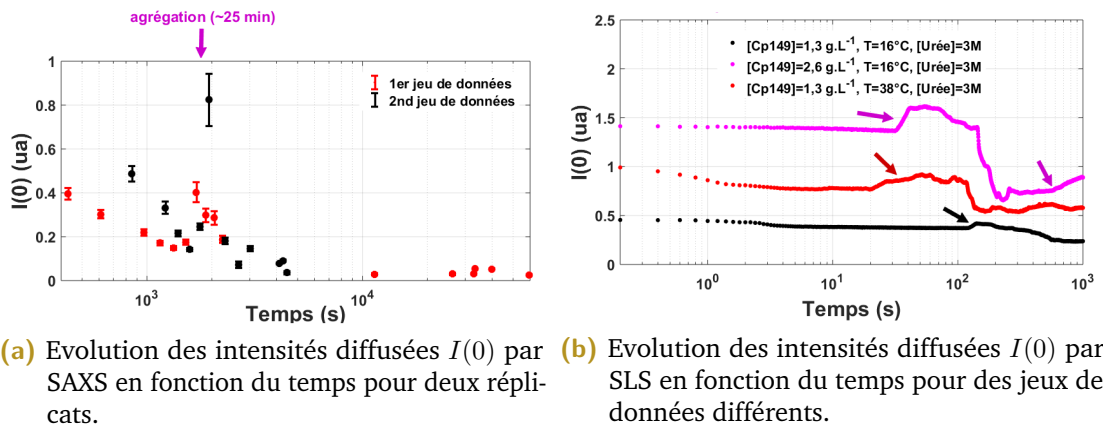
Si le modèle permettant de décrire la cinétique semble correct, il reste encore une question : quelle est la nature des agrégats qui apparaissent au cours de ce désassemblage ? Il existe plusieurs interrogations liées à la nature de cette agrégation :

- L'agrégation est-elle reproductible ou provient-elle d'un "accident" au cours de l'expérience présentée ?
- L'agrégation est-elle due à des dégâts d'irradiation ?

#### 3.3.2.1 Reproductibilité du phénomène d'agrégation

La figure 5.15a montre l'évolution des intensités de rayons X diffusées en fonction du temps pour deux jeux de données identiques : en bleu, le jeu de données précédemment présenté et, en noir, son réplique. Nous pouvons observer sur ce graphique que les deux cinétiques suivent la même dynamique. En effet, nous retrouvons une diminution lente des intensités dans les premières minutes puis une augmentation soudaine de l'intensité autour de 25 min après le mélange. La reproductibilité du processus d'agrégation avec la même technique d'observation montre que celui-ci n'est pas artéfactuel, mais bien révélateur d'un évènement survenant lors de la cinétique.

La figure 5.15b montre l'évolution des intensités de lumière visible diffusées en fonction du temps pour plusieurs jeux de données. Les données présentées sur



**Figure 5.15:** Illustration des essais visant à étudier la nature de l'intermédiaire de désassemblage

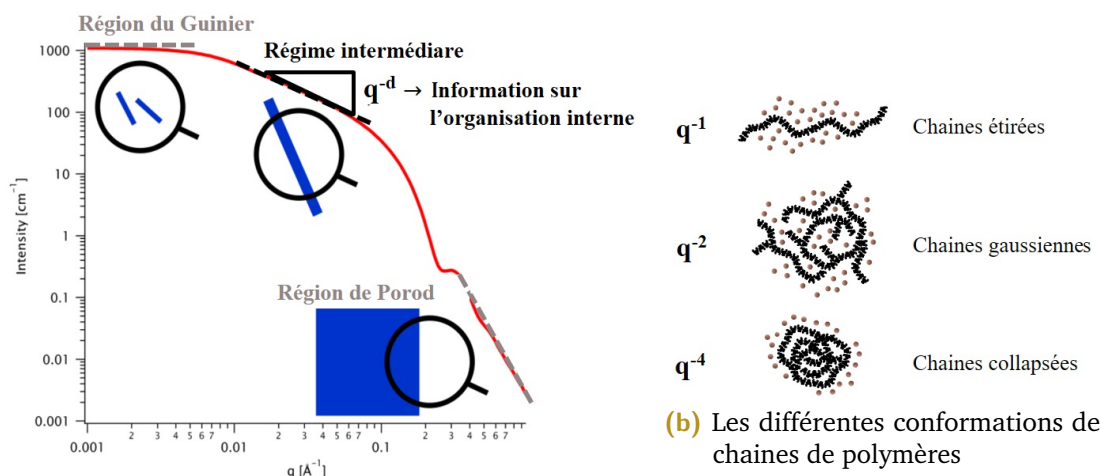
le graphique ont été obtenues par diffusion statique de lumière (SLS), i.e. une méthode non-destructrice qui permet de suivre l'évolution du poids moléculaire moyen de la solution en fonction du temps. Sur le graphique, nous pouvons constater que les cinétiques présentent toutes une augmentation de l'intensité mais qui se produit à différents instants après le mélange. Cette variabilité temporelle peut s'expliquer par le fait que les cinétiques ont été réalisées à différentes concentrations en urée et en capsides de VHB. Cependant, la reproductibilité du processus d'agrégation avec une technique non-destructive montre que celui-ci n'est pas dû à des dégâts d'irradiation mais bien révélateur d'un évènement issu du mélange.

### 3.3.2.2 Ordre interne des agrégats

Habituellement, lorsque nous voulons étudier des données obtenues par TR-SAXS, nous utilisons la région du Guinier, qui permet d'estimer des valeurs telles que le rayon de giration ou l'intensité lorsque le nombre d'onde  $q$  tend vers zéro (voir sous-section 1.5). Dans le cas présent, cette région caractérisée par un plateau, n'est pas accessible. En effet, la taille des objets est trop grande pour la gamme de  $q$  accessible. Il n'est donc pas possible d'obtenir ces informations. Néanmoins, si l'on s'intéresse à la région dite "intermédiaire" présentée sur la figure 5.16a, il est possible d'obtenir des informations sur l'organisation interne de ces agrégats.

En effet, la pente d'une courbe tracée en échelle log-log donnera des informations à propos de la nature fractale de l'objet. Dans le cas des polymères, en régime intermédiaire, lorsque l'intensité est proportionnelle à  $q^{-1}$ , le terme de chaînes





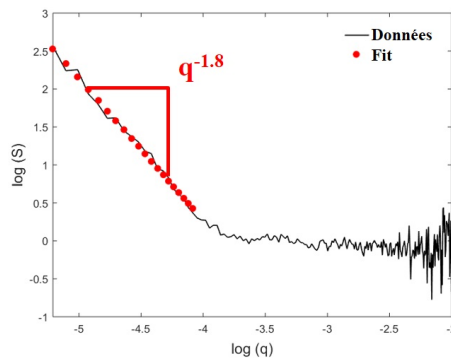
(a) Les différents régimes d'une courbe de SAXS. Adapté de [109].

**Figure 5.16:** Illustration schématique des informations obtenues à partir d'une courbe SAXS. La figure 5.16a présente un exemple de courbe obtenue par SAXS (en rouge). Chaque régime agit comme une loupe : aux petites valeurs de  $q$ , le régime du Guinier sonde la taille et la masse des objets. Le régime intermédiaire sonde l'organisation interne des objets. Aux grandes valeurs de  $q$ , le régime de Porod vient de la diffusion issue des sous-unités formant les complexes et sonde la rugosité des interfaces. La figure 5.16b montre une représentation schématique de la signification de la dimension fractale pour des chaînes de polymères

étirées peut être utilisé. Dans ce cas, les chaînes s'étendent pour favoriser l'interaction avec la solution. Lorsque la pente est de l'ordre de  $-2$ , les chaînes sont dites gaussiennes. Dans ce cas, les interactions attractives sont contrebalancées par les interactions répulsives, et inversement, du fait de la solution. Lorsque la pente est de l'ordre de  $q^{-4}$ , le terme de chaînes compactes est employé. Dans ce cas, les chaînes collapsent pour éjecter la solution et former un objet globulaire et compact. (5.16b).

Dans notre cas, nous ne parlons pas de polymères mais d'une structure où les protéines se sont agrégées. Cependant, nous pouvons nous inspirer de cette approche pour investiguer l'organisation interne de l'intermédiaire de désassemblage. Pour cela, nous allons nous intéresser au facteur de structure des agrégats. Celui-ci a été calculé en divisant son intensité mesurée par le facteur de forme des dimères, comme expliqué dans la section 1.3 et est présenté sur la figure 5.17. Sur celle-ci, nous pouvons tout d'abord constater une augmentation aux petites valeurs de  $q$  qui indique une interaction attractive, typique d'un phénomène d'agrégation. C'est cette zone de la courbe qui a été ajustée en échelle log-log et nous a permis de

déterminer une dimension fractale de -1,8. Cette valeur suggère que les agrégats peuvent être décrits comme des structures fractales branchées.

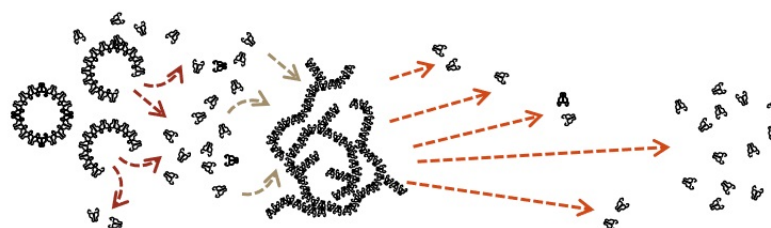


**Figure 5.17:** Facteur de structure des agrégats (courbe noire) et son ajustement (points rouges).

Concernant l'origine de cette agrégation, si l'urée est une molécule capable de dissocier la capsidie par diminution de l'attraction hydrophobe, celle-ci est également capable de casser les liaisons hydrogènes et donc de déstructurer une protéine. Le changement de conformation des protéines peut entraîner une modification des interactions entre les dimères libres en solution ce qui peut générer de l'agrégation.

### 3.4 Conclusion sur le désassemblage

Dans les conditions expérimentales de l'étude, nous avons pu identifier une cinétique de désassemblage, de la capsidie vide du VHB, avec deux phases : une première d'environ 25 min et une seconde de plusieurs heures. Au cours de cette cinétique, nous avons pu observer trois espèces en solution dont un intermédiaire de désassemblage. L'ensemble donne le schéma suivant :



**Figure 5.18:** Illustration schématique de l'évolution des espèces au cours de la cinétique de désassemblage.

Les capsides se dissocient en dimères via un processus de dépolymérisation. Les dimères libres en solution s'agglomèrent ensuite très rapidement pour former des agrégats pouvant être décrits comme des structures fractales branchées. La dernière étape est simplement la dissociation de ces agrégats en dimères.

# Assemblage du VHB

Dans ce chapitre, nous étudions l'auto-assemblage de la capsid vide du VHB, formée des protéines tronquées (CP149). Les expériences ont été réalisées par diffusion de rayons X aux petits angles résolue en temps à l'ESRF sur la ligne ID02. Un des objectifs de cette approche était d'obtenir des informations sur ce processus pour différentes salinités et différentes concentrations en protéines.

## 4.1 Données expérimentales

Dans cette section, je présenterai la préparation des échantillons ainsi que les conditions expérimentales des cinétiques d'assemblages réalisées. Je ferai également des remarques succinctes concernant les jeux de données ainsi obtenus.

### 4.1.1 Préparation des échantillons

Les capsides formées des protéines CP149, et utilisées pour l'ensemble des expériences, ont été exprimées et purifiées selon les protocoles présentés en annexes. Ces étapes ont été réalisées au MMSB de Lyon, par Shishan Wang et Lauriane Lecoq. Les solutions tampons et les échantillons ont été également préparés par Shishan Wang et Lauriane Lecoq. Les solutions tampons utilisées dans cette étude étaient :

- La solution tampon de stockage des dimères, composée de : 50 mM CHES à pH 9, 1 mM DTT
- La solution tampon d'assemblage, composée de : 50 mM CHES à pH 9, 1 mM DTT + 2 M NaCl.

La concentration en dimères de la solution stock était de  $1,17 \text{ g.L}^{-1}$  avec une pureté telle que  $A_{260}/A_{280} = 0,57$ . L'échantillon avait été dialysé en tampon de stockage, sur la nuit, deux jours avant les expériences à l'ESRF.

## 4.1.2 Les cinétiques réalisées par TR-SAXS

A partir des solutions précédemment présentées, plusieurs conditions expérimentales ont pu être explorées afin d'étudier l'assemblage de la capsid vide du VHB. Ceci a pu être réalisé en faisant varier les volumes injectés des solutions tampons de stockage et d'assemblage, ainsi que celui de l'échantillon. L'ensemble de ces cinétiques, réalisées à  $T=20^\circ\text{C}$ , est résumé dans le tableau ci-dessous (5.1).

CONDITIONS EXPÉRIMENTALES						
Concentration en protéines ( $\text{g}\cdot\text{L}^{-1}$ )	0,3		0,6		0,9	
Concentration en NaCl (M)	0,5*	1*	1,5*	0,5	1*	0,5

**Table 5.1:** Tableau récapitulatif des conditions expérimentales explorées pour l'étude de l'assemblage de la capsid vide du VHB. Les astérisques sont associés aux conditions ayant été réalisées plusieurs fois.

La distance échantillon-détecteur était de 1,3 ou 3 mètres en fonction des cinétiques.

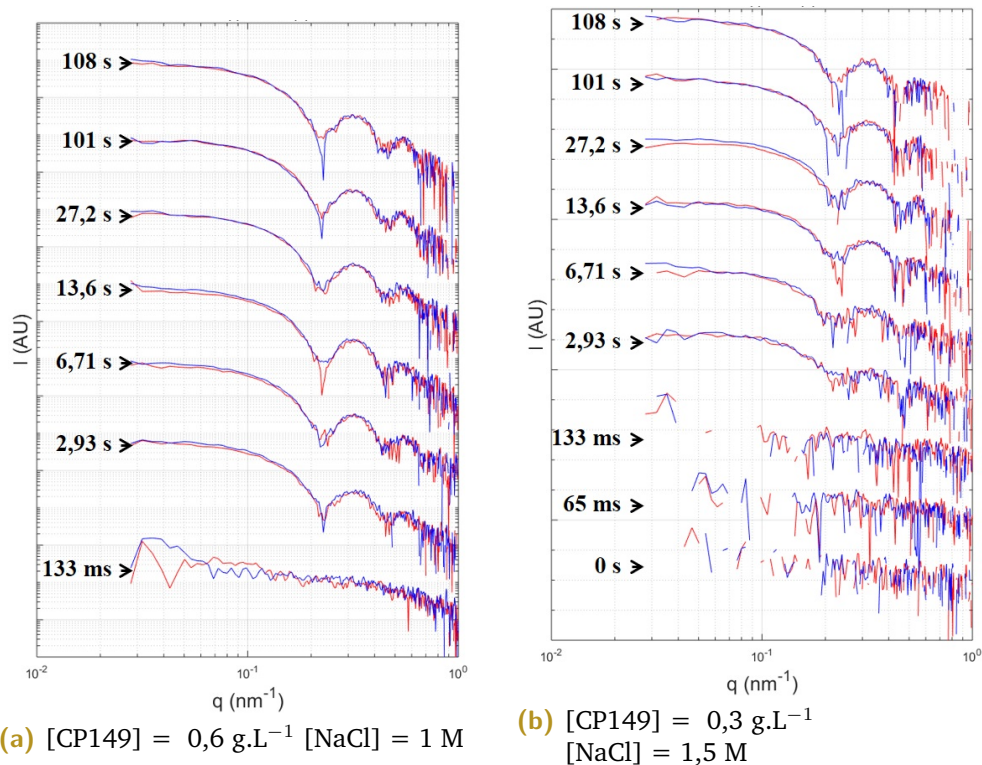
## 4.1.3 Les données

Au total, six cinétiques d'assemblages à différentes concentrations en NaCl et protéines ont été réalisées. Les données ont été collectées sur différents intervalles de temps, allant de 6 minutes pour le plus court à 19 minutes pour le plus long.

### 4.1.3.1 Reproductibilité des données

La figure 5.19 présente deux graphiques, chacun montrant les courbes superposées de cinétiques d'assemblages identiques. Le graphique 5.19a montre les cinétiques réalisées avec une concentration en protéine de  $0,6\text{ g}\cdot\text{L}^{-1}$  et une concentration en NaCl de 1 M. Le graphique 5.19b montre les cinétiques réalisées avec une concentration en protéine de  $0,3\text{ g}\cdot\text{L}^{-1}$  et une concentration en NaCl de 1,5 M. Pour les deux graphiques, les courbes rouges correspondent à une cinétique obtenue sur un intervalle de 6 min, tandis que les courbes bleues correspondent à une cinétique obtenue sur un intervalle de 9 min. Sur ces graphiques, ne sont présentées que les mesures réalisées aux mêmes temps d'acquisition pour plus de clarté.

Sur cette figure, nous pouvons voir une très bonne reproductibilité des expériences. Nous pouvons tout de même noter quelques divergences dans les valeurs ponctuelles lorsque  $q < 4 \cdot 10^{-2} \text{ nm}^{-1}$  ou  $q > 6 \cdot 10^{-1} \text{ nm}^{-1}$ . Celles-ci peuvent s'expliquer



**Figure 5.19:** Graphiques superposant les courbes obtenues lors des cinétiques d'assemblages réalisées par TR-SAXS. A gauche : réplicat d'une cinétique avec un faible rapport entre la concentration en sel et la concentration en protéine. A droite : réplicat d'une cinétique avec un fort rapport entre la concentration en sel et la concentration en protéine.

soit par le faible nombre de points de mesures pour le premier cas ou par des imperfections issues de la soustraction du tampon pour le second cas.

#### 4.1.3.2 Impact des concentrations en protéines et en sel sur la qualité des données

Il peut également être noté, sur la figure 5.19, la différence de qualité des courbes obtenues entre les deux jeux de données. A gauche, les cinétiques présentées ont un faible rapport entre la concentration en sel et la concentration en protéine tandis qu'à droite, les cinétiques ont un fort rapport entre la concentration en sel et la concentration en protéine. Dans la section 1.2 de la première partie du manuscrit, j'avais abordé les notions de contraste et de densité de longueur diffusée

( $\Delta\rho = \rho(\vec{r}) - \rho_s$ ). Celles-ci mettaient en avant l'importance de mesurer les solvants seuls afin de pouvoir soustraire leur contribution et étudier uniquement le signal des molécules d'intérêt. En sachant que  $\rho(\vec{r})$  diminue avec la concentration en protéines et que  $\rho_s$  augmente avec la concentration en sel, il devient évident que le ratio de la concentration en protéines vis-à-vis de la concentration en sel aura un impact majeur sur la qualité des courbes obtenues et donc sur le traitement et l'analyse qui s'en suivra.

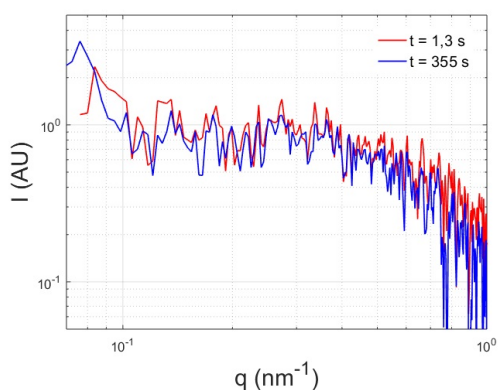
## 4.2 Premières analyses des données de TR-SAXS

Dans cette section, je présenterai les analyses réalisées afin de déterminer le nombre ainsi que l'évolution des concentrations des espèces en solution durant chacune des cinétiques d'assemblage présentées précédemment. Pour ce faire, je commencerai par présenter les résultats obtenus par la décomposition en valeurs singulières avant de les comparer avec la littérature existante. Je présenterai ensuite les reconstructions des cinétiques d'assemblage et discuterai de l'influence de la concentration en NaCl et en protéines sur les fractions finales en capsides.

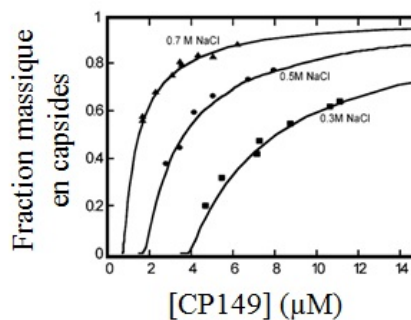
### 4.2.1 Concentration critique d'assemblage

Avant d'aborder l'estimation du nombre d'espèces contribuant au signal pour chaque cinétique, nous pouvons d'abord nous arrêter sur la cinétique d'assemblage avec une concentration en protéines de  $0,3 \text{ g.L}^{-1}$  et une concentration en sel de  $0,5 \text{ M}$ . La première et la dernière courbe de cette cinétique sont représentées sur la figure 5.20a. On remarque clairement qu'il n'y a pas eu d'assemblage, dans ces conditions, dans l'intervalle de temps étudié.

Ce résultat n'est pas surprenant. En effet, comme Ceres et al. l'avaient prédit en 2002, il n'y a pas d'assemblage possible pour une concentration en dimères de CP149 trop faible [102]. Cependant, comme nous pouvons le voir sur le graphique 5.20b, pour des concentrations en NaCl et protéines similaires aux nôtres, leurs expériences conduisent à une fraction massique finale en capsides d'environ 0,8.



(a) Courbes superposées.



(b) Graphique résumé des fractions massiques de capsides obtenus en fonction de la concentration en dimères. Adapté de [102]

**Figure 5.20:** Figures illustrant l'influence de la concentration en protéines et en NaCl sur l'assemblage des capsides vides de CP149. La figure 5.20a présente la première et dernière courbe de la cinétique d'assemblage avec une concentration en protéines et en NaCl de  $0,3 \text{ g.L}^{-1}$  et  $0,5 \text{ M}$ , respectivement. La figure 5.20b montre l'évolution de la fraction massique de capsides formées au bout de 24h en fonction de la concentration en dimères. L'assemblage a été induit par addition de  $0,3 \text{ M}$  (carrés),  $0,5 \text{ M}$  (cercles) ou  $0,7 \text{ M}$  (triangles) de NaCl. Les courbes sont tracées à partir l'équation d'équilibre de formation des capsides :  $K_{capsides} = [capsides]/[dimères]$ . (Adapté de [102])

Plusieurs différences dans le mode opératoire peuvent cependant expliquer ces résultats contradictoires :

- La première explication possible est l'intervalle de temps entre le mélange déclenchant l'assemblage et l'observation. En effet, nos cinétiques ont été réalisées sur un intervalle de temps de 6 minutes, ce qui est très court, tandis que les observations de Ceres et al. ont été réalisées 24 heures après le mélange.
- La seconde explication porte sur la différence de  $pH$  entre les deux types d'expériences : celui de nos solutions tampon est situé à 9 tandis que celui des expériences de Ceres et al. est situé à 7,5. Or, les groupes carboxyles des protéines sont plus ionisés à  $pH$  9 qu'à  $pH$  7,5. Il y a donc une densité de charges négatives plus importante à  $pH$  9. De plus, toujours dans cette même étude, les auteurs avaient estimé que la réaction d'assemblage était contrôlée par entropie, en accord avec les contacts en grande partie hydrophobes trouvés dans la structure cristalline. Une augmentation de la force électrostatique de répulsion entre les protéines, alors que les interactions entre les dimères sont faibles ( $\epsilon = -5,9 k_B T$ ), pourrait modifier le rapport entre les différentes



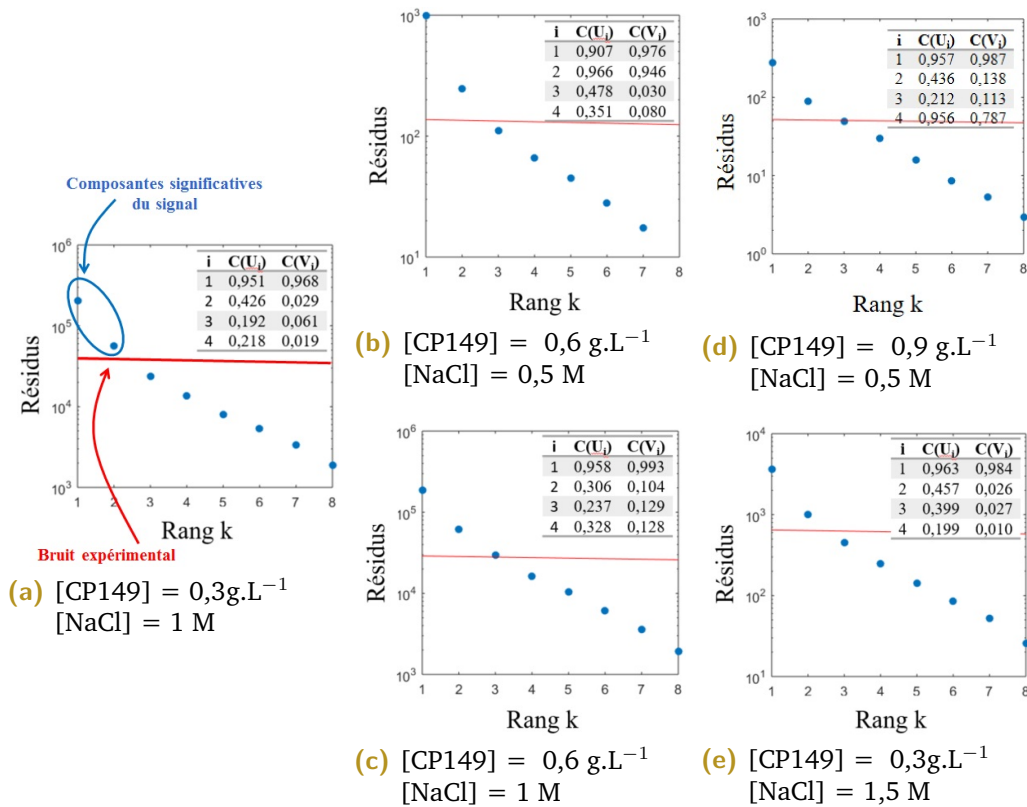
forces mises en jeu dans le processus et expliquer ces différences dans les résultats.

## 4.2.2 Analyse par SVD

Comme nous avons pu le constater sur la figure 5.19, les jeux de données ont un rapport signal-sur-bruit faible du fait des conditions expérimentales. Il est alors difficile d'estimer la présence de points isobestiques qui pourraient suggérer la présence de seulement deux espèces contribuant significativement au signal. De ce fait, la première analyse a été de réaliser une SVD comme présentée précédemment (section 3.2.1.1).

Les résultats de ces analyses par SVD sont présentés sur la figure 5.21. Comme indiqué précédemment dans la sous-section 4.1.3.1, les réplicats pour chacune des conditions expérimentales montrent une bonne reproductibilité permettant ainsi de ne présenter qu'une seule analyse par réplicat. Ainsi, chaque sous-figure représente une condition expérimentale et chacune des sous-figures présente le tableau d'auto-corrélation des vecteurs de base  $\vec{U}_i$  et  $\vec{V}_i$  ainsi que le graphique de comparaison des résidus avec le bruit expérimental de chacune de ces conditions.

Sur la sous-figure 5.21b, nous pouvons observer deux valeurs significatives, aussi bien sur le tableau des auto-corrélations des vecteurs de base  $\vec{U}_i$  et  $\vec{V}_i$  que sur le graphique de comparaison des résidus avec le bruit expérimental. Sur les autres sous-figures, seul le graphique de comparaison des résidus avec le bruit expérimental indique deux valeurs significatives. Les tableaux des auto-corrélations des vecteurs de base des matrices  $U$  et  $V$  n'indiquent qu'une seule valeur significative. Cependant, compte tenu de l'évolution des cinétiques d'assemblage, il est raisonnable d'estimer qu'il existe au moins deux espèces contribuant au signal, i.e. la capsid et le dimère. Les valeurs très similaires, à partir des deuxièmes rangs des tableaux des auto-corrélation, peuvent s'expliquer par la difficulté à séparer les contributions significatives du bruit du fait du faible contraste de densité de longueur diffusée de nos conditions expérimentales. Ainsi, l'ensemble de ces éléments nous permet de conclure que seules deux espèces contribuent significativement au signal pour les intervalles étudiés.

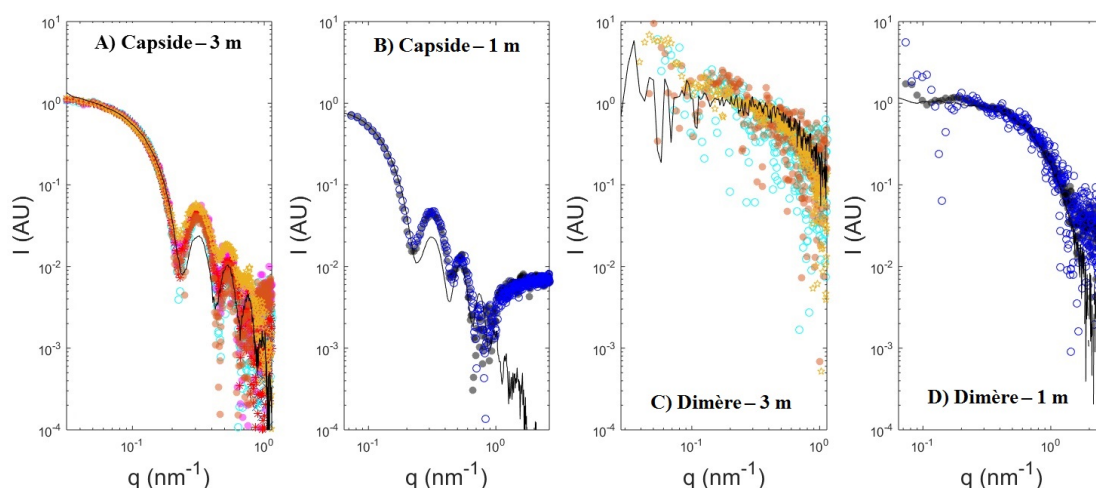


**Figure 5.21:** Présentation résumé des critères de SVD utilisés afin de déterminer le nombre de composantes suffisantes pour décrire les données pour chaque condition expérimentale.

### 4.2.3 Comparaison des espèces reconstruites et expérimentales

A partir de ces conclusions, nous avons utilisé l'analyse par SVD pour reconstruire ces espèces. Sur la figure 5.22, les graphiques présentent les courbes superposées de facteurs de forme : les courbes discontinues représentent les données reconstruites à partir de l'analyse par SVD, les courbes en trait plein les courbes expérimentales des échantillons de capsides ou de dimères.

Malgré des signaux parfois très bruités, notamment les espèces reconstruites et présentées sur le deuxième graphique en partant de la droite, nous pouvons noter un bon accord entre les facteurs de formes reconstruits et expérimentaux ce qui indique que les deux espèces contribuant significativement aux différentes cinétiques sont effectivement la capside et le dimère.

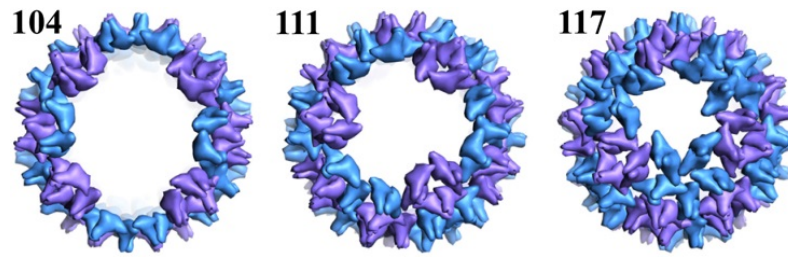


**Figure 5.22:** Comparaison des facteurs de formes des espèces reconstruites à partir de l'analyse par SVD avec les courbes expérimentales mesurées sur des échantillons purs en capsides et en dimères. Sur la figure, de gauche à droite, nous pouvons voir : A) les facteurs de forme des premières espèces reconstruites et comparées avec celui de la capside. La distance échantillon-détecteur est de 3 mètres. B) Les facteurs de forme des premières espèces reconstruites et comparées avec celui de la capside. La distance échantillon-détecteur est de 1,3 mètres. C) Les facteurs de forme des secondes espèces reconstruites et comparées avec celui du dimère. La distance échantillon-détecteur est de 3 mètres. D) Les facteurs de forme des secondes espèces reconstruites et comparées avec celui du dimère. La distance échantillon-détecteur est de 1,3 mètres.

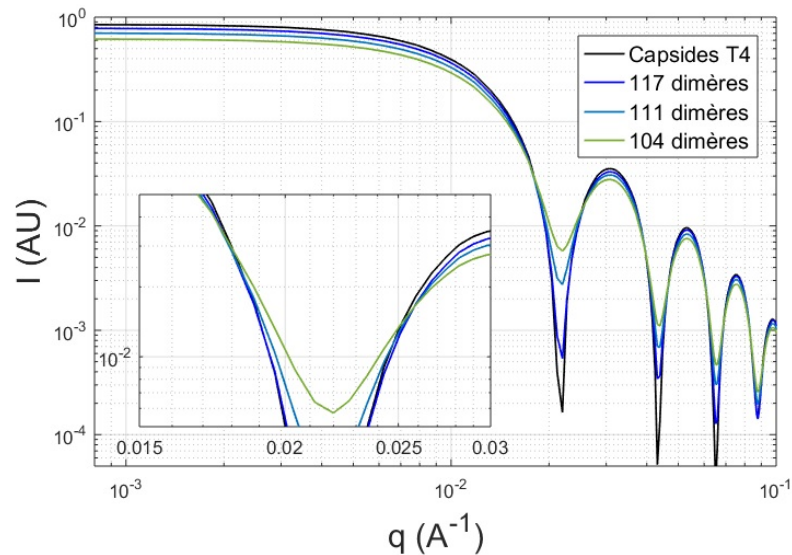
Cette conclusion peut sembler étonnante. En effet, en 2014, Pierson et al. publiaient un article indiquant la présence de gros intermédiaires composés de 104 à 117 dimères lors d'assemblages étudiés par spectrométrie de masse à détection de charge [98]. Ces conclusions étaient appuyées par des images de cryo-microscopie indiquant la présence de capsides incomplètes. Ces intermédiaires sont présentés sur la figure 5.23a.

A l'aide du logiciel PyMOL<sup>1</sup>, nous avons reconstitué les modèles d'intermédiaires à partir de la structure de la capside du VHB 1QGT disponible dans la Protein Data Bank<sup>2</sup>[36]. Les intensités diffusées ont ensuite été calculées à l'aide du logiciel CRY SOL de la suite ATSAS<sup>3</sup> [86]. Ces intensités diffusées théoriques sont présentées sur la figure 5.23b et comparées avec celle de la capside  $T = 4$ .

1. Logiciel libre de visualisation de structures chimiques en 3D.  
 2. Collection mondiale de données sur la structure tridimensionnelle de macromolécules biologiques.  
 3. Suite de programmes pour l'analyse des données de diffusion aux petits angles obtenues à partir de macromolécules biologiques



(a) Modèles possibles des structures des intermédiaires (Adaptées de[98]).



(b) Courbes superposées des intensités diffusées de la capside  $T = 4$  et des intermédiaires.

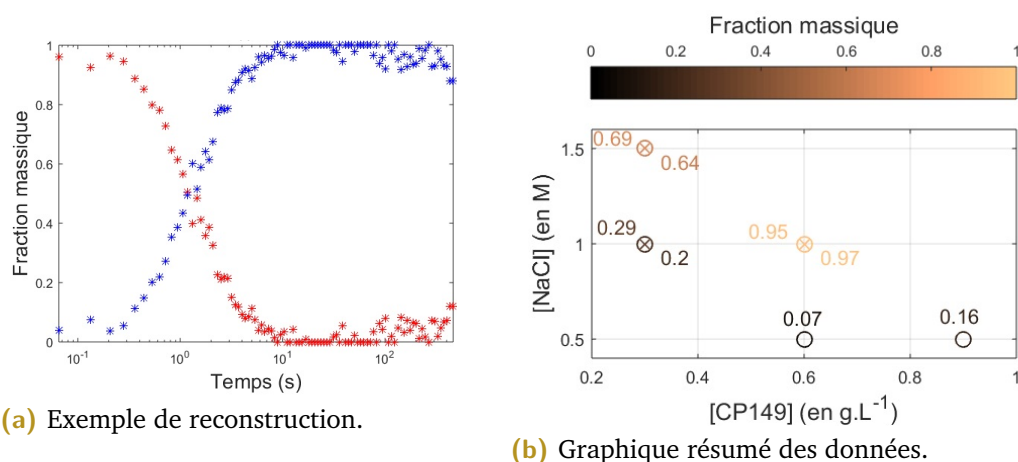
**Figure 5.23:** Mise en évidence de la similitude des courbes d'intensités diffusées théoriques des intermédiaires identifiés par Pierson et al. avec celle de la capside  $T = 4$ .

La figure 5.23b met en évidence la similitude des intensités diffusées entre les intermédiaires identifiés par Pierson et al. et l'intensité diffusée par la capside  $T = 4$ . En effet, les courbes montrent toutes une forme présentant des arches dont les minima sont localisés aux mêmes valeurs de  $q$ . Seule l'amplitude des courbes permet de les distinguer. Or, comme l'explique Williamson et al. dans un article datant de 2008, la décomposition la plus efficace nécessite que les formes des courbes de diffusion des différentes espèces en solution soient distinctes [110]. Du fait de la proximité des signaux des intermédiaires et de la capside, cette analyse ne nous permet pas de valider ou d'invalider la présence des intermédiaires proposés par Pierson et al. Il est donc raisonnable de conclure que nos résultats ne sont pas contradictoires.

De plus, à partir des observations précédentes, nous n'avons pas pu déterminer la présence d'autres intermédiaires probables de plus petites tailles. Nous pouvons alors conclure raisonnablement que s'ils existent, soit leurs concentrations, soit leurs temps d'existence ne sont pas suffisants pour contribuer significativement au signal, dans nos conditions expérimentales. Les équations d'assemblage de la capside peuvent alors être simplifiées de façon à ne dépendre que des dimères libres en solution et des capsides complètes.

#### 4.2.4 Reconstruction des cinétiques

A l'aide de l'équation 3.25 présentée dans la sous-section 3.2.2.2 et des observations précédentes, nous avons déterminé l'évolution de la fraction massique des espèces en solution au cours de la cinétique. La figure 5.24a est un exemple



**Figure 5.24:** La figure 5.24a présente l'évolution de la fraction en capsides (en bleu) et en dimères (en rouge) en fonction du temps pour la cinétique avec des concentrations en protéines et NaCl de 0,6 g.L<sup>-1</sup> et 1 M, respectivement. La figure 5.24b montre les fractions massiques finales en capsides en fonction de concentrations initiales en dimères et NaCl (cercles). Les croix représentent les expériences ayant été réalisées plusieurs fois.

des reconstructions réalisées par SVD. Sur celle-ci est représentée l'évolution, en fonction du temps, de la fraction massique des capsides (en bleu) ainsi que celle des dimères (en rouge). La figure 5.24b présente les fractions massiques finales en capsides obtenues en fonction de la concentration initiale en dimères et de la concentration en NaCl, pour l'ensemble des cinétiques.

Sur la figure 5.24b, nous pouvons observer que pour une faible concentration en NaCl, la fraction massique finale obtenue est nulle ou très faible (0,07 et 0,16)

avec une légère augmentation de fraction en fonction de la concentration initiale en dimères. Pour de faibles concentrations initiales en dimères, l'influence de la concentration en NaCl est plus visible. Alors qu'à 0,5 M de sel, il y a une fraction massique nulle en capsides dans nos conditions expérimentales, celle-ci augmente pour se situer entre 0,2 et 0,29 à 1 M et entre 0,64 et 0,69 à 1,5 M. Seules les cinétiques avec 0,6 g.L<sup>-1</sup> de protéines et 1 M de NaCl présentent des fractions massiques situées à 0,95 et 0,97, indiquant une réaction quasi-totale.

Ces observations montrent une dépendance de la fraction massique finale en capsides aux concentrations initiales en protéines et NaCl. Celles-ci peuvent s'expliquer assez simplement dans le cas où les équations d'assemblage de la capsides ne dépendent que de deux espèces, i.e. le dimère et la capsides, tel que  $\rho_T = \rho_1 + N\rho_N$ . En 2007, Van der Schoot et Zandi proposent un modèle cinétique permettant de décrire l'assemblage et le désassemblage des capsides de virus [111]. Celui-ci permet de relier  $f_c$  la fraction massique finale en capsides à  $\rho_T$  la concentration totale en dimères :

$$\begin{aligned} f_c &\approx \left(\frac{\rho_T}{\rho^*}\right)^N \ll 1 \quad \text{pour } \rho_T \ll \rho^* \\ &\approx 1 - \frac{\rho^*}{\rho_T} \quad \text{pour } \rho_T \gg \rho^* \end{aligned} \quad (4.1)$$

où  $f_c = N\rho_N / \rho_T$

Dans ces équations,  $\rho^*$  représente la concentration critique en dimères, i.e. la concentration minimale nécessaire à la formation de capsides. Etant donné que  $\rho^*$  est une concentration constante pour des conditions expérimentales données, lorsque  $\rho_T \gg \rho^*$ , l'équation 4.1 met en évidence l'augmentation de la capacité du système à former des capsides avec la concentration totale en dimères. La fraction massique finale en capsides est donc bien dépendante des concentrations initiales en protéines. De plus,  $\rho^*$  dépend de  $G_N^{cap}$ , l'énergie libre d'interaction au sein d'une capsides, tel que :

$$\rho^* v_0 \approx \exp(G_N^{cap} / Nk_B T) \quad (4.2)$$

Or, cette énergie dépend des potentiels chimiques de chaque constituant du mélange, autrement dit de la concentration en NaCl. Ainsi, lorsque la force ionique augmente, l'énergie libre d'interaction au sein d'une capsides diminue, entraînant une baisse de la concentration critique en dimères. Il y a donc un déplacement des

équilibres réactionnels avec la concentration en NaCl expliquant l'augmentation de la fraction massique finale en capsides avec la concentration en NaCl.

## 4.3 Construction d'un modèle d'assemblage de la capside vide

A partir des conclusions précédentes, je proposerai dans cette section un modèle cinétique d'assemblage de la capside vide du VHB. Pour ce faire, je présenterai les analyses réalisées sur chacune des cinétiques d'assemblage afin de déterminer leurs temps caractéristiques associés et discuterai des valeurs obtenues. Je présenterai également les différentes phases de l'assemblage déterminées par nos expériences.

### 4.3.1 Dépendance de la cinétique d'assemblage aux concentrations en protéines et NaCl

Les rayons de giration ainsi que les intensités quand  $q$  tend vers zéro ont été estimés, pour chaque courbe, à l'aide du logiciel DATRG<sup>4</sup> de la suite ATSAS [86]. Dans le cas des cinétiques avec une faible concentration en protéines et une forte concentration en NaCl, i.e. les cinétiques avec un faible rapport signal-sur-bruit, les estimations ont été confirmées manuellement sur plusieurs courbes, via un ajustement dans la région du Guinier (voir section 1.5), à l'aide du logiciel SaSView<sup>5</sup>.

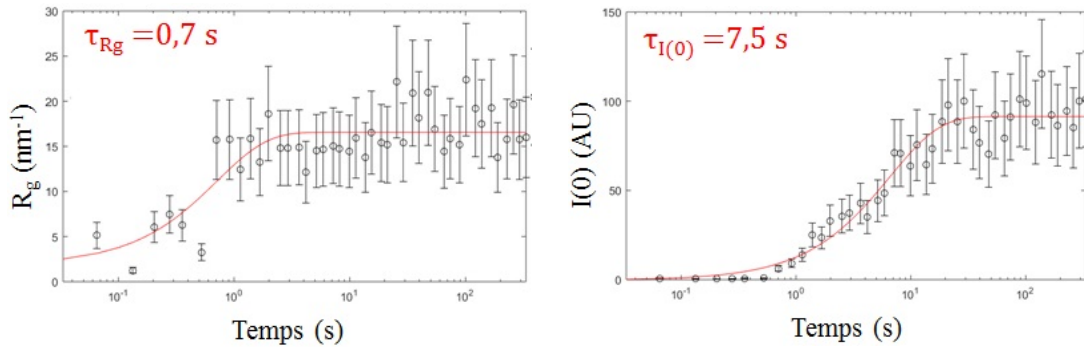
Afin d'estimer les temps caractéristiques d'évolution des rayons de giration et des intensités pour chacun des jeux de données, nous avons ensuite réalisé un ajustement à l'aide d'une exponentielle simple associée à une constante. De plus, les courbes d'intensités diffusées ont été converties en unités absolues. Elles ont ensuite été normalisées de telle façon que l'intensité  $I(q \rightarrow 0)$  nous permette de suivre l'évolution de la masse moyenne de l'ensemble des espèces en solution à chaque instant. Autrement dit, si la solution n'est composée que de dimères, l'intensité  $I(0)$  sera égale à 1 tandis que si la solution n'est composée que de

---

4. Outil permettant d'estimer, pour un grand nombre de courbes de diffusion, leurs  $R_g$  et  $I(0)$  ainsi que les erreurs associées

5. Logiciel d'analyse des données de diffusion aux petits angles. Site web : <https://www.sasview.org/>

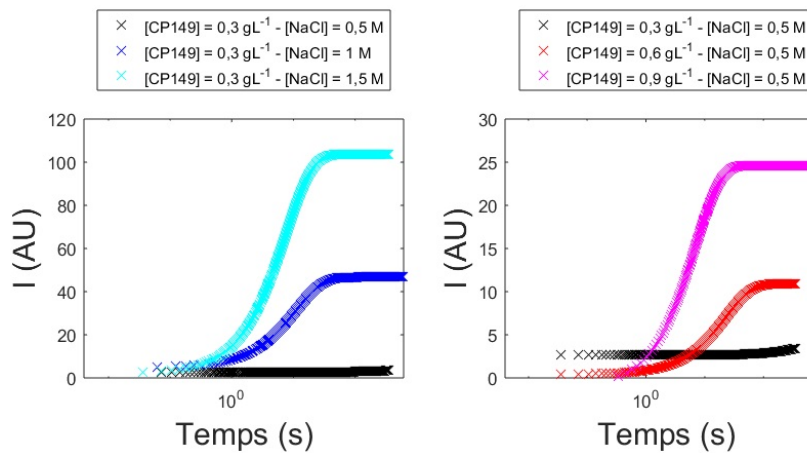
capsides, l'intensité  $I(0)$  sera égale à 120. Deux ajustements, associés à leurs données, sont montrés en exemple sur la figure 5.25.



(a) Évolution du rayon de giration en fonction du temps (b) Évolution de l'intensité  $I(0)$  en fonction du temps

**Figure 5.25:** Exemple de cinétique d'auto-assemblage de la capside du VHB. La cinétique présentée a une concentration en protéines et en NaCl de  $0,3 \text{ g.L}^{-1}$  et  $1,5 \text{ M}$ , respectivement. Les points noirs représentent les données expérimentales, les courbes rouges les ajustements.

Sur la figure 5.26, à gauche, nous montrons une comparaison des courbes ajustées de l'intensité des cinétiques réalisées à une concentration en protéines de  $0,3 \text{ g.L}^{-1}$



**Figure 5.26:** Comparaison de l'évolution, en fonction du temps, de l'intensité  $I(0)$  pour différentes concentrations en NaCl (à gauche) et en protéines (à droite).

et pour différentes concentrations en NaCl. Les courbes peuvent être divisées en plusieurs phases : une phase de faible croissance, une phase de croissance rapide et une phase asymptotique. A mesure que la concentration en NaCl augmente, la pente de la phase de croissance se raidit et la valeur du plateau augmente. Sur la même figure, à droite, nous montrons les courbes ajustées de l'intensité des cinétiques réalisées à une concentration en NaCl de  $0,5 \text{ M}$  et pour différentes



concentrations en protéines. Les courbes peuvent également être divisées en trois phases (faible croissance, croissance et asymptotique). Comme pour la figure précédente, à mesure que la concentration en protéines augmente, la pente de la phase de croissance se raidit et la valeur du plateau augmente. Nous pouvons également noter que la durée de la phase de faible croissance semble varier avec la concentration en protéines et en NaCl. Les valeurs des temps caractéristiques issus des ajustements et associées à ces figures sont résumées dans les tableaux suivants (Tab. 5.2).

TEMPS CARACTÉRISTIQUES D'ÉVOLUTION DES $I(0)$			
Concentration en protéines (g.L <sup>-1</sup> )	0,3	0,6	0,9
0,5	—	23 ± 2	7,1 ± 0,5
Concentration en NaCl (M)	1	7,1 ± 1,1	2,2 ± 0,2
		11,1 ± 0,6	1,6 ± 0,8
	1,5	7,2 ± 0,5	
		9,0 ± 0,6	
TEMPS CARACTÉRISTIQUES D'ÉVOLUTION DES $R_g$			
Concentration en protéines (g.L <sup>-1</sup> )	0,3	0,6	0,9
0,5	—	13,4 ± 1,1	2,2 ± 0,2
Concentration en NaCl (M)	1	5,0 ± 0,5	0,98 ± 0,04
		6,7 ± 0,7	0,72 ± 0,03
	1,5	0,61 ± 0,06	
		0,56 ± 0,03	

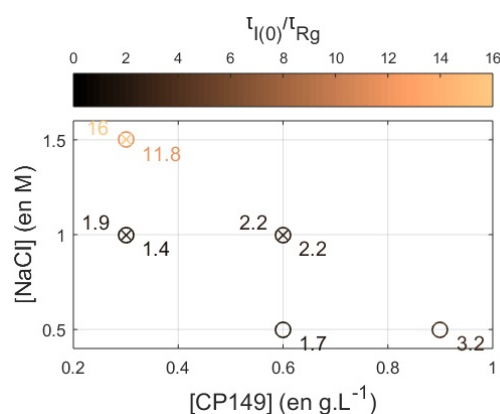
**Table 5.2:** Tableaux récapitulatifs des temps caractéristiques d'évolution des rayons de giration et des intensités pour chacun des jeux de données (en secondes).

A partir de ces tableaux, nous pouvons effectuer plusieurs remarques. La première est qu'il y a un bon accord des temps caractéristiques estimés pour les cinétiques avec des réplicats. De plus, nous pouvons noter que les temps caractéristiques d'évolution des rayons de giration diminuent aussi bien avec l'augmentation de la concentration en NaCl que celle en protéines. Concernant les temps caractéristiques d'évolution de l'intensité, ceux-ci évoluent de manière similaire, sauf pour la cinétique avec une concentration en protéines de 0,3 g.L<sup>-1</sup> et une concentration en NaCl de 0,5 M. Ceci, associé aux observations réalisées à partir de la figure 5.26, mettent en évidence la dépendance de la dynamique de l'assemblage à la concentration en protéines et en NaCl. En effet, les vitesses de réactions sont favorisées par l'augmentation de ces concentrations. Des conclusions identiques avaient été publiées dans des articles présentant des expériences réalisées par de la diffusion de lumière couplée avec de la chromatographie d'exclusion de taille avec

soit des solutions ayant un  $pH$  9,5 soit, des solutions à  $pH$  7,5. Dans les deux cas, les gammes de concentrations en NaCl étaient identiques aux nôtres [95] [102].

### 4.3.2 Chemin cinétique d'auto-assemblage des capsides vides

La dépendance de l'assemblage à la concentration en protéines et NaCl n'est pas la seule conclusion qui peut être extraite des courbes et des valeurs des temps caractéristiques d'évolution. En effet, en comparant les courbes d'évolution de l'intensité  $I(0)$  et du rayon de giration, nous pouvons constater qu'il existe deux groupes différents de cinétiques d'auto-assemblage de la capside. Plus précisément, le rapport des temps caractéristiques  $\tau_{I(0)}/\tau_{R_g}$  est situé entre 1,4 et 3,2 pour toutes les cinétiques en dehors de celle avec une concentration en protéines de  $0,3 \text{ g.L}^{-1}$  et une concentration en NaCl de  $1,5 \text{ M}$  qui est situé à 11,8 ou 16 en fonction du réplicat (Fig. 5.27). Nous présenterons donc séparément ces deux cas.



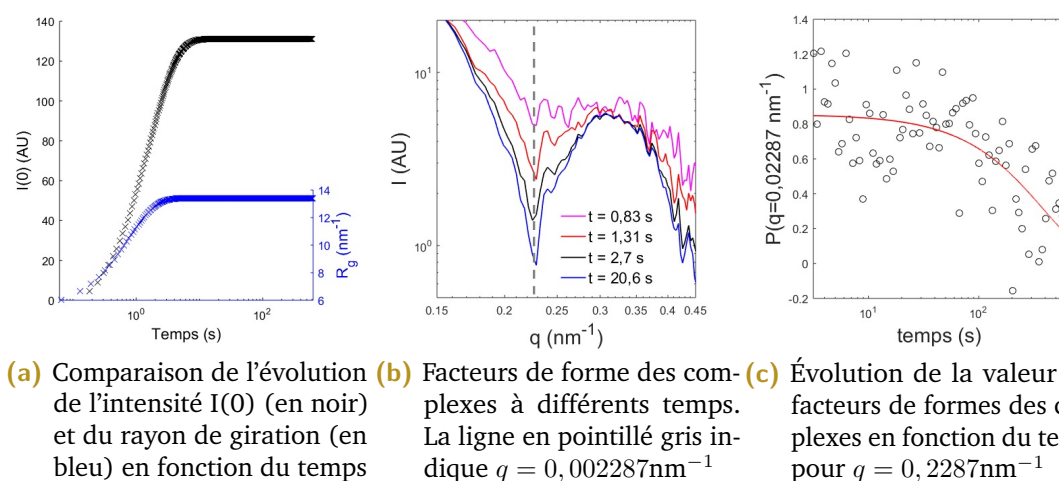
**Figure 5.27:** Rapport  $\tau_{I(0)}/\tau_{R_g}$  en fonction des concentrations initiales en dimères de protéines et en NaCl.

#### 4.3.2.1 Les différentes phases de l'assemblage

Dans le premier cas, nous pouvons constater qu'une cinétique peut être scindée en trois phases différentes. Pour plus de clarté, seule la cinétique avec une concentration en protéine de  $0,6 \text{ g.L}^{-1}$  et une concentration en NaCl de  $1 \text{ M}$  sera présentée et commentée. L'ensemble des cinétiques étudiées dans cette section peuvent être décrites de la même façon (aux constantes de temps près).

Comme nous pouvons le voir sur la figure 5.28a, il existe une première phase où la taille, ainsi que la masse des objets en solution, augmentent conjointement mais à des vitesses différentes. Les pentes des courbes d'évolution des rayons

de giration sont plus raides que celles des intensités  $I(0)$ . Cette phase est suivie d'une deuxième où la masse des objets continue d'augmenter alors que le rayon de giration n'évolue plus significativement. Associée à l'augmentation de l'intensité, nous pouvons noter une évolution du facteur de forme des objets en fonction du temps. Celle-ci est visible par une augmentation de l'amplitude des arches avec l'avancement de la cinétique (Fig. 5.28b). Cette augmentation perdure alors même que l'intensité  $I(0)$  n'évolue plus significativement. Nous pouvons ainsi définir la troisième phase de l'assemblage.



**Figure 5.28:** Dynamique d'assemblage et de structuration des dimères en capsides

Ces observations suggèrent que, dans un premier temps, plusieurs dimères s'agglomèrent rapidement pour former des complexes de la taille de la capside mais de masse inférieure à celle-ci. Dans un deuxième temps, ces complexes s'auto-organisent tout en capturant des dimères libres en solution. Une fois une masse correspondant à celle de la capside atteinte, les complexes continuent de s'auto-organiser lentement pour tendre vers la forme icosaédrique classique de la capside. En suivant l'évolution du premier minimum du facteur de forme en fonction du temps, nous pouvons estimer un temps caractéristique de relaxation de la capside. Un ajustement à l'aide d'une exponentielle simple permet d'estimer une valeur de  $\tau_P$ , pour la cinétique présentée, de  $390 \pm 70 \text{ s}$  ce qui est 2 ordres de grandeur plus grand que les deux autres temps caractéristiques. Pour les autres cinétiques, les valeurs sont situées entre  $320 \pm 65$  et  $650 \pm 70$  secondes.

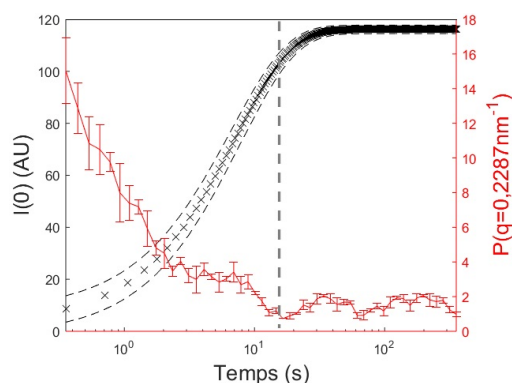
Dans la littérature, l'auto-assemblage des capsides vides de VHB a pu être décrit comme un mécanisme de nucléation-croissance (voir section 1.1.1). Nos obser-

vations ne nous permettent pas de parler d'un tel mécanisme. En effet, nous observons plutôt un mécanisme en trois phases : agglomération, croissance et relaxation. Dans ce mécanisme, le terme d'agglomération fait référence à la formation rapide de complexes formés de plusieurs dimères où les complexes ne présentent pas de facteurs de forme, i.e. de structures identifiables. Dans cette phase, la cinétique est initialisée par les interactions hydrophobes. Lorsque la concentration en NaCl est très faible, la répulsion électrostatique est forte. De ce fait, les forces hydrophobes ne sont pas suffisantes pour les compenser et conduire à la formation du complexe initial. Le terme de croissance est défini comme précédemment, i.e. ajout séquentiel de dimères dans une structure s'auto-organisant. Les temps caractéristiques de cette phase diminuent avec la répulsion électrostatique. La dynamique de cette phase est dirigée par une compétition entre la répulsion électrostatique et l'attraction hydrophobe. Le troisième terme, relaxation, fait référence à l'intervalle de temps pendant lequel les capsides vont se restructurer pour tendre vers la forme icosaédrique classique de la capside. Même si nous ne voyons pas d'évolution de l'intensité  $I(0)$  pendant cette phase, il ne peut être exclu que quelques dimères manquants soient ajoutés au cours de cette réorganisation. Les temps caractéristiques élevés de cette phase,  $\tau_P \gg \tau_{I(0)}$  et  $\tau_{R_g}$ , peuvent s'expliquer comme ceci : à mesure que l'on s'approche du nombre total de dimères pour former une capside complète, l'énergie nécessaire pour insérer un dimère croît du fait de l'encombrement stérique, ce qui ralentit le processus d'assemblage.

#### 4.3.2.2 Cas particulier

Contrairement aux autres cinétiques, celle avec une concentration de  $0,3 \text{ g.L}^{-1}$  et  $1,5 \text{ M}$  en protéines et en NaCl respectivement, ne présente pas trois mais deux phases différentes identifiables. Le temps caractéristique d'évolution pour  $I(0)$  est similaire à celui de la cinétique à  $1 \text{ M}$  tandis que celui pour  $R_g$  est environ 10 fois plus petit que celui de cette dernière. Ainsi, lors de la première phase de cette cinétique, l'agglomération des dimères pour former les complexes est beaucoup plus rapide. Ceci peut s'expliquer par l'augmentation de la force ionique qui diminue fortement la répulsion électrostatique entre les dimères favorisant, ainsi, la formation des complexes. Comme présenté précédemment, lors d'une seconde phase, ces complexes s'auto-organisent tout en capturant des dimères

libres en solution. Dans ce cas, le temps caractéristique de la réaction est peu modifié par la forte concentration en NaCl. Ceci peut s'expliquer par le fait que l'auto-organisation, ainsi que la capture séquentielle des dimères, est dirigée par l'attraction hydrophobe qui n'est pas modifiée par la forte concentration saline.



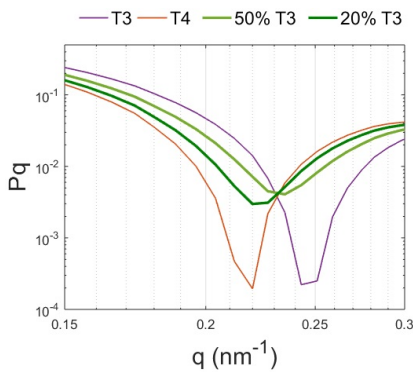
**Figure 5.29:** Comparaison de l'évolution de l'intensité  $I(0)$  (en noir) et de la valeur du premier minimum des facteurs de formes (en rouge).

La figure 5.29 met en évidence l'absence de troisième phase, i.e. pas de restructuration des capsides. En effet, sur cette figure est présentée l'évolution de l'intensité  $I(0)$  et de la valeur du premier minimum des facteurs de formes. Nous pouvons constater que contrairement aux cinétiques précédentes, il n'y a pas d'évolution significative des minima. Ceci peut s'expliquer de deux façons, la première est que l'intervalle de temps nécessaire à la restructuration, pour cette cinétique, est très supérieur à celui de nos expériences (9 min). La seconde explication est que, du fait de la forte concentration saline, l'attraction hydrophobe devient très grande devant la répulsion électrostatique. De ce fait, les dimères ne peuvent plus se réarranger. Les capsides formées dans ces conditions sont piégées cinétiquement et contiennent des défauts. Dans un article en préparation, Li et al. observent un comportement similaire de certaines capsides lors de simulations informatiques[112]. Dans cet article, les auteurs utilisent la méthode de Monte-Carlo pour étudier le rôle de l'énergie élastique dans la dynamique d'assemblage des capsides de virus. Pour un régime particulier, où les interactions hydrophobes sont très grandes devant les énergies élastiques, les auteurs ont mis en évidence un régime lors duquel les capsides sont désordonnées et ne peuvent pas relaxer vers la forme icosaédrique. Les capsides sont également piégées cinétiquement.

### 4.3.2.3 Morphologie des capsides

Comme nous avons pu le voir dans la section 1.1.3, une étude récente de Sun et al met en évidence la possibilité de déplacer la proportion entre les formes de capsides  $T = 3$  et  $T = 4$ , à l'aide de cations tels que  $\text{Na}^+$  ou  $\text{Ca}^{2+}$  [101]. Dans cet article, quel que soit le  $\text{pH}$  des solutions, les auteurs observaient une diminution du pourcentage de capsides  $T = 4$  avec l'augmentation en sel. Par exemple, lorsque la concentration en  $\text{NaCl}$  de la solution augmente de 0,5 M à 1 M, le pourcentage de capsides  $T = 4$  diminue de 80% à 50%. Nous n'observons pas le même phénomène.

Sur la figure 5.30, nous pouvons voir un zoom sur le premier minimum des intensités diffusées théoriques des capsides  $T = 3$  et  $T = 4$  ainsi que des mélanges



**Figure 5.30:** Intensité diffusées théoriques des capsides  $T = 3$  et  $T = 4$  ainsi que des mélanges de  $T = 3$  et  $T = 4$ .

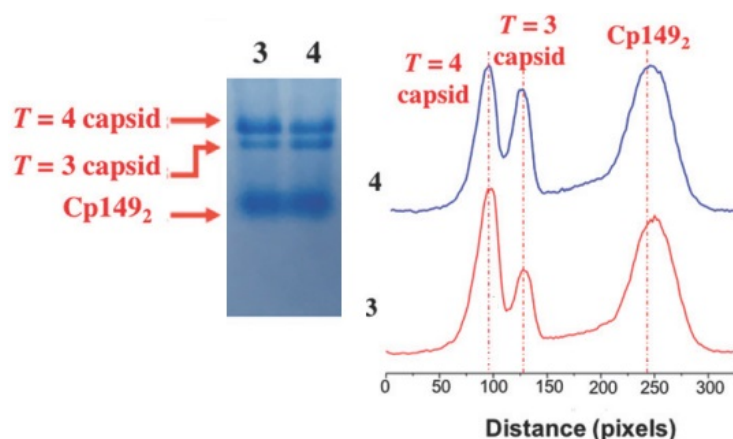
		Concentration en protéines (g.L <sup>-1</sup> )		
		0,3	0,6	0,9
Concentration en NaCl (M)	0,5	—	0,225	0,229
	1	0,228	0,225	—
		0,232	0,229	—
1,5	0,225	—	—	
		0,229	—	—

**Table 5.3:** Tableaux récapitulatifs du premier minimum des facteurs de forme des capsides issues des cinétiques d'assemblage (en  $\text{nm}^{-1}$ ).

de  $T = 3$  et  $T = 4$ . Comme il est visible sur cette figure, les premiers minima sont distincts. Pour la capside  $T = 4$  (courbe orange), celui-ci se situe à  $q = 0,22 \text{ nm}^{-1}$  tandis que pour la capside  $T = 3$  (courbe violette), il est légèrement supérieur à  $q = 0,24 \text{ nm}^{-1}$ . En théorie, par décomposition en valeurs singulières, nous devrions être en mesure de séparer ces deux signaux. En pratique, nous n'en avons pas été capables. Cela peut être dû au fait que la concentration en capsides  $T = 3$  de nos solutions est trop faible pour contribuer de manière significative au signal ou parce que notre rapport signal-sur-bruit est trop faible. Si notre incapacité à discerner les facteurs de forme des deux capsides provient de la faible contribution des capsides  $T = 3$ , et donc du fait que le signal de celle-ci est mélangé à celui de la capside

$T = 4$ , nous pouvons tout de même nous intéresser à l'évolution du facteur de forme de cette dernière. En effet, si la composition du mélange évolue avec la salinité, alors l'intensité diffusée devrait également être modifiée. Comme nous pouvons le voir sur la figure 5.30, le premier minimum se déplace vers les plus grandes valeurs de  $q$  avec l'augmentation de la quantité de  $T = 3$  en solution. Pour une solution contenant 20% de capsid  $T = 3$ , le minimum est situé autour de  $q = 0,22 \text{ nm}^{-1}$  tandis que pour une solution contenant 50% de capsid  $T = 3$ , le minimum atteint  $0,23 \text{ nm}^{-1}$ . Or, comme nous pouvons le voir dans le tableau 5.3, nous n'avons pas de déplacement significatif des valeurs des minima.

Ci-dessous, nous pouvons voir une figure présentant un des gels d'électrophorèse réalisés par Sun et al ainsi que l'analyse par densité optique des bandes du gel. Sur le gel, sont montrés les assemblages déclenchés avec 1 M de  $\text{K}^+$  dans une solution de dimères à  $1 \text{ g.L}^{-1}$  via addition successive de deux aliquotes de sel à  $\text{pH } 7,5$  et via un unique ajout à  $\text{pH } 8,5$ . Les expériences sont notées 3 et 4, respectivement.



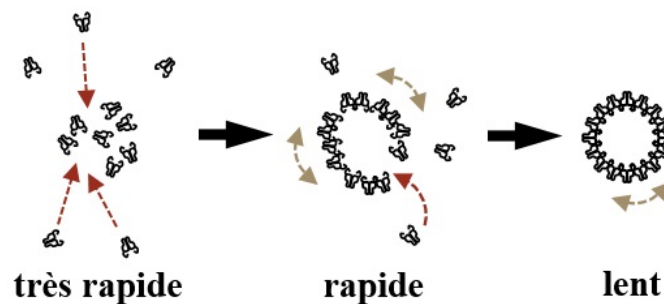
**Figure 5.31:** Figure illustrant le dimorphisme observé par Sun et al. A gauche, le gel d'électrophorèse des capsides issues des expériences d'assemblage. A droite, l'analyse par densité optique des bandes du gel. (Adapté de [101]).

Sur cette figure, nous pouvons remarquer que les proportions de capsides  $T = 3$  analysées ont tendance à être inférieures à celles obtenues par d'autres méthodes de caractérisation *in situ*. Les auteurs expliquent cette différence par le fait que les capsides  $T = 3$  ont sans doute été dissociées en dimères sous le champ électrique en raison de leur plus faible stabilité. Cependant, sur les images de microscopie électronique présentées dans l'article, des capsides incomplètes sont visibles et les auteurs n'abordent pas cette question. Les capsides incomplètes ayant également

une stabilité plus faible que les capsides  $T = 4$ , celles-ci peuvent aussi être à l'origine de ces dissociations. De plus, leurs masses, leurs densités et leurs charges globales sont également inférieures à celles de la capside  $T = 4$ , ce qui modifie leur migration, en direction de la bande  $T = 3$ , aussi bien sur un gel d'électrophorèse que lors de la centrifugation sur gradient de sucre. De ce fait, il est possible que les proportions de capsides  $T = 3$  aient été surestimées.

## 4.4 Conclusion

Dans les conditions expérimentales de l'étude, nous avons pu identifier des cinétiques d'assemblage, de la capside vide du VHB, avec trois phases : agglomération, croissance et relaxation. Ces phases sont dirigées par l'attraction hydrophobe qui



**Figure 5.32:** Illustration schématique de la dynamique d'assemblage des protéines CP149.

est modulée par la répulsion électrostatique. En effet, les vitesses de réaction sont favorisées par l'augmentation de la concentration en NaCl. Cependant, lorsque l'écrantage ionique est trop important, la phase de relaxation disparaît et les capsides formées sont piégées cinétiquement et comportent très certainement des défauts. Nous avons également pu observer l'influence de la concentration en protéines sur la dynamique d'assemblage. Les vitesses de réaction sont favorisées par l'augmentation de celle-ci. Au cours des cinétiques, nous n'avons pas identifié d'intermédiaire d'assemblage ou pu confirmer l'influence du sel sur le dimorphisme des capsides. Ceci peut être imputable aux très faibles concentrations en protéines utilisées lors de nos expériences qui, du fait du rapport signal-sur-bruit, rendent difficile l'analyse.



# Conclusion du chapitre

Dans ce chapitre, nous avons étudié les interactions au sein des capsides vides du VHB par CTPF ainsi que le désassemblage et l'auto-assemblage de ces mêmes capsides par la technique de TR-SAXS couplée à une analyse par SVD.

Dans le premier cas, nous avons pu estimer une énergie d'interaction des dimères, entre eux, de  $-6 k_B T$ , ce qui est similaire à ce qui avait été rapporté précédemment. La même technique nous a également permis d'estimer la charge effective du dimère de protéine, i.e.  $1,6e$  en valeur absolue à un  $pH$  7,5, qui n'avait pas été identifiée auparavant. Dans le deuxième cas, nous avons pu observer une cinétique de désassemblage de plusieurs heures composée de deux phases et identifier trois espèces en solution, dont une espèce transitoire de désassemblage [26]. La présence de cet agrégat avait déjà été observée auparavant mais sa nature n'avait pas été déterminée. Dans nos conditions expérimentales, nous avons identifié celui-ci comme étant un agrégat de dimères pouvant être décrit comme une structure fractale-branchée. Dans le troisième cas, nous avons pu identifier une cinétique d'assemblage courte (quelques minutes à dizaine de minutes) en trois phases, i.e. agglomération, croissance et relaxation, dirigée par l'attraction hydrophobe et modulée par la répulsion électrostatique. Cette description, proche du mécanisme de nucléation-croissance utilisé dans la littérature, ajoute à celui-ci l'idée d'une réorganisation possible au sein de la capside à l'issue de l'assemblage. La charge effective du dimère étant faible, la phase de relaxation peut être facilement inhibée, par une augmentation de la salinité, piégeant ainsi cinétiquement les capsides dans une conformation contenant des défauts.

L'ensemble de ces éléments montre des dynamiques d'assemblage et de désassemblage de la capside vide du VHB différentes, tant du point de vue de la temporalité que des chemins cinétiques empruntés. Ces observations mettent donc en évidence une hystérésis. D'un point de vue biologique, et donc en dehors de conditions extrêmes (par exemple, 1,5 M de sel), celle-ci peut être décrite comme ceci : l'assemblage associé à la capacité de restructuration sont favorisés par les interactions ce qui conduit de manière reproductible à la formation de capsides icosaédriques.

De manière opposée, le désassemblage est défavorable et permet de préserver la capsidie d'environnements néfastes à la survie du virus.



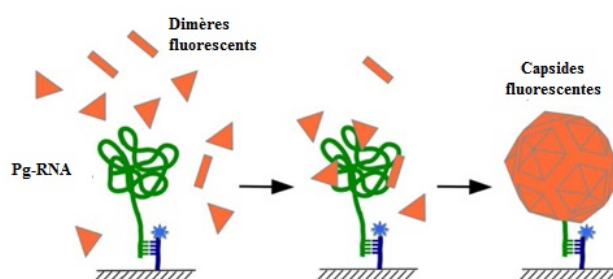
# VI

---

Stratégie de marquage pour l'étude en  
molécule unique de l'auto-assemblage  
de la capsid pleine du VHB

Comme nous avons pu le voir dans le chapitre 3, comprendre l'auto-assemblage de la capsid du VHB est une stratégie prometteuse pour envisager de nouveaux traitements. L'étude, présentée précédemment et réalisée par TR-SAXS, a permis d'obtenir des informations temporelles et structurales. Ces informations nous ont permis de parvenir à une description d'un chemin dynamique d'assemblage possible. Cependant, cette étude ne s'intéressait qu'à la protéine CP149, i.e. composée uniquement du domaine d'assemblage et ne tenait donc pas compte des interactions entre le génome et les protéines. De plus, la technique de TR-SAXS ne permet de suivre qu'une variation moyennée des espèces en solution en fonction du temps.

En parallèle, nous avons donc cherché à étudier les mêmes processus, mais, au niveau de la capsid seule, par microscopie de fluorescence par réflexion interne totale (TIRFM). Cette technique ayant une grande sensibilité, elle permet de suivre le nombre de dimères se liant à des molécules d'ARN uniques (Fig. 6.1). De ce fait, elle permet d'estimer le rapport de charge entre le génome et les dimères en fonction du temps. De plus, les résolutions attendues, spatiale et temporelle, de cette technique sont de quelques dimères au maximum et quelques centaines de millisecondes, respectivement. L'ensemble en fait une technique parfaitement complémentaire du TR-SAXS.



**Figure 6.1:** Illustration schématique du principe des expériences TIRFM. L'ARN est immobilisé sur une surface et des dimères fluorescents s'y lient progressivement afin de former une capsid fermée. L'intensité de la fluorescence, dépendante du nombre de dimères marqués à la surface de l'ARN, est enregistrée à mesure que le temps varie.

Cette partie du manuscrit sera divisée en trois chapitres. Dans un premier temps, je présenterai l'ensemble des outils nécessaires à sa compréhension : la microscopie de fluorescence par réflexion totale interne, les différentes protéines mutantes utilisées, la chromatographie d'exclusion de taille ainsi que les protocoles de

production. Je présenterai ensuite les différentes stratégies mises en place pour obtenir des dimères de protéines fluorescents, en détaillant les résultats obtenus pour chacune des protéines mutantes. Le dernier chapitre sera consacré aux résultats obtenus, en TIRFM, pour le génome seul.

# Matériels et méthodes

Dans cette partie, nous verrons les différents outils nécessaires à l'étude *in vitro* de l'auto-assemblage des capsides par TIRFM. Pour ce faire, je commencerai par présenter le principe de la microscopie de fluorescence par réflexion totale interne. Celle-ci nécessitant des systèmes d'intérêt fluorescents, je présenterai ensuite les différents mutants de la protéine native pouvant répondre à cette problématique. La section suivante sera consacrée à la chromatographie d'exclusion de taille. Cette technique, usuellement employée par les biologistes afin de purifier les protéines, a également été utilisée lors de ma thèse pour déterminer la meilleure stratégie à adopter afin d'obtenir des dimères de protéines fluorescents. Je terminerai par la présentation des solutions tampons ainsi que des différents protocoles expliquant comment produire, purifier et désassembler les dimères de protéines.

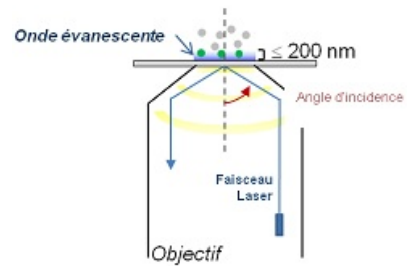
## 1.1 TIRFM

La microscopie de fluorescence par réflexion totale interne (TIRFM, pour Total Internal Reflection Fluorescence Microscopy, en anglais), est un type particulier de microscopie optique à fluorescence permettant d'observer un volume mince d'échantillon (moins de 200 nm) au voisinage d'une surface. Ce type de microscopie convient particulièrement bien à l'observation de phénomènes biologiques tels que l'auto-assemblage de particules virales.

### 1.1.1 Principe du TIRFM

Le principe de l'appareil consiste à exciter la fluorescence sur une très faible profondeur, i.e. l'environnement échantillon immédiatement adjacent au support. Ce volume d'excitation est suffisant pour observer des systèmes de tailles nanométriques et, également, suffisamment sélectif pour que le bruit de fluorescence soit limité. L'excitation est due à une onde évanescente. Celle-ci est générée lorsque la lumière incidente, passant d'un milieu à indice de réfraction élevé (par ex. du verre) à un milieu à indice de réfraction plus faible (par ex. de l'eau), est totalement réfléchi à

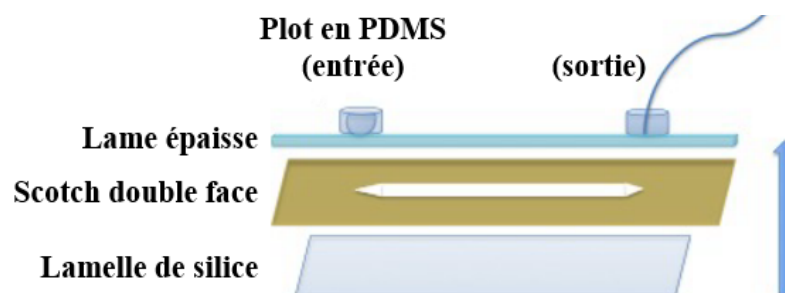
l'interface entre ces milieux. Cette réflexion totale a lieu pour un angle d'incidence donné et génère un champ électromagnétique sur une faible profondeur au-delà de la surface. Celui-ci décroît exponentiellement en fonction de la distance par rapport à la surface et donne l'énergie nécessaire à l'excitation des fluorophores. Il est alors possible d'observer la fluorescence des molécules individuellement dans la couche mince (Fig. 6.2<sup>1</sup>).



**Figure 6.2:** Principe de la microscopie de fluorescence par réflexion totale interne (TIRFM).

### 1.1.2 Préparation de la cellule micro-fluidique

Notre objectif étant d'étudier des processus dynamiques, nous devons être en mesure d'injecter, dans les cellules d'observations, le milieu qui contiendra les protéines. Et, cette injection doit être quasi-instantanée, pour ne pas introduire une incertitude sur la définition du temps initial,  $t_0$ . La cellule micro-fluidique que nous avons utilisée, et qui répond aux critères, a été mise au point par David Dulin<sup>2</sup>. Celle-ci est partiellement réutilisable et peut donc être démontée et remontée pour chaque expérience (Fig. 6.3).



**Figure 6.3:** Illustration schématique d'une cellule micro-fluidique (Adaptée de [113]).

Avant tout traitement les lamelles doivent être nettoyées, séchées puis soumises à un plasma d'oxygène (de 0,6 mbar) pendant 10 min. Le plasma d'oxygène détruit les impuretés résiduelles du premier lavage et active les fonctions hydroxyles

1. Adaptée du site : <http://bfa.univ-paris-diderot.fr/microscopie-a-fluorescence-excitee-par-onde-evanescente/> (23 septembre 2019)

2. Thèse soutenue en 2009 à l'Institut d'Optique Graduate School (IOGS)



de la silice ( $\text{SiO}_2$ ) en surface. La partie réutilisable est constituée d'une lame de microscope en verre percée en deux points. L'un des trous est surplombé par un réservoir en PDMS de volume interne  $100 \mu\text{L}$ . Sur le second est greffé un capillaire en PTFE terminé par un embout en plastique reliable à une seringue. Sur cette base, une couche de scotch double face est collée pour étanchéifier l'espace entre la lame et la lamelle de microscope. Un canal, entre le réservoir et le point d'entrée du tuyau en plastique, est découpé dans le scotch. Celui-ci a un volume d'environ  $18 \mu\text{L}$ <sup>3</sup>. La cellule micro-fluidique est ensuite fermée en apposant la lamelle sur le scotch. Du liquide peut alors être déposé dans le réservoir et aspiré dans le canal à l'aide d'une seringue adaptée. La zone d'observation est d'environ de  $100 \mu\text{m} \times 100 \mu\text{m}$ , ce qui permet de réaliser plusieurs acquisitions sur différentes zones dans un même canal. Après l'expérience, le scotch et la lamelle sont enlevés et le reste du système nettoyé avant une prochaine utilisation.

### 1.1.3 Accroche des systèmes biologiques

Pour pouvoir observer nos échantillons en TIRFM, ceux-ci doivent être confinés dans le volume d'excitation, i.e. les 100 premiers nanomètres à la surface de la lamelle. Il faut donc "accrocher" nos systèmes d'intérêts, en surface de la lamelle, pour éviter qu'ils ne diffusent librement hors de l'onde évanescente. Pour ce faire, un protocole de chimie de surface a été développé et optimisé par l'équipe molécule unique de l'IOGS (Fig. 6.4). Cette chimie de surface fonctionnelle, essentielle à la réalisation de nos expériences, doit répondre à deux contraintes :

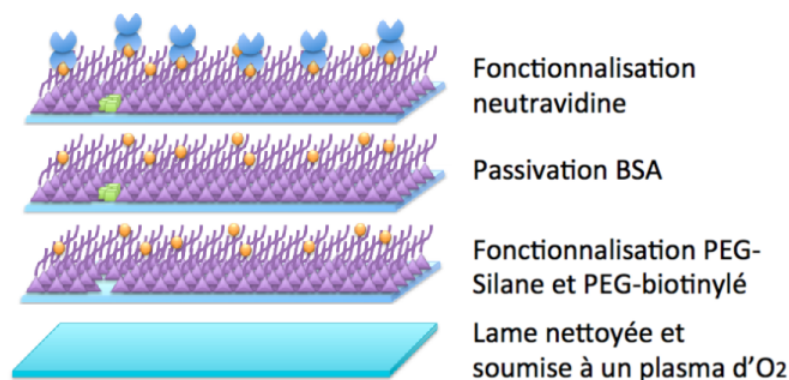
- passiver la surface pour empêcher l'accroche d'impuretés fluorescentes qui dégraderaient le rapport signal/bruit.
- assurer une accroche fiable et spécifique de nos systèmes biologiques.

La passivation est assurée par du PEG-Silane<sup>4</sup>. La fonction silane de celui-ci réagit avec les fonctions hydroxyles activées de la silice de la lamelle pour former une liaison covalente. Le PEG va ainsi former une couche homogène et électriquement neutre sur la surface, ce qui empêchera d'éventuelles molécules de venir s'y accrocher de manière non-spécifique.

3. largeur  $\times$  Longueur  $\times$  hauteur =  $4 \text{ mm} \times 45 \text{ mm} \times 0,1 \text{ mm}$

4. Polyéthylène-glycol triéthoxysilane (5 kDa)

Dans le même temps, des sites permettant l'accroche spécifique de nos systèmes d'intérêts sont créés par le couple biotine/neutravidine. Les sites possédant des biotines sont introduit par un mélange de PEG-silane partiellement biotinylé dans un tampon d'acide borique<sup>5</sup>. Les proportions de ce mélange et la constitution du tampon d'acide borique ont été optimisées, lors de la thèse de Nathalie Barbier<sup>6</sup>, pour assurer une passivation efficace et un nombre de sites d'accroche suffisant en surface. Après incubation du mélange, une solution de BSA<sup>7</sup> en tampon Tris est injectée. Celle-ci permet de combler les éventuelles aspérités dans la couche de PEG ce qui termine de passiver parfaitement l'ensemble de la surface. Une solution de neutravidine est ensuite injectée dans la cellule micro-fluidique. Celle-ci, ayant une grande affinité avec la biotine, permet la formation d'un complexe très stable et l'interaction avec notre système d'intérêt, i.e. l'ARN prégénomique du VHB. La production de ce dernier sera développée dans la section 3.1.1.



**Figure 6.4:** Illustration schématique des différentes étapes de chimie de surface (du bas vers le haut) (Adaptée de [113]).

## 1.2 Les protéines mutantes de la CP183

### 1.2.1 Les différentes mutations

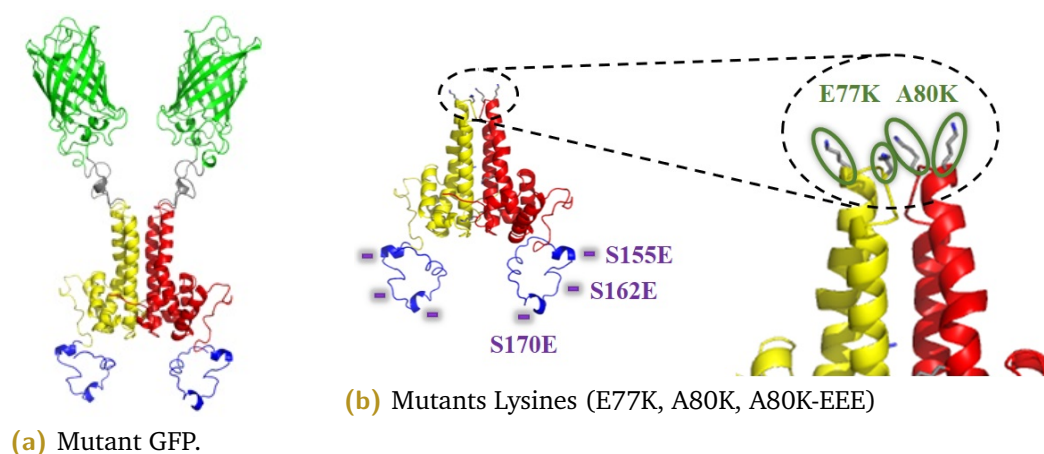
Afin de sonder l'auto-assemblage des protéines de la capsid du VHB autour de son ARN-pré génomique, nous avons eu recours à différents mutants de la protéine native :

5. 100 mM, pH=8,8

6. Thèse soutenue en 2017 à l'Institut d'Optique Graduate School (IOGS)

7. Bovin Serum Albumin

- Un mutant généré par l'insertion d'une protéine intrinsèquement fluorescente au niveau de la boucle du spicule de la protéine native. Sur la figure 6.5a, nous pouvons observer que la localisation de cette insertion autorise l'assemblage des protéines sous forme de capsides[114][115]. Cette protéine fluorescente est appelée eGFP, pour enhanced Green Fluorescent Protein en anglais, et a la propriété d'émettre sa fluorescence dans le vert. La protéine mutante est notée GFP-CP183 dans le manuscrit.
- Des mutants ponctuels générés par substitution d'un acide aminé par un autre, également au niveau de la boucle du spicule de la protéine. Ces mutants, notés A80K-CP183 et E77K-CP183 dans le manuscrit, sont présentés sur la figure 6.5b, à droite. Les mutations du glutamate 77 et de l'alanine 80 en lysine ont été réalisées afin de pouvoir marquer chimiquement la protéine avec un fluorophore.
- Le même mutant ponctuel comportant également des modifications dans le domaine d'interaction avec le génome. Ces modifications permettent d'imiter les propriétés physico-chimiques de la protéine au moment de l'empaquetage. En effet, la mutation de trois sérines en glutamates ( $S \rightarrow E$ ), visible à gauche sur la figure 6.5b, permet d'introduire des charges négatives sur le bras flexible. Ce sont ces charges qui permettent de mimer l'état phosphorylé de ces acides aminés. Ce mutant, comportant trois glutamates (E), est noté A80K-EEE-CP183 dans le manuscrit (Fig. 6.5b, à gauche).



**Figure 6.5:** Structures 3D des dimères de protéines mutées à partir de la protéine native de la capsid du VHB. 6.5a Structure du mutant généré par l'insertion de la eGFP visible en vert. 6.5b Structure des mutants ponctuels générés par substitution d'un ou plusieurs acides aminés.

## 1.2.2 Avantages et inconvénients

**Table 6.1:** Tableau récapitulatif des avantages et inconvénients de chacune des protéines mutantes.

MUTANT	AVANTAGES	INCONVÉNIENTS
GFP-CP183	Nombre de fluorophores par dimère connu, Protéine bien caractérisée par les biologistes.	Taille du fluorophore et de la protéine d'intérêt comparable pouvant modifier la dynamique d'assemblage.
E77K-CP183	Fluorophore petit devant la protéine, Protéine peu modifiée.	Nombre de fluorophores par dimère pouvant varier (0 à 2), Modification de la charge négative portée par l'acide aminé en charge positive.
A80K-CP183	Fluorophore petit devant la protéine, Protéine peu modifiée.	Nombre de fluorophores par dimère pouvant varier (0 à 2). Ajout d'une charge positive.
A80K-EEE-CP183	Avantages identiques au mutant A80K-CP183, Mutant phospho-mimétique.	Nombre de fluorophores par dimère pouvant varier (0 à 2).

## 1.3 La chromatographie d'exclusion de taille

La chromatographie est une méthode physico-chimique permettant de séparer les constituants d'un mélange sur une phase stationnaire immobile, à l'aide d'une phase mobile, liquide ou gazeuse, de nature différente. Chaque molécule est retenue selon ses propriétés physico-chimiques ce qui permet la séparation des différents constituants. Il existe plusieurs types de chromatographie avec différents facteurs physico-chimiques intervenant comme critères de séparation. Ceux-ci peuvent être la taille, la charge, l'hydrophobicité, etc.

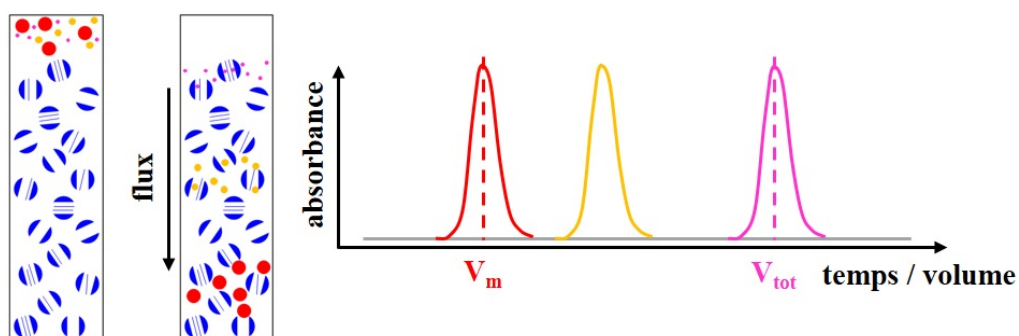
### 1.3.1 Principe de la SEC

La chromatographie d'exclusion de taille, SEC pour size exclusion chromatography en anglais, est une méthode de chromatographie en phase liquide permettant de séparer des molécules en fonction de leur rayon hydrodynamique lors de leur passage à travers un gel poreux. Cette technique présente deux phases :

- une phase stationnaire constituée de billes creuses qui laisseront plus ou moins diffuser les molécules à analyser.

- une phase mobile qui permet le déplacement de ces molécules dans la phase stationnaire.

Les billes contiennent un maillage créant un réseau avec des pores plus ou moins étroits. La matrice formant le réseau est supposée inerte vis à vis des protéines, i.e. il n'y a pas d'affinité chimique entre celle-ci et les protéines. De ce fait, les molécules d'intérêt vont être plus ou moins ralenties selon leur taille, uniquement. Les molécules ayant un diamètre supérieur à celui des pores ne pourront pénétrer la matrice et seront donc éluées en premier, au niveau du volume mort (molécules en rouge sur la figure 6.6). Ce dernier, noté  $V_m$ , correspond au volume autour des billes. Les molécules de tailles intermédiaires peuvent pénétrer partiellement la matrice et traversent plus de volume. Leur migration étant freinée, celles-ci seront donc éluées plus tardivement (molécules en jaune). Le volume nécessaire pour faire sortir une molécule de taille intermédiaire de la colonne est appelé volume d'éluion et est noté  $V_e$ . Une molécule totalement incluse dans la matrice sera éluée avec un volume d'éluion tel que  $V_e = V_{total} = V_m + V_i$ , où  $V_i$  est le volume de tampon interne au maillage. C'est le cas des solutés (molécules en magenta).



**Figure 6.6:** Illustration schématique du principe de la chromatographie d'exclusion de taille.

La SEC permet donc de séparer les molécules en fonction de leur taille mais également de déterminer leur masse moléculaire. En effet, dans l'hypothèse où toutes les protéines ont la même densité, la taille apparente des protéines correspond à celle de protéines globulaires de même masse moléculaire, ce qui donne une relation de proportionnalité entre le volume d'éluion et le logarithme de la masse moléculaire. Attention, cette relation n'est valable que sur la gamme de séparation de la matrice. En effet, au-delà d'une certaine taille, toutes les grosses molécules sortent au volume mort. De la même manière, en-deçà d'une certaine taille, toutes les petites molécules sortent au volume total.

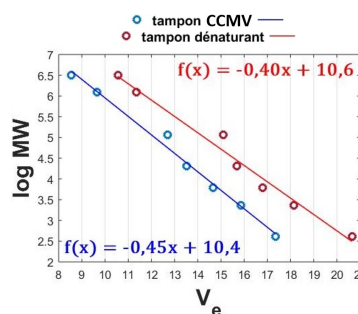
### 1.3.2 Calibration

Il est nécessaire de déterminer la relation de proportionnalité entre le volume d'élution et le logarithme de la masse moléculaire. En effet, celle-ci dépend du type de colonne utilisée ainsi que de la phase mobile, i.e. la solution tampon utilisée. Dans notre cas, nous avons utilisé des colonnes superdex S200 dont la matrice est formée d'un composite de polymères dextran et de chaînes d'agarose avec une gamme de séparation située entre 10 et 600 kDa. Sur la figure 6.7, sont présentées deux calibrations mettant en évidence l'impact de la solution tampon sur le volume d'élution d'une protéine. Les étalons utilisés pour ces calibrations sont présentés dans le tableau 6.2. Sur la figure, les points et les droites représentent, respectivement, les données expérimentales et les ajustements de ces dernières. Dans le cas du tampon CCMV<sup>8</sup> (données en bleues), le volume d'élution attendu

LISTE DES ÉTALONS

Thyroglobuline	669 kDa
Ferritin	440 kDa
Aldolase	158 kDa
Conalbumine	75 kDa
Ovalbumine	44 kDa
Anhydrase carbonique	29 kDa
Ribonucléase	13,7 kDa

**Table 6.2:** Tableau récapitulatif des étalons utilisés pour les calibrations.



**Figure 6.7:** Courbes de calibration de la colonne S200 10/300 GL.

pour une protéine de 42 kDa est attendu autour de 14,8 mL. Dans le cas du tampon dénaturant<sup>9</sup> (données en rouges), le volume d'élution attendu, pour la même protéine, est déplacé vers 17,2 mL.

8. Composition du tampon CCMV : 50 mM Tris HCl pH 7,5 ; 0,5 M NaCl ; 1 mM EDTA ; 1 mM DTT ; 0,5 mM PMSF

9. Composition du tampon dénaturant utilisé : 20 mM Tris pH 7,5 ; 2 M chlorure de guanidium ; 0,5 M LiCl ; 10 mM DTT

## 1.4 Protocoles détaillés de production de la protéine native

Cette section présente les protocoles utilisés dans notre laboratoire afin de transformer les bactéries qui serviront à exprimer les protéines avant d'être purifiées. Ceux-ci nous ont été fournis par Lauriane Lecoq du MMSB<sup>10</sup>, après adaptation des protocoles de Michael Nassal.

### 1.4.1 Les solutions tampons usuelles

**Table 6.3:** Tableau récapitulatif des solutions tampons utilisées pour la production des dimères de protéines.

NOM	COMPOSITION
Tampon de lyse (TN50)	50 mM Tris pH 7,5; 50 mM NaCl; 5 mM EDTA; 5 mM DTT
Tampon de purification	50 mM Tris pH 7,5; 5% sucrose; 2,5 mM EDTA; 1 mM DTT
Tampon de stockage	50 mM Tris pH 7,5; 5% sucrose; 5 mM EDTA; 2 mM DTT
Tampon de désassemblage	20 mM Tris pH 7,5; 2 M chlorure de guanidium; 0,5 M LiCl; 10 mM DTT

**Table 6.4:** Tableau récapitulatif des fonctions des produits utilisés pour les solutions tampons.

PRODUIT	FONCTION
Tris	maintient le pH (gamme d'action : 6,5 à 9,7)
chlorure de sodium	sert à ajuster la force ionique, soit pour créer un choc osmotique afin de déstabiliser la membrane cellulaire, soit pour moduler les interactions entre les protéines.
EDTA	bloque l'activité des nucléases dépendantes des ions $Mg^{2+}$
DTT	empêche la création de ponts disulfures
chlorure de guanidinium	désassemble les capsides par déstabilisation des interactions hydrophobes entre les protéines
chlorure de lithium	précipite les acides nucléiques
sucrose	protège les protéines contre la déshydratation provoquée par la formation de la glace pendant la congélation.

10. Après avoir été formée, pendant une semaine, au MMSB, sur les différents protocoles de biochimie.

## 1.4.2 Cellules compétentes

En biologie, la compétence est la capacité d'une cellule à absorber une molécule d'ADN libre dans son environnement. Celle-ci peut être naturelle ou acquise par traitement en laboratoire, comme c'est le cas pour celles que nous utilisons. De plus, nous utilisons des bactéries modifiées (BL21\**CordonPlus*) qui contiennent un plasmide pRARE qui code des tRNA rares chez les bactéries et leur confère une résistance au chloramphénicol. Pour ce protocole, l'essentiel des manipulations doit être réalisé dans une zone stérile, soit dans un rayon de 10 cm autour de la flamme d'un bec bunsen.

1. 2 tubes Falcon de pré-culture stériles sont préparés en ajoutant 5 mL de milieu LB puis 5  $\mu\text{L}$  de chloramphénicol<sup>11</sup> à 34  $\text{mg}\cdot\text{mL}^{-1}$ . 1 colonie de cellules est prélevée à l'aide d'un cône stérile sur une boîte de Pétri puis ajoutée à l'un des 2 tubes. Le second est un témoin de stérilité. L'ensemble est mis à incuber sur la nuit ( $\sim 16$  h) à 200 rpm à 37°C.
2. Après incubation, si le milieu du tube de culture est trouble et le témoin de stérilité ne l'est pas, 200  $\mu\text{L}$  de chloramphénicol à 34  $\text{mg}\cdot\text{mL}^{-1}$  ainsi que les 5 mL de pré-culture sont ajoutés dans 200 mL de milieu LB. L'ensemble est mis à incuber, sur une table d'agitation à 200 rpm, à 37°C. La croissance bactérienne est suivie par turbidité du milieu via l'absorbance à 600 nm ( $\text{DO}_{600}$ ). Pour cela, nous utilisons des cuves à usage unique de trajet optique 10 mm. L'incubation est stoppée lorsque la  $\text{DO}_{600}$  atteint 0,5.
3. Après incubation, les 200 mL de culture sont répartis dans 4 tubes Falcon stériles de 50 mL, préalablement refroidis sur glace. L'ensemble est mis à incuber 20 min sur glace puis centrifugé à 2 800 rpm à 4°C pendant 10 min (rotor A-4-62).
4. Après centrifugation, les surnageants sont enlevés et 10 mL de  $\text{CaCl}_2$  à 0,1 M, préalablement refroidi à 4°C, sont ajoutés dans chaque tube Falcon. Les bactéries sont re-suspendues en pipetant doucement avant d'incuber 10 mn à 4°C.
5. Après incubation, l'ensemble est centrifugé à 2 800 rpm à 4°C pendant 10 min (rotor A-4-62).

---

11. La résistance au chloramphénicol permet la sélection des bactéries BL21\*CP, uniquement.



6. Après centrifugation, les surnageants sont enlevés et les culots re-suspendus dans 1 mL de  $\text{CaCl}_2$  à 0,1 M. L'ensemble est mis à incuber à 4°C pendant 1 h.
7. Après incubation, les solutions de bactéries sont récupérées et transférées dans un seul tube Falcon stérile.
8. Si les bactéries sont utilisées immédiatement, le volume de bactéries nécessaire peut être aliquoté dans des tubes eppendorf stériles. Si les bactéries vont être congelées, 140  $\mu\text{L}$  de DMSO sont ajoutés à la solution puis incubés 15 min à 4°C. L'opération est réalisée 2 fois avant d'aliquoter les bactéries dans des tubes eppendorf stériles de 600  $\mu\text{L}$ . Les bactéries sont conservées à -80°C.

### 1.4.3 Transformation des bactéries

La transformation des bactéries est le processus par lequel un plasmide<sup>12</sup> est introduit dans une cellule. Dans notre cas, le plasmide d'expression de la protéine porte un gène de résistance à la kanamycine. Pour ce protocole, l'essentiel des manipulations doit être réalisé dans une zone stérile, soit dans un rayon de 10 cm autour de la flamme d'un bec bunsen.

1. Les cellules compétentes, stockées dans le congélateur à -80°C, sont décongelées dans un bac à glace pendant environ 10 min.
2. Dans un tube eppendorf, 1  $\mu\text{L}$  de plasmide de concentration égale à 10  $\text{ng}\cdot\mu\text{L}^{-1}$  est ajouté à 100  $\mu\text{L}$  de cellules compétentes puis mélangé en pipetant délicatement. Le tube est ensuite mis à incuber 20 à 30 min dans un bac à glace.
3. Afin de faire pénétrer le plasmide, codant pour la protéine d'intérêt et pour la résistance à la kanamycine<sup>13</sup>, dans la bactérie, un choc thermique est réalisé : le mélange est placé dans un bain-marie à 42°C pendant 2 min avant de le remettre dans le bac à glace pendant 10 min.
4. Après incubation, 1 mL de milieu LB est ajouté puis pipeté délicatement avant d'être mis à incuber 1 h, sur table d'agitation à 180 rpm, à 37°C.

---

12. ADN étranger codant pour la protéine d'intérêt ainsi que pour un gène de résistance à un antibiotique permettant de marquer les bactéries porteuses du gène.

13. La résistance à la kanamycine permet la sélection des protéines d'intérêt, uniquement.

5. Après incubation, le tube eppendorf est centrifugé 2 min à 4000 rpm (rotor FA-45-30-11).
6. Après centrifugation, 950  $\mu\text{L}$  de surnageant est enlevé et le culot est resuspendu dans le surnageant restant. 50 à 150  $\mu\text{L}$  de la solution de bactéries sont déposés sur une boîte de Pétri préalablement préparée<sup>14</sup> et dispersés à l'aide d'un râteau stérile. La boîte de Pétri est ensuite mise à incuber, à l'étuve, à 37°C sur la nuit puis stockée à 4°C.

#### 1.4.4 Expression des protéines

L'expression est le processus par lequel les cellules hôtes sont amenées à produire les protéines d'intérêt en grand quantité. Pour ce protocole, l'essentiel des manipulations doit être réalisé dans une zone stérile, soit dans un rayon de 10 cm autour de la flamme d'un bec bunsen.

1. Les pré-cultures sont préparées en séparant 10 mL de milieu LB dans 2 tubes Falcon stériles de 15 mL (2x5 mL) puis en ajoutant 5  $\mu\text{L}$  de chloramphénicol et 5  $\mu\text{L}$  de kanamycine dans chaque tube. A l'aide d'un cône de pipette stérile, une colonie issue de la transformation présentée précédemment est ajoutée aux tubes Falcon. Un témoin de stérilité est également préparé. L'ensemble est mis à incuber sur la nuit ( $\sim 16$  h) à 180 rpm à 37°C.
2. Après incubation, si le milieu de la pré-culture est trouble et le témoin de stérilité ne l'est pas, 1 mL de chloramphénicol et de kanamycine ainsi que les 2x5 mL de pré-culture sont ajoutés dans 2x500 mL de milieu LB répartis dans 2 erlenmeyers de 2 L. L'ensemble est mis à incuber à 160 rpm à 37°C pendant 4 à 6 h et la  $\text{DO}_{600}$  doit être mesurée régulièrement.
3. Lorsque la  $\text{DO}_{600}$  est comprise entre 0,8 et 1,2, 1 mM d'IPTG est ajouté pour induire l'expression de la protéine puis mis à incuber sur la nuit ( $\sim 16$  h) à 180 rpm à 18°C.
4. Après incubation, les solutions de bactéries sont récupérées et transférées dans des bouteilles de centrifugation de contenance 400 mL et centrifugées à 6000 g à 4°C pendant 20 min (rotor JLA 10.500).

---

14. 15 mL de LB-agar + 15  $\mu\text{L}$  chloramphénicol à 34  $\text{mg}\cdot\text{mL}^{-1}$  + 15  $\mu\text{L}$  kanamycine à 50  $\text{mg}\cdot\text{mL}^{-1}$

5. Après centrifugation, le surnageant est éliminé. Dans le cas d'une utilisation à court terme, le culot est re-solubilisé dans du tampon de lyse (TN 50<sup>15</sup>) puis conservé à -80°C.

### 1.4.5 Purification des protéines

Le protocole est écrit pour la purification des capsides formées des protéines CP183 contenues dans un culot de 6 g re-suspendu dans 15 mL de tampon de lyse (TN50). L'essentiel des manipulations doit être réalisé à 4°C, donc en chambre froide ou à défaut en gardant les solutions dans un bac à glace.

1. Le culot de bactérie, stocké dans le congélateur à -80°C, est décongelé dans un bac à glace en chambre froide à 4°C sur la nuit.
2. Le lendemain matin, 250 µL de lysozyme<sup>16</sup> à 40 g.L<sup>-1</sup>, 200 µL d'un cocktail d'inhibiteurs de protéases<sup>17 18</sup> 50× et 50 µL de Triton-×-100<sup>19</sup> à 100 % sont ajoutés au culot puis l'ensemble est mis à incuber sur carrousel pendant 1 h.
3. Après incubation, 20 µL de benzonase nucléase<sup>20</sup> sont ajoutées. L'ensemble est mis à incuber à température ambiante pendant 45 min sur carrousel.
4. Après incubation, le culot est soniqué pendant 15 s en mode pulse puis refroidi dans le bac à glace pendant 50 s. Cette étape est réalisée 10 fois<sup>21</sup>.
5. Après sonication, la solution est transférée dans un tube adapté et centrifugée 30 min à 35 000 g à 4°C (rotor 70 Ti).
6. Après centrifugation, le surnageant est récupéré et un aliquote de 10 µL, noté S, est prélevé puis réservé pour l'étape 8. Une couche de 8 mL du surnageant est soigneusement ajoutée sur chacun des gradients de sucrose préparés la veille (voir figure ci-contre). L'ensemble est centrifugé 3 h à 100 000 g à 4°C (rotor SW-32 Ti).

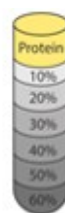


Figure 6.8: Schéma du gradient de sucrose.

---

15. Un culot de 5 à 6 g est généralement re-suspendu dans 15 mL de tampon.  
16. Le lysozyme permet de dégrader la paroi cellulaire bactérienne.  
17. L'inhibiteur de protéase permet de protéger les protéines d'intérêt de la dégradation par des protéases endogènes libérées lors de la purification.  
18. Produit Sigma, référence : P2714  
19. C'est un lipide utilisé pour perturber les membranes cellulaires et permettre la libération des matériaux intracellulaires sous forme soluble.  
20. La benzonase nucléase permet de digérer les acides nucléiques  
21. La sonication sert à casser les bactéries afin de récupérer le maximum de protéines.

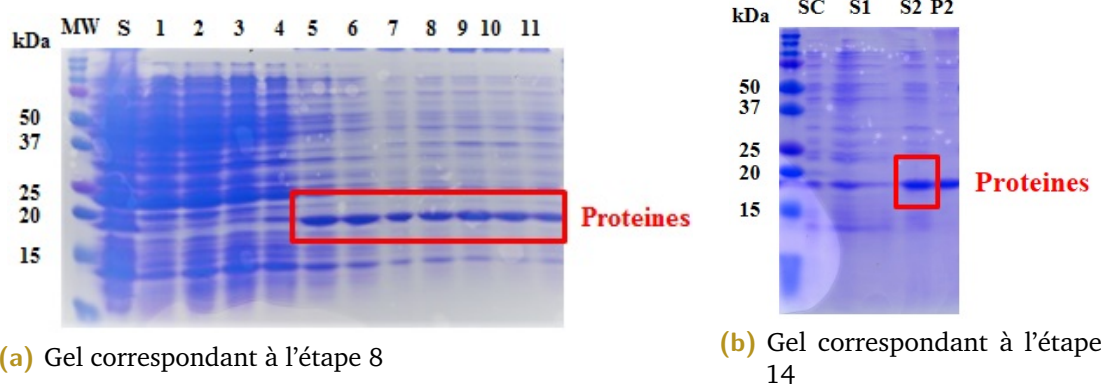
7. Après centrifugation, le surnageant est récupéré, en 13 fractions de 2,7 mL, en prenant garde à toujours prélever la partie la plus haute de la solution. Les fractions sont stockées dans des tubes séparés et conservées à 4°C. Un aliquote de 10  $\mu$ L de chaque fraction est prélevé, noté de 1 à 13 et réservé pour l'étape suivante.
8. Les fractions contenant les protéines sont identifiées en utilisant un gel SDS-PAGE<sup>22</sup> contenant 14% d'acrylamide<sup>23</sup> (voir figure 6.9a) puis mélangées. Un aliquote de 10  $\mu$ L du mélange est prélevé, noté SC et réservé pour l'étape .
9. Du sulfate d'ammonium est ajouté lentement à la solution de protéine pour obtenir un pourcentage final de 40%. L'ensemble est mélangé 1 h à température ambiante puis laissé au repos 2 h dans un bac à glace.
10. Après incubation, le mélange protéine-AMSO<sub>4</sub> est récupéré et centrifugé pendant 40 min à 20 000g à 4°C (rotor JA 25.50).
11. Après centrifugation, le surnageant est éliminé et le culot est solubilisé dans 11 mL de tampon de purification puis centrifugé pendant 20 mn à 20 000g à 4°C (rotor JA 25.50). Un aliquote de 10  $\mu$ L du surnageant est prélevé, noté S1 et réservé pour l'étape 14.
12. Après centrifugation, le surnageant est récupéré. Un aliquote de 10  $\mu$ L du surnageant est prélevé, noté S2 et réservé pour l'étape 14. Le culot est solubilisé dans 11 mL de tampon de purification. Un aliquote de 10  $\mu$ L du culot est prélevé, noté P2 et réservé pour l'étape 14. Les absorbances des solutions S1, S2 et P2 sont mesurées au Nanodrop puis la concentration en capsides est calculée<sup>24</sup>
13. La pureté de l'échantillon est contrôlée via un gel SDS-PAGE contenant 14% d'acrylamide.
14. La solution S2, qui contient les protéines, est ensuite dialysée, sur la nuit, dans 1 L de tampon de purification pour enlever le sulfate d'ammonium restant puis stockée à 4°C après vérifié la concentration.

---

22. Acronyme de l'anglais poly-acrylamide gel electrophoresis. C'est un méthode de séparation des molécules basée sur la différence de leurs poids moléculaires respectifs.

23. La pourcentage d'acrylamide définit le maillage du gel.

24. Usuellement, la concentration des protéines est calculée à partir de l'absorbance à 280 nm. Cependant, dans le cas des capsides de virus, la diffusion de la lumière augmente le rapport  $A_{260}/A_{280}$ . De ce fait, celle-ci est corrigée par les absorbances à 260, 340 et 360 nm pour tenir compte des acides nucléiques et de la diffusion de la lumière [116].



**Figure 6.9:** Gel SDS-PAGE 14% issu de la purification de la protéine native

## 1.4.6 Désassemblage des capsides

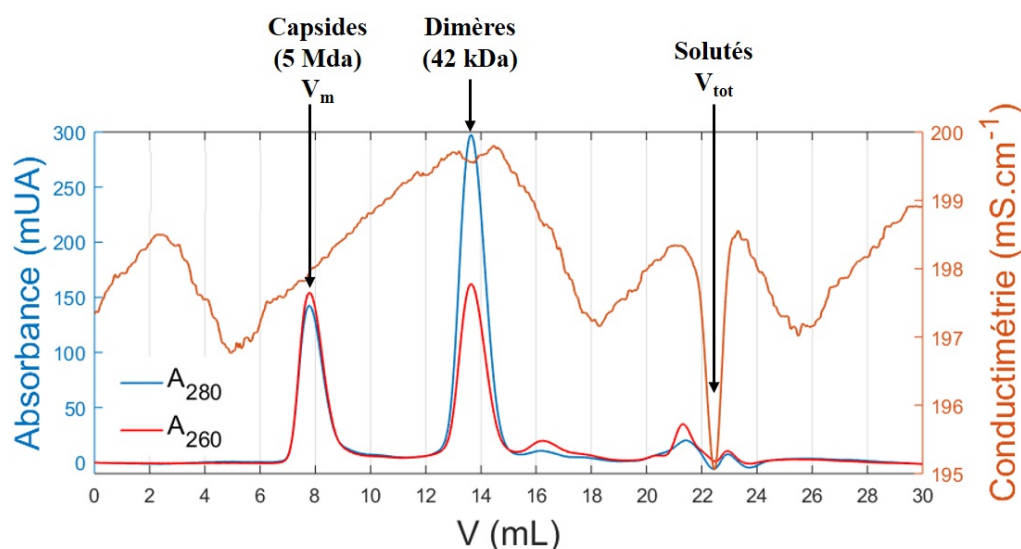
Le protocole est écrit pour le désassemblage des capsides formées des protéines CP183 préalablement purifiées. Les manipulations doivent être réalisées à 4°C, donc en chambre froide ou à défaut en gardant les solutions dans un bac à glace.

### 1.4.6.1 Protocole

1. La solution de protéines est dialysée dans 2 L de tampon de désassemblage toute la nuit à 4°C. Attention, la concentration de la solution doit être inférieure à 2 mg.mL<sup>-1</sup> pour que les capsides se dissocient.
2. Après dialyse, la solution de protéines est centrifugée 5 min à 1 500 g à 4°C (rotor FA-45-30-11) pour culotter les ARN précipités par le chlorure de lithium. La concentration de l'échantillon doit être vérifiée au Nanodrop.
3. Une chromatographie d'exclusion de taille est ensuite réalisée pour séparer les dimères des capsides. Un chromatogramme typique du désassemblage de la protéine native est visible sur la figure 6.10.
4. Après la SEC, la concentration du mélange correspondant aux fractions contenant les dimères est vérifiée au Nanodrop.

### 1.4.6.2 Remarques complémentaires

Sur la figure 6.10, la courbe bleue correspond à l'absorbance à 280 nm qui permet de quantifier la présence de protéines. La courbe rouge se rapporte à l'absorbance à



**Figure 6.10:** Chromatogramme classique de la protéine native (CP183) après dialyse dans le tampon de désassemblage.

260 nm qui permet de quantifier la présence d'acides nucléiques. La courbe orange permet de suivre l'évolution de la conductimétrie. La figure 6.10, met en évidence deux pics d'absorbance largement majoritaire. Le premier pic du chromatogramme, avec des absorbances à 260 et 280 nm équivalentes, correspond aux capsides non-dissociées. Celles-ci pèsent environ 5 MDa et sortent donc au volume mort, i.e. autour de 8 mL. La forte absorbance à 260 nm est due aux ARN bactériens encapsidés. Le second pic, avec une absorbance à 260 nm moitié moindre que celle à 280 nm, correspond aux dimères de protéines. Ceux-ci sont élués autour de 14 mL ce qui correspond à une protéine de 145 kDa, selon la calibration présentée dans la section 1.3. Cette masse étant calculée à partir d'une approximation globulaire basée sur le rayon hydrodynamique de la protéine, moins celle-ci est globulaire plus la masse est surestimée, ce qui est notre cas. Le pic inversé sur la courbe de conductimétrie correspond au volume d'éluion du soluté, i.e. le volume total.

# Les différentes stratégies pour avoir des protéines fluorescentes

Dans cette partie, nous verrons les différentes stratégies mises en oeuvre afin d'obtenir des dimères de protéines fluorescents en vue de l'étude *in vitro* de l'auto-assemblage des capsides par TIRFM. Pour ce faire, je présenterai les adaptations mises en place sur les différents protocoles pour palier aux écueils du fait des mutations. La première section sera consacrée au mutant GFP-CP183 tandis que la seconde exposera les résultats liés aux mutants lysines.

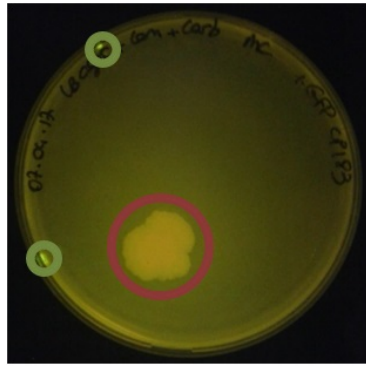
## 2.1 La protéine mutante GFP-CP183

Issue d'une méduse, la eGFP est constituée de 238 acides aminés pour une masse moléculaire d'environ 27 kDa. Celle-ci étant de masse comparable à celle de la protéine native, la production du mutant GFP-CP183 diffère de celle de la protéine native.

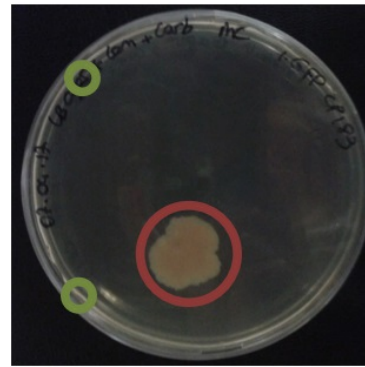
### 2.1.1 Impact de la GFP sur les premières étapes de production

#### 2.1.1.1 Transformation en bactéries

Lors de la transformation en bactérie (cf. sous-section 1.4.3), seules les bactéries qui intègrent le plasmide codant pour la protéine d'intérêt possèdent le gène de résistance à l'antibiotique associé et peuvent pousser. Le gène de résistance est utilisé comme marqueur des bactéries d'intérêt. Cependant, dans le cas où la protéine est toxique pour la bactérie, cela peut conduire à une instabilité du plasmide, i.e. une altération de la séquence, se traduisant par la croissance de bactéries résistantes à l'antibiotique mais n'exprimant pas la protéine d'intérêt. Sur la figure 6.11, est présentée une boîte de Pétri contenant des colonies après l'étape



(a) Utilisation de la lumière UV et du filtre orange



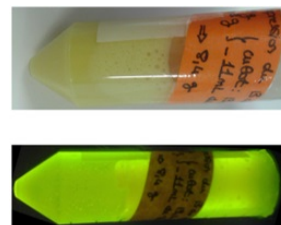
(b) Utilisation de la lumière visible

**Figure 6.11:** Photos d'une boîte de Pétri contenant des colonies BL21\*CP exprimant le gène de la protéine GFP-CP183.

de transformation bactérienne. A gauche, la boîte est exposée à de la lumière UV ; A droite, la boîte est exposée à de la lumière visible. Nous pouvons remarquer deux types de colonies différentes : les bactéries non-transformées sont visibles en rouge à la lumière visible mais n'émettent pas de fluorescence à la lumière UV. Les bactéries contenant les protéines GFP-CP183 émettent une fluorescence avec une lumière UV et un filtre orange, mais ne sont pas visibles avec une lumière visible. Quelques jours après la transformation bactérienne, il est important de vérifier sur plaque UV avec un filtre orange que les colonies fluorescent.

### 2.1.1.2 Expression des protéines

Lors de l'expression des protéines, une surexpression continue d'une protéine toxique induit un stress pour les bactéries. C'est le cas de la protéine GFP-CP183. Les températures d'incubation sont donc abaissées vers 23°C afin de réduire la pression sur les bactéries et les durées de culture sont allongées. Or, les cultures sur de longues périodes de temps peuvent sélectionner des bactéries non-productrices de GFP. En effet, les bactéries sécrètent de la  $\beta$ -lactamase, i.e. une enzyme qui consomme les antibiotiques en solution.



**Figure 6.12:** Observations par lumière visible (en haut) et UV (en bas) de la qualité du culot bactérien.



De ce fait, dans le cas d'une culture longue, la quantité d'antibiotiques diminue de manière non-négligeable au cours du temps. Il y a alors deux possibilités : Ajouter 1 mL d'antibiotiques frais avant l'induction ou culotter les bactéries et les re-suspendre dans du milieu LB frais contenant les antibiotiques. Dans tous les cas, il faut toujours vérifier sur la table UV que le culot bactérien obtenu fluoresce bien (Fig. 6.12).

### 2.1.1.3 Purification des protéines

Si les premières étapes de production de capsides, composées des protéines mutantes GFP-CP183, ont dû être grandement modifiées, celle de purification ne l'a que très peu été. Celles-ci peuvent être résumées ainsi :

- Le maillage des gels SDS-PAGE est modifié à 12 % du fait de la masse de la protéine.
- Les étapes 9 à 13 du protocole, i.e. étapes de précipitation au sulfate d'ammonium et de re-solubilisation à l'aide du tampon de purification, ont été supprimées.

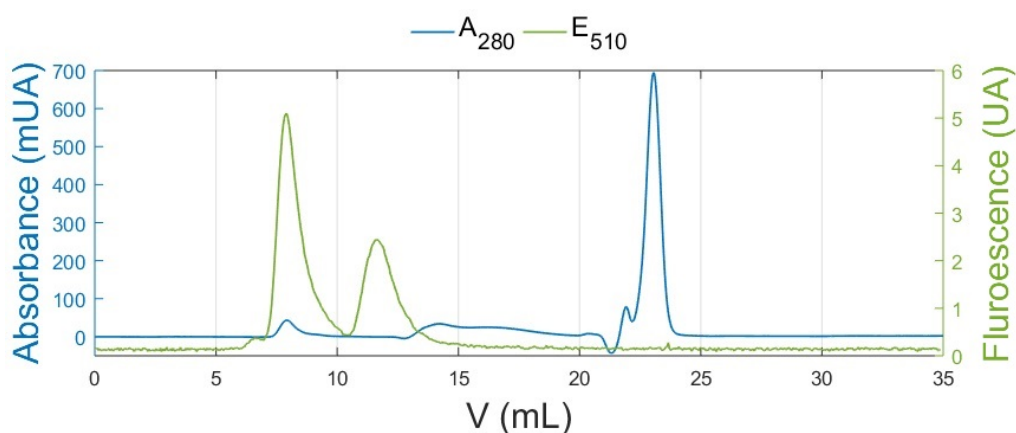
Cette seconde modification est due au fait que la précipitation au sulfate d'ammonium est une technique de purification, par précipitation, qui utilise la solubilité différentielle des protéines. Chaque protéine étant plus ou moins soluble en solution, celles-ci peuvent être séparées en fonction de leur tendance à précipiter plus ou moins vite en changeant la force ionique de la solution. Or, l'insertion eGFP modifie de manière conséquente la solubilité de la protéine, rendant cette étape inopérante.

## 2.1.2 Impact du protocole de désassemblage sur la fluorescence des dimères GFP-CP183

### 2.1.2.1 Désassemblage en 2 M chlorure de guanidinium

Comme nous avons pu le voir dans la section 1.4.6, le protocole de désassemblage des capsides composées des protéines natives est relativement simple : Nous dialysons les capsides dans une solution de désassemblage avant de séparer les

capsides restantes et les dimères par une SEC. Nous avons donc procédé de manière identique avec le mutant GFP-CP183 et obtenu le chromatogramme<sup>1</sup> suivant :



**Figure 6.13:** Chromatogramme de la protéine GFP-CP183 après dialyse dans le tampon de désassemblage classique.

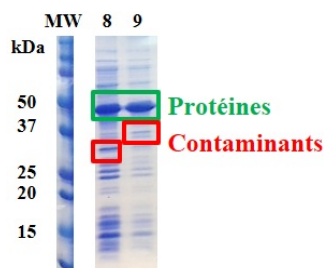
Sur la figure 6.13, nous pouvons observer deux courbes, bleue et verte, montrant, respectivement, l'évolution de l'absorbance à 280 nm et de la fluorescence à 510 nm en fonction du volume d'élution. La courbe A<sub>280</sub> présente un premier pic, avec une absorbance faible, situé autour de 8 mL. Comme nous avons pu le voir dans la section 1.4.6, un pic situé au niveau du volume mort correspond aux capsides de VHB non-dissociées. Sur ce chromatogramme, celui-ci est associé à un pic de fluorescence du fait de l'insertion eGFP. La courbe A<sub>280</sub> présente un second pic, beaucoup plus intense, autour de 23 mL. Celui-ci, même s'il n'est pas situé vers 17 mL comme prévu par la calibration présentée section 1.4.6, correspond bien aux dimères dissociés. La différence vient du fait que la matrice n'est pas totalement inerte vis à vis des protéines. Les protéines mutantes GFP-CP183 ayant une affinité avec la résine, l'élution de celles-ci est retardée.

Nous pouvons également constater que la courbe E<sub>510</sub> présente un second pic autour de 13 mL<sup>2</sup> qui ne semble pas associé à un pic d'absorbance. Ce pic pourrait être dû à une faible quantité de protéines GFP-CP183 tronquées, i.e. ayant subi une rupture au niveau d'un des linkers entre la eGFP et le spicule de la protéine native. Les structures 3D des deux monomères possibles de la protéine tronquée sont présentées sur la figure 6.15. Sur ces structures, ce sont les parties N-ter ou

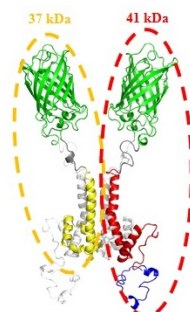
1. Colonne utilisée : S200 10/300 GL (24 mL)

2. Ce volume d'élution équivaut à une masse d'environ 150 kDa dans le cas d'une protéine globulaire.

C-ter de la protéine, en gris sur la figure, qui pourraient avoir été déletées. Sur la figure 6.14 est présenté un extrait du Gel SDS-PAGE correspondant à l'étape 8 du protocole de purification qui montre la présence de contaminants de tailles similaires aux structures proposées (37 et 41 kDa).

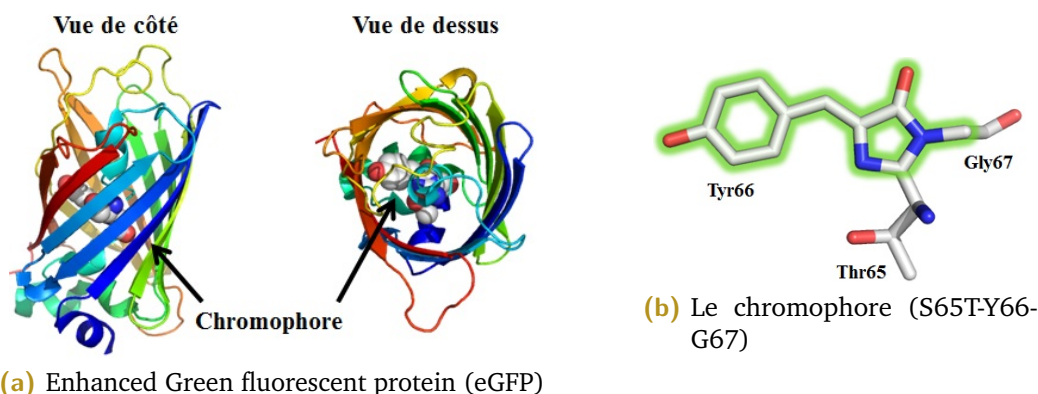


**Figure 6.14:** Gel SDS-PAGE correspondant à l'étape 8 du protocole de purification.



**Figure 6.15:** Structure 3D des possibles monomères de GFP-CP183 tronquées

Enfin, nous pouvons remarquer que le pic de dimères n'est pas associé à de la fluorescence. Ceci peut être expliqué par la structure de la eGFP<sup>3</sup>, montrée sur la figure 6.16. La partie responsable de la fluorescence s'appelle le chromophore. Elle



**Figure 6.16:** Structure 3D de la eGFP et de son chromophore.

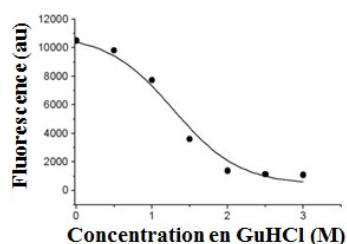
est située au centre du tonneau et mise en évidence par les billes grises, bleues et rouges. Les atomes du squelette de la eGFP forment une liaison spéciale qui donne une structure cyclique (Fig. 6.16b), essentielle pour la capacité de la protéine à absorber la lumière bleue (488 nm) et à émettre une fluorescence verte (510 nm). Or, comme rappelé au début de la section, nous utilisons un agent chaotrope afin de désassembler les capsides, c.à.d une molécule qui détruit la structure spatiale des macromolécules biologiques. Le chlorure de guanidinium dénature également

3. PDB ID : 2YOG

la structure tridimensionnelle du chromophore qui perd sa capacité à émettre de la fluorescence.

### 2.1.2.2 Désassemblage avec 0,5 M chlorure de guanidinium

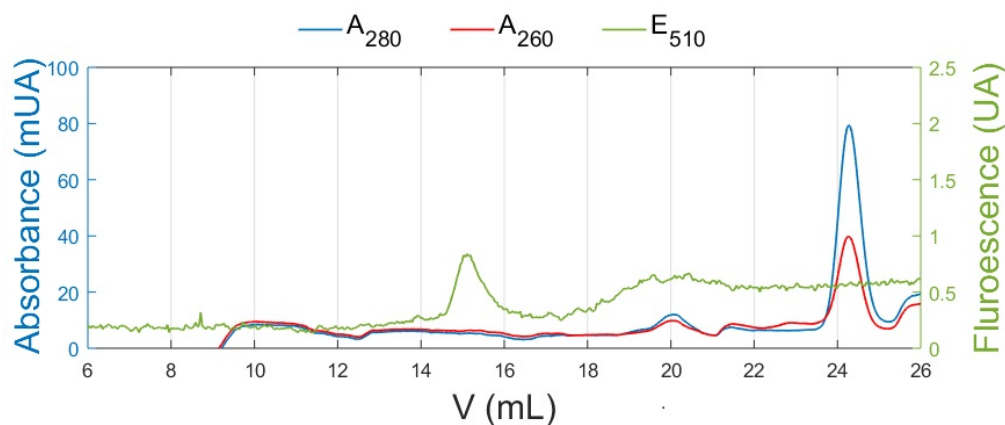
La stabilité de la GFP et de ses mutants ont été étudiés par différents groupes et depuis de nombreuses années [117][118]. En 2010, M. Palfi et al. ont étudié la dénaturation des protéines mutantes eGFP à différentes concentrations de chlorure de guanidinium, après une incubation à température ambiante pendant 16 heures. Leurs résultats sont présentés sur la figure 6.17. Sur celle-ci, adaptée de [119], la demi-dénaturation



**Figure 6.17:** Fluorescence de la eGFP à différentes concentrations de GuHCl.

est située à 1,3 M de GuHCl tandis que Jung et al. avaient rapporté une concentration au moins deux fois supérieure [120]. Cette différence montre que le phénomène de diminution de la fluorescence est encore méconnu et que des études supplémentaires sont nécessaires. Dans la mesure où nous souhaitons obtenir, de manière certaine, des dimères de protéines fluorescents, nous avons choisi d'essayer de désassembler les capsides de GFP-CP183 à une concentration en chlorure de guanidinium de 0,5 M. De plus, nous n'avons pas ajouté de chlorure de lithium dans le tampon de désassemblage. Les résultats sont présentés sur la figure 6.18.

Sur la figure 6.18, nous pouvons observer trois courbes. Les courbes rouge et bleue montrent l'évolution de l'absorbance à 260 et 280 nm en fonction du volume d'élution. Celles-ci présentent un pic vers 24 mL et le rapport des deux abondances indique que celui-ci correspond bien aux dimères. Cependant, la courbe verte, i.e. la fluorescence à 510 nm, ne présente toujours pas de pic associé indiquant que le protocole de désassemblage ne nous a pas permis d'obtenir de dimères fluorescents.



**Figure 6.18:** Chromatogramme de la protéine GFP-CP183 après dialyse dans le tampon avec 0,5 M de chlorure de lithium.

Nous pouvons noter la présence d'un pic faible de fluorescence autour de 14 mL. Celui-ci peut correspondre aux protéines GFP-CP183 tronquées présentées précédemment. En effet, les capsides utilisées proviennent de la même purification que celle présentée dans la section précédente.

### 2.1.2.3 Les différents tampons de désassemblage

La concentration en chlorure de guanidinium ne permettant pas d'expliquer l'absence de fluorescence sur le précédent chromatogramme, la production de protéine de GFP-CP183 a été entièrement refaite depuis l'étape de transformation. A l'issue de l'étape de purification, plusieurs tampons de désassemblage différents ont été testés et sont répertoriés dans le tableau 6.5 ci-dessous. Seuls les chromatogrammes présentant un résultat positif sont présentés sur la figure 6.19.

**Table 6.5:** Tableau récapitulatif des solutions tampons de désassemblage testées.

NOM	MODIFICATION
Tampon de désassemblage	20 mM Tris pH 7,5 ; 2 M GuHCl ; 0,5 M LiCl ; 10 mM DTT
D	0,5 M GuHCl
B	0,5 M GuHCl ; 0 M Sel
A	0,5 M GuHCl ; 0,5 M NaCl
C	pH9 ; 0,5 M GuHCl ; 0 M Sel
E	pH9 ; 0,5 M GuHCl

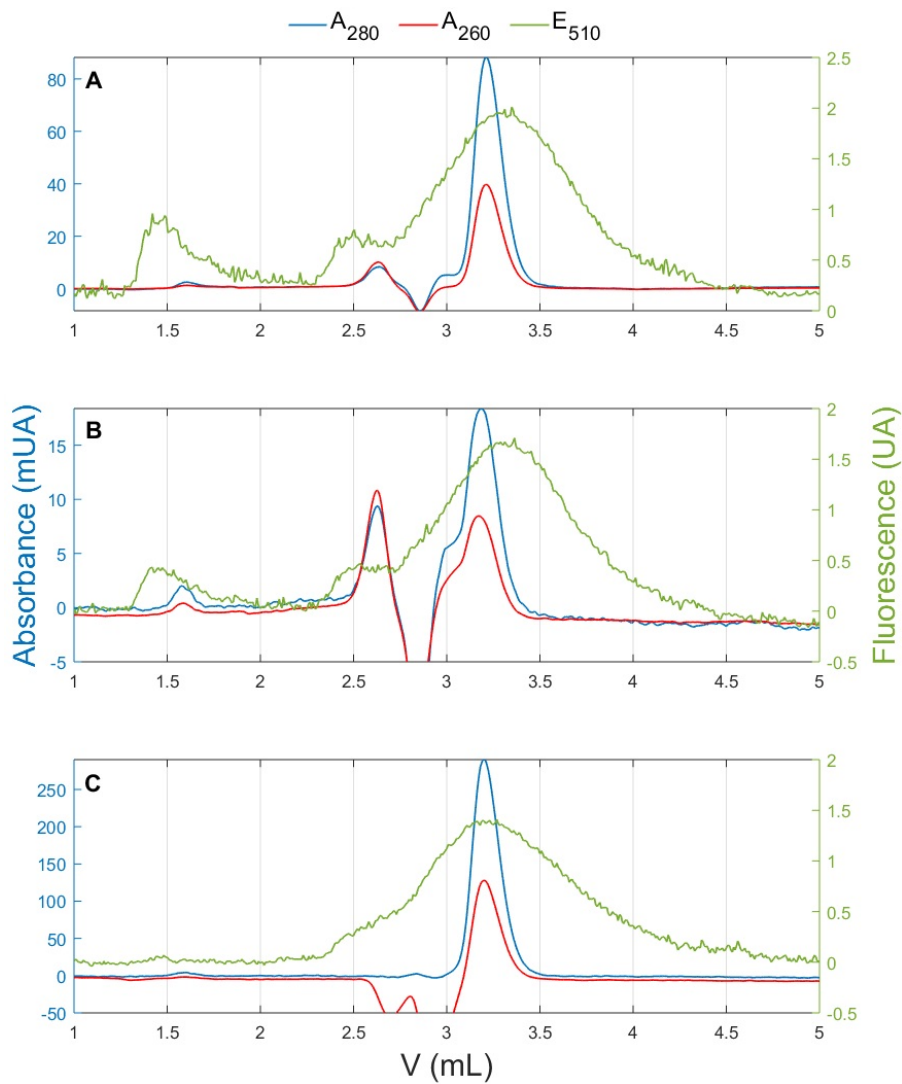
La colonne utilisée pour réaliser ces chromatogramme est la S200 5/150 GL, avec un volume mort de 1 mL et un volume total de 3 mL. La gamme de séparation

est identique à celle de la colonne 10/300 utilisée précédemment. Sur les trois chromatogrammes, nous pouvons observer trois courbes. Les courbes rouge et bleue montrent l'évolution de l'absorbance à 260 et 280 nm en fonction du volume d'élution ; La courbe verte présente la fluorescence à 510 nm. Les trois chromatogrammes présentent deux pics d'absorbances : le premier, situé un peu après 1,5 mL et associé à une fluorescence faible, correspond au pic de capsides non-dissociées. Le second, situé vers 3,2 mL et associé à une fluorescence légèrement plus importante, a un rapport  $A_{260}/A_{280}$  indiquant que celui-ci correspond bien aux dimères. Ces chromatogrammes indiquent que les tampons de désassemblage A, B et C permettent d'obtenir de dimères fluorescents.

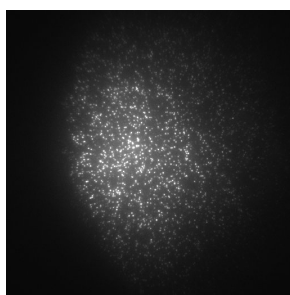
Dans un premier temps, nous pouvons noter que contrairement à la figure 6.18, le chromatogramme B de la figure 6.19 présente des dimères fluorescents. Or, le désassemblage a été effectué dans des tampons ayant la même composition. La différence entre les deux chromatogrammes ne provient pas de l'étape de désassemblage mais du stockage entre les protocoles d'expression et de purification. En effet, dans le premier cas, les culots bactériens avaient été congelés immédiatement après l'expression tandis que dans le second, ceux-ci ont été stockés pendant 48h en chambre froide. Or, la formation du chromophore final nécessite une étape d'oxydation d'une durée de plusieurs heures [121]. Les bactéries ne doivent donc pas être congelées immédiatement, il faut les laisser environ 24 h dans un bac à glace ou en chambre froide avant congélation.

Nous pouvons ensuite noter que si l'amplitude du signal de fluorescence est similaire pour les trois chromatogrammes (1,5 mUA), celle des courbes  $A_{260}$  et  $A_{280}$  ne l'est pas. En effet, les maxima des absorbances à 280 nm des chromatogrammes A, B et C sont situés autour de 90, 20 et 290 mUA, respectivement. La fluorescence par dimère n'est donc pas équivalente. Nous avons testé deux échantillons par TIRFM et obtenu les deux images présentées sur la figure 6.20. L'image 6.20a et l'image 6.20b correspondent aux chromatogrammes A et C, respectivement.

Lors de ces tests, les concentrations en dimères utilisées sont équivalentes mais pas le signal de fluorescence. Sur la figure de gauche (tampon pH neutre), le nombre de taches lumineuses est très supérieur à celui sur la figure de droite (tampon basique). Cette différence ne peut être imputée à la différence de pH entre les deux échantillons. En effet, comme nous pouvons le voir sur la figure



**Figure 6.19:** Chromatogrammes de la protéine GFP-CP183 après dialyse dans différents tampons avec 0,5 M de GuHCl. Les conditions de  $pH$  et de concentrations en sel sont résumées dans le tableau 6.5.

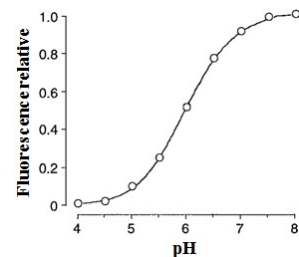


(a) Image correspondant au chromatogramme A



(b) Image correspondant au chromatogramme C

**Figure 6.20:** Image en TIRFM des protéines désassemblées des tampons avec 0,5 M de GuHCl



**Figure 6.21:** Fluorescence de la eGFP en fonction du  $pH$

6.21, la fluorescence de la GFP est stable entre  $pH$  7,5 et 9. Celle-ci provient plus probablement de la concentration en sel des solutions de désassemblage. En effet, dans une publication de 2004, les auteurs étudient la cinétique de dénaturation d'un mutant<sup>4</sup> de la GFP et remarquent que celle-ci est moins prononcée pour de plus fortes salinités. Les auteurs expliquent cette différence par la liaison des ions chlorure à la protéine dénaturée, qui minimiserait la répulsion des charges intramoléculaires et stabiliserait la structure secondaire retenue dans la protéine dénaturée [122].

Dans les deux cas, nous avons également pu constater que le signal de fluorescence n'est pas stable, i.e. les GFP scintillent. La photo-instabilité des protéines pourrait être due à la composition du milieu nutritif utilisé pendant l'étape d'expression des protéines. En effet, des études récentes ont montré que la photo-stabilité de la protéine eGFP pouvait être améliorée en l'absence de riboflavine dans le milieu nutritif ou en augmentant la densité cellulaire [123] [124].

## 2.2 Les mutants Lysines

Comme indiqué dans la section 1.2.1, nous avons également travaillé avec des mutants ponctuels, générés par substitution du glutamate 77 ou de l'alanine 80 en lysine, afin de pouvoir étiqueter chimiquement la protéine avec un marqueur fluorescent.

### 2.2.1 Présentation du fluorophore

Le fluorophore que nous avons utilisé est l'ATTO 647N NHS-ester qui absorbe et émet dans la partie rouge du spectre visible. Les longueurs d'onde d'absorbance et de fluorescence maximum sont de 647 et 667 nm, respectivement. Le marqueur porte un groupe réactif NHS-ester qui interagit avec les groupements amines des résidus des lysines pour former une liaison covalente. La réaction de marquage se produit comme dépeint sur la figure 6.22a. Le groupement amine de la chaîne latérale ayant un  $pK_a$  de 9, pour des  $pH$  situés entre 7 et 9, ce groupe est partiellement déprotoné<sup>5</sup>. Dans ces conditions, le réactif NHS-ester peut former une

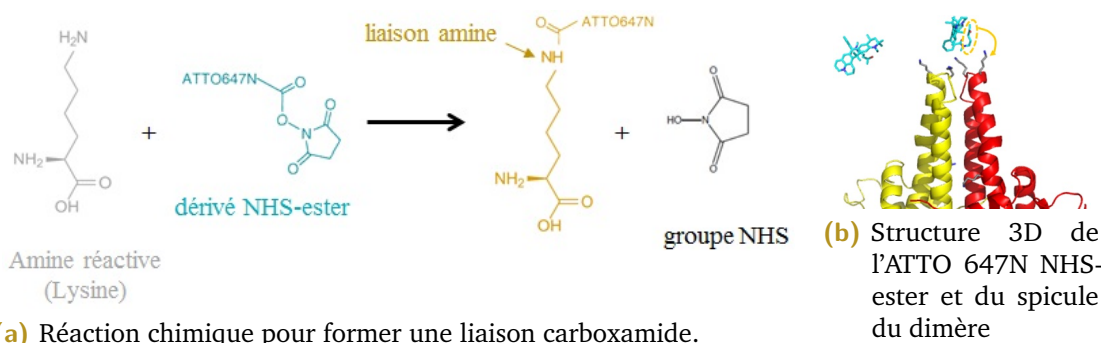
---

4. F99S/M153T/V163A

5. par ex.  $pH$  9 50% déprotoné,  $pH$  8 10% déprotoné



liaison carboxamide (Fig. 6.22b). Cependant, pour des  $pH$  trop basiques ( $pH > 9$ ) la protéine peut se dénaturer. De ce fait, le  $pH$  de travail choisi est un compromis entre ces deux contraintes.



**Figure 6.22:** Réaction de marquage du dimère de protéines du VHB par l'ATTO 647N NHS-ester.

## 2.2.2 Protocole de marquage

Si, contrairement au mutant GFP-CP183, les mutations ponctuelles n'ont pas nécessité d'adaptation des différents protocoles, il a tout de même fallu modifier le tampon utilisé lors de la purification et donc pour le marquage. En effet, le Tris a dû être remplacé par d'autres tampons (par ex.  $KH_2PO_4$ ,  $NaHCO_3$ ) puisqu'il contient un groupement amine qui peut réagir avec le marqueur fluorescent.

Le présent protocole est écrit pour étiqueter les capsides mutées du VHB avec des marqueurs fluorescents ATTO 647N NHS-ester. Ceux-ci ont un poids moléculaire de  $843 \text{ g.mol}^{-1}$  et sont stockés dans 100% de DMSO à une concentration telle que  $c_{ATTO} = 2,372 \text{ mM}$ . Le marquage est réalisé lorsque les protéines sont sous forme de capside de façon à éviter que les deux lysines situées dans le corps de la protéine ne soient exposées. Afin de vérifier le bon déroulement du marquage, un aliquote de  $20 \mu\text{L}$  est prélevé à chaque étape du protocole.

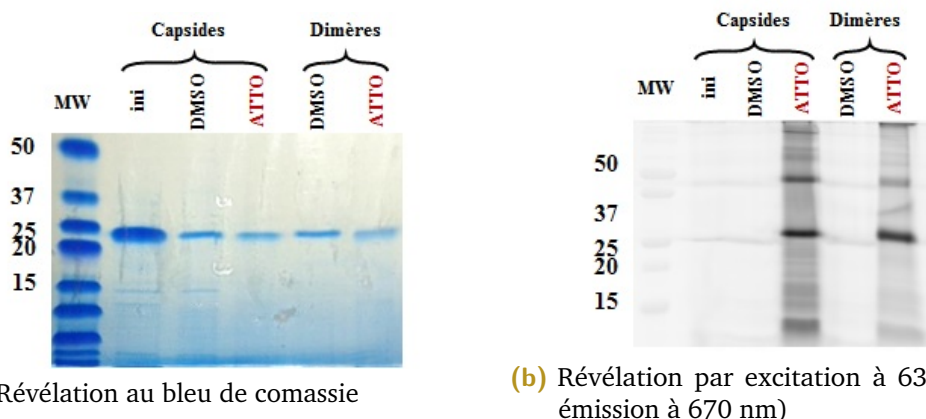
1. Le volume de marqueur à prélever est calculé en fonction de trois paramètres : le taux de marquage souhaité  $\tau$ , la concentration en protéines  $c_{prot}$  ainsi que le volume de la solution de protéines  $V_{prot}$  tel que :

$$V_{ATTO} = \frac{c_{prot} \times \tau \times V_{prot}}{c_{ATTO}} \quad (2.1)$$

La quantité ainsi calculée est mélangée à la solution de protéines puis mise à incuber à  $22^\circ\text{C}$  à 450 rpm pendant 45 min (bloc ThermoMixer<sup>TM</sup> C). Un

contrôle négatif est réalisé en mélangeant un volume équivalent de DMSO à la solution de protéines.

- Après incubation, l'ensemble est dialysé dans 1 L de tampon de désassemblage, sur la nuit à 4°C. Attention, la concentration de la solution doit être inférieure à 2 mg.mL<sup>-1</sup> pour que les capsides se dissocient.
- Après dialyse, la solution de protéines est centrifugée 5 min à 1 500 g à 4°C (rotor F-45-24-11) pour culotter les ARN précipités par le chlorure de lithium.
- Après centrifugation, une chromatographie d'exclusion de taille est réalisée pour séparer les dimères des capsides.
- Le bon déroulement du marquage est contrôlé en déposant les 5 aliquotes, résumé dans le tableau 6.6, sur un gel SDS-PAGE Tris-Tricine contenant 16,5% d'acrylamide. Le volume de dépôt est ajusté pour obtenir la même concentration théorique dans chaque puits. Autrement dit, si la quantité de protéine à chaque étape ne varie pas, alors l'intensité de chaque bande sera équivalente.



**Figure 6.23:** Gel SDS-PAGE Tris-Tricine, contenant 16,5% d'acrylamide, du marquage de la protéine A80K-CP183

**Table 6.6:** Tableau récapitulatif des aliquotes prélevés lors du marquage et fonctions.

ALIQUOTES DE PROTÉINES PRÉLEVÉS	FONCTION
Post-purification	Référence
Mélange protéines-DMSO	Contrôle négatif
Mélange protéines-marqueur	Vérification de l'efficacité du marquage
Post-désassemblage (DMSO et ATTO)	Vérification de la dissociation des capsides et du maintien de marquage

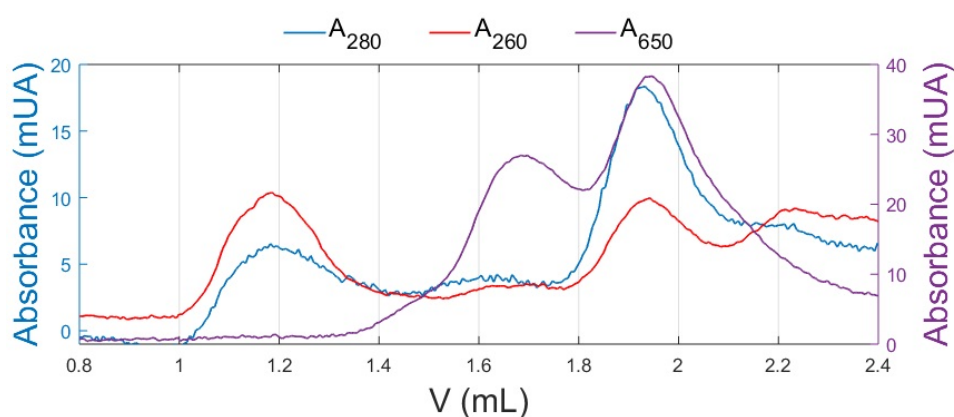
Comme nous pouvons le voir sur la figure 6.23 le marquage semble fonctionner. En effet, sur la figure 6.23b, lorsque le gel est observé par fluorescence, nous retrouvons deux bandes, indiquant la présence de protéines d'environ 20 kDa, dans les puits correspondant à la protéine marquée et non dans ceux des contrôles négatifs. Sur la figure 6.23a, lorsque le gel est observé par coloration classique, nous pouvons observer une bande correspondant à la protéine dans chaque puits.

## 2.2.3 Conditions de marquage des mutants lysines

Dans cette section, je présenterai les différentes conditions de marquages utilisées ainsi que leur impact sur la dissociation.

### 2.2.3.1 Le tampon phosphate de potassium

Les premières conditions de marquage que nous avons utilisées sont les suivantes : le tampon utilisé pour la purification ainsi que le désassemblage est le phosphate de potassium ( $\text{KH}_2\text{PO}_4$ ) et le ratio de marquage était 1 protéine pour 5 ATTO. Ce ratio a été choisi en se basant sur les recommandations fournies avec le marqueur. Nous avons ainsi obtenu le chromatogramme présenté sur la figure 6.24. Celui-ci a été réalisé avec la colonne S200 5/150 GL (3 mL) qui permet de réaliser, rapidement, des tests préliminaires.



**Figure 6.24:** Chromatogramme de la protéine A80K-CP183 après désassemblage dans le tampon phosphate de potassium.

Sur ce chromatogramme, et les suivants, nous pouvons observer trois courbes :

- Les courbes rouge et bleue montrent l'évolution de l'absorbance à 260 et 280 nm en fonction du volume d'élution. Celles-ci présentent un premier pic vers 1,2 mL et le rapport des deux abondances indique que celui-ci correspond aux capsides de VHB non-dissociées. Le second pic, situé vers 1,9 mL, correspond aux dimères.
- La courbe violette présente l'évolution de l'absorbance à 650 nm, caractéristique de la présence du marqueur fluorescent, en fonction du volume d'élution. Sur celle-ci, nous pouvons également observer deux pics. Cependant, seul le second est identifiable puisqu'il est associé au pic de dimères, indiquant que ceux-ci sont marqués.

Nous pouvons remarquer que les pics du chromatogramme sont peu intenses indiquant des concentrations finales très inférieures à celles attendues. Cette observation est confirmée par la figure 6.23a de la section précédente. Sur celle-ci, nous pouvons observer que les intensités des bandes correspondant aux aliquotes prélevés après l'incubation avec le marqueur stockés dans le DMSO ou, le DMSO seul, sont toutes très inférieures à celle de la première bande, i.e. avant le marquage. Ajouté à cela, un spectre d'absorbance a été réalisé sur le pic de dimères, à l'issue de la SEC. Celui-ci indique une concentration en protéine inférieure à  $0,1 \text{ mg.mL}^{-1}$ . Cette diminution de la quantité de protéine peut être due au tampon de stockage du marqueur, i.e. le DMSO. En effet, celui-ci déstabilise la structure secondaire des protéines ce qui peut générer de l'agrégation. De ce fait, après marquage, lorsque la solution de protéines est centrifugée pour culotter les ARN précipités par le chlorure de lithium, il est possible qu'une partie des protéines ait également précipité. De ce fait, le protocole de marquage a été modifié. Ainsi, après incubation, les protéines marquées et le contrôle négatif sont d'abord dialysés dans  $2 \times 500 \text{ mL}$  de tampon de purification pendant  $2 \times 2 \text{ h}$  à  $4^\circ\text{C}$  afin de retirer les marqueurs surnuméraires libres en solution ainsi que le DMSO.

### 2.2.3.2 Le tampon bicarbonate de sodium

Si la colonne 5/150 permet une première approche, celle-ci est peu résolutive et les résultats peuvent être difficiles à interpréter. De ce fait, une confirmation de ces résultats préliminaires doit être obtenue avec la colonne 10/300 (24 mL). Pour ce faire, nous avons réitéré l'expérience à l'identique. Nous avons également testé un

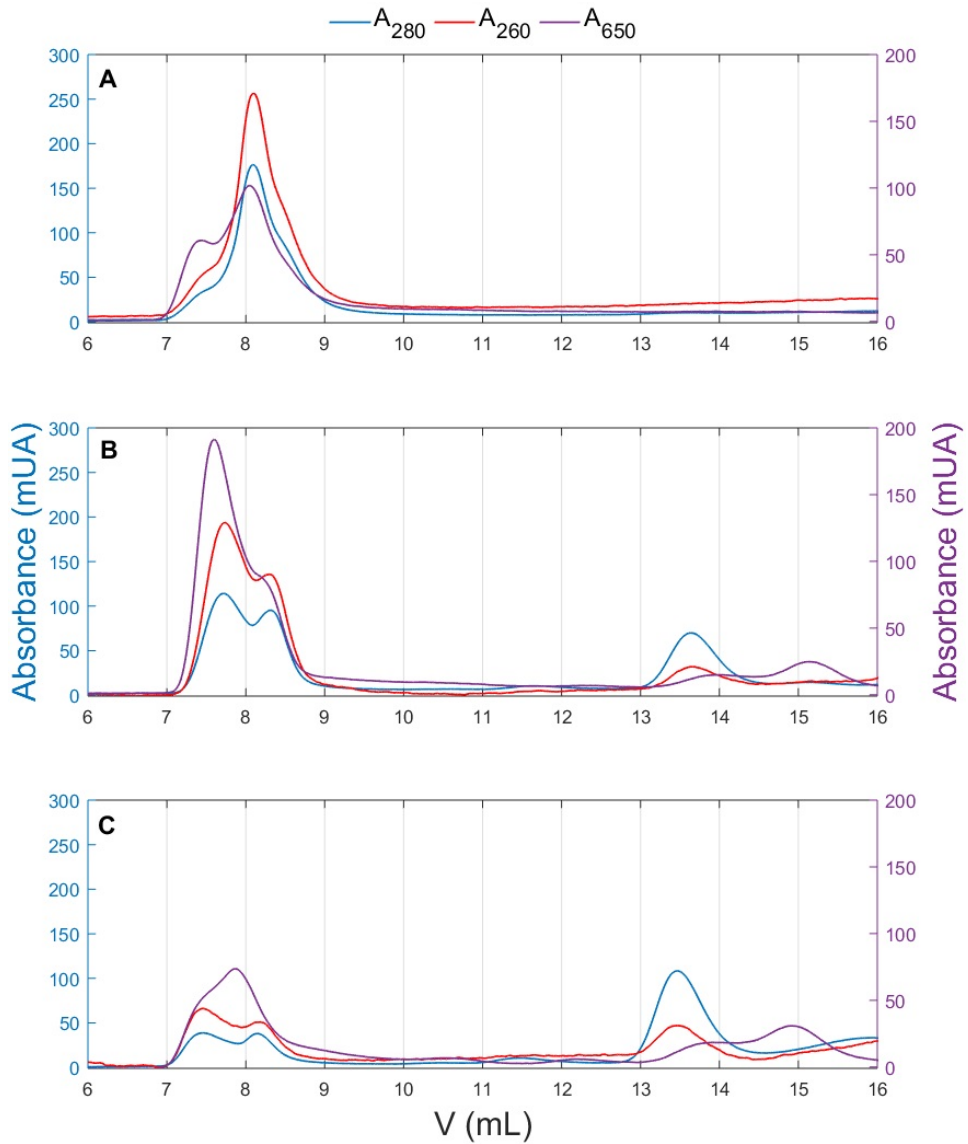
autre tampon, le bicarbonate de sodium pour comparer les résultats. Les conditions de marquage que nous avons utilisées sont les suivantes :

- Le ratio de marquage a été conservé, i.e. 1 protéine pour 5 ATTO.
- Le premier chromatogramme montre les résultats lorsque le tampon utilisé pour la purification et le désassemblage est le phosphate de potassium ( $\text{KH}_2\text{PO}_4$ ).
- Le deuxième chromatogramme montre les résultats lorsque le tampon utilisé pour la purification est le phosphate de potassium et le tampon de désassemblage est le bicarbonate de sodium ( $\text{NaHCO}_3$ )
- Le troisième chromatogramme montre les résultats lorsque le tampon utilisé pour la purification et le désassemblage est le bicarbonate de sodium.

L'ensemble de ces tests de marquage a été réalisé, en parallèle, dans les mêmes conditions, avec des protéines issues de la même purification. Les différences ne proviennent donc que des tampons utilisés.

Sur le chromatogramme A, nous pouvons observer la présence d'un seul pic avec un épaulement, situé vers 8 mL, qui correspond aux capsides marquées non dissociées. Nous n'observons donc pas de dissociation contrairement au résultat présenté précédemment. Cette différence, entre les deux chromatogrammes, est difficile à expliquer. Celle-ci provient peut être du fait que le nombre de protéines marquées par capside est supérieur à celui de l'expérience précédente. Une autre hypothèse peut être une répartition plus inégale des marqueurs à la surface de la capside. Dans les deux cas, les marqueurs ayant tendance à agréger lorsqu'ils sont très concentrés, ces hétérogénéités pourraient rendre plus difficile le désassemblage et expliquer cette différence.

Sur le deuxième chromatogramme, B, nous pouvons observer la présence de trois pics à 260 et 280 nm. Les deux premiers, situés entre 7,5 et 8,5 mL correspondent aux capsides non-dissociées. Nous pouvons remarquer que, si les deux pics sont associés à de l'absorbance à 650 nm, celle du premier est deux fois plus importante que celle du second. Ceci semble indiquer deux populations de capsides distinctes, i.e. avec un nombre de protéines marquées par capside très différent. Nous pouvons également observer un troisième pic à 280 nm, situé vers 13,5 mL, avec un maximum d'environ 70 mUA. Celui-ci, deux fois moins intense que celui des capsides, correspond aux dimères libres. Le pic d'absorbance à 650 nm, associé aux



**Figure 6.25:** Chromatogrammes avec la protéine A80K-CP183 purifiée et désassemblée dans différents tampons. **A** Purification et désassemblage en  $\text{KH}_2\text{PO}_4$ , **B** Purification en  $\text{KH}_2\text{PO}_4$  et désassemblage en  $\text{NaHCO}_3$ , **C** Purification et désassemblage en  $\text{NaHCO}_3$ .

dimères marqués, a une amplitude faible ( $\approx 14$  mUA). Nous pouvons noter que le maximum de celui-ci est décalé vers 14 mL. Ceci semble indiquer deux populations de dimères, marqués et non-marqués. Nous pouvons également noter la présence d'un pic d'absorbance à 650 nm vers 15 mL que nous ne pouvons identifier.

Sur le troisième chromatogramme, C, nous pouvons également observer deux pics de capsides, autour de 8 mL et un pic de dimères libres autour de 13,5 mL. Dans ce cas, les pics des capsides à 280 nm sont deux fois moins intenses que celui des dimères ( $\approx 2 \times 33$  mUA), indiquant que la dissociation a été plus effective. Nous

pouvons également observer que l'absorbance à 650 nm est similaire à ce que nous avons pu décrire sur le chromatogramme B. Deux populations de dimères, marqués et non-marqués, sont présents en solution. Le maximum d'absorbance à 650 nm pour les dimères est d'environ 18 mUA.

Ces résultats, associés à l'équation 2.2, permettent d'estimer approximativement le nombre de marqueurs fluorescents par protéine pour chaque expérience.

$$DOL = \frac{c_{ATTO}}{c_{prot}} = \frac{A_{650} \times \epsilon_{280}}{A_{280} \times \epsilon_{650}} \quad (2.2)$$

où  $\epsilon_{650}$  est le coefficient d'extinction molaire de l'ATTO à 650 nm. Dans le cas de l'ATTO libre, celui-ci vaut  $1,5 \cdot 10^5 \text{ M}^{-1}\text{cm}^{-1}$ . Dans le cas de l'ATTO conjugué à la protéine du VHB, celui-ci a été estimé à environ  $4 \cdot 10^5 \text{ M}^{-1}\text{cm}^{-1}$

Dans le cas où le marquage a été effectué en phosphate de potassium et le désassemblage a été réalisé en phosphate de potassium ( $\text{NaHCO}_3$ ), nous pouvons estimer qu'un peu moins de 1,5 % des protéines, sous forme de dimères, ont été marquées. Dans le cas où le marquage et le désassemblage ont été réalisés en phosphate de potassium, nous pouvons évaluer qu'environ 1,7% des protéines ont été marquées. Dans la mesure où les dimères peuvent être marqués 0, 1 ou 2 fois, cela signifie qu'au mieux, 2 dimères parmi les 120 nécessaires à la formation de la capsidie peuvent être détectés par TIRFM, ce qui est insuffisant.

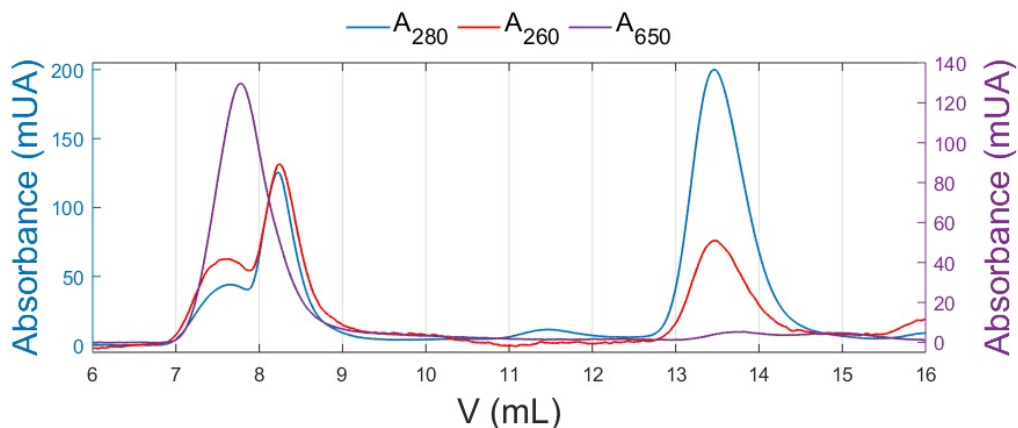
Même sans DMSO pour déstabiliser la structure secondaire, celles-ci, ainsi que les marqueurs, ont tendance à s'agréger. Encore une fois, le faible ratio final provient sans doute d'un problème d'agrégation. Une possibilité pour améliorer ce résultat est de modifier le ratio protéine-marqueur lors du marquage. En effet, réduire drastiquement la quantité de marqueurs pourrait permettre une répartition plus homogène de ceux-ci sur la capsidie. Cela aurait pour effet d'augmenter l'efficacité du désassemblage et donc d'obtenir, *in fine*, plus de dimères marqués.

### 2.2.3.3 Ratio protéines-marqueurs

Afin de confirmer l'hypothèse proposée précédemment, nous avons réitéré l'expérience. Pour ce faire, nous avons choisi les meilleures conditions<sup>6</sup> et fixé le ratio

6. marquage et désassemblage en tampon phosphate de potassium

de marquage à 0,1 ATTO pour 1 protéine. Le temps d'incubation a également été modifié, du fait de la faible concentration en marqueurs. Nous avons testé trois temps d'incubation, 1h15, 2h30 et 5h. Les résultats pour ces trois temps d'incubation étant identiques, seul le chromatogramme avec un temps d'incubation de 1h15 est présenté.



**Figure 6.26:** Chromatogramme avec la protéine A80K-EEE-CP183 et un ratio de marquage de 10%.

Le chromatogramme, présenté sur la figure 6.26, est très similaire à ceux présentés précédemment. Nous pouvons observer deux pics de capsides, autour de 8 mL, et un pic de dimères libres autour de 13,5 mL. Comme pour le chromatogramme C de la figure 6.25, les pics des capsides à 280 nm sont moins intenses que celui des dimères. Les maximums d'absorbance à 280 nm et à 650 nm, du pic des dimères, sont de 200 mUA et 5 mUA, respectivement. Ceci indique que moins de 0,2 % des protéines, sous forme de dimères, ont été marquées. Le protocole de marquage employé a donc été moins efficace que le précédent.

#### 2.2.3.4 Conclusion

Dans cette section, nous avons effectué des tests de marquage sur les protéines mutantes du VHB. Nous avons pu constater que seule une faible proportion des protéines, marquées ou non, ressortent des différentes dialyses puis de la SEC. De plus, une grande partie des échantillons sont peu ou pas dissociés. L'ensemble de ces éléments indiquent une forte propension à l'agrégation des protéines marquées, qu'elles soient sous forme de capsides ou de dimères. Cette agrégation diminue fortement la quantité finale obtenue de dimères marqués. Néanmoins,



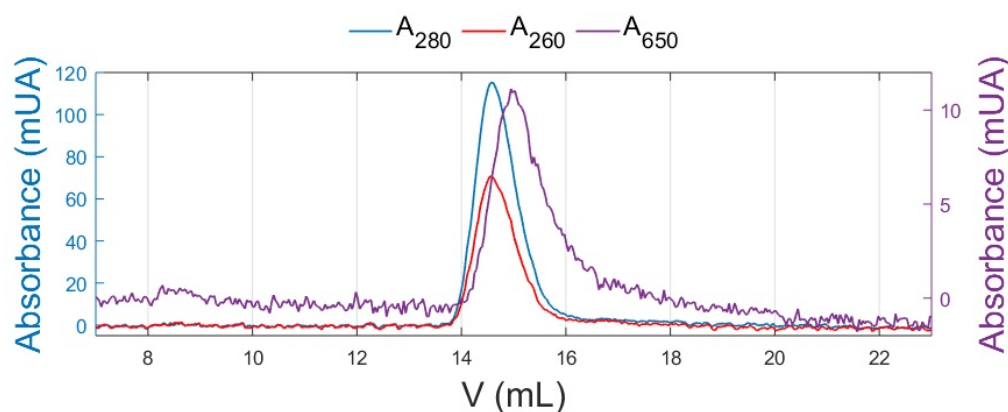
plusieurs possibilités sont envisageables pour essayer d'améliorer l'efficacité du marquage :

- Préalablement, diluer le marqueur dans la solution tampon de marquage puis ajouter un volume équivalent à la solution de protéine.
- Réaliser le mélange plus lentement, au goutte à goutte par exemple.
- Diminuer le  $pH$  afin de limiter la réaction de marquage.

Dans les deux premiers cas, l'idée serait de voir l'impact d'un mélange plus homogène sur le marquage. Le dernier cas pourrait permettre de limiter le phénomène d'agrégation.

## 2.2.4 Marquage du CCMV

Afin de mieux comprendre les écueils précédents, nous avons réitéré l'expérience précédente à l'identique mais avec un autre virus. Nous avons donc utilisé le CCMV, présenté chapitre 2, au lieu du VHB. Le chromatogramme obtenu est présenté ci-dessous.



**Figure 6.27:** Chromatogramme avec la protéine du CCMV et un ratio de marquage de 10%.

Sur la figure 6.27, nous pouvons observer la présence d'un seul pic d'absorbance à 260 et 280 nm, situé autour de 14,5 mL. Le rapport des absorbances nous indique que le pic correspond bien aux dimères libres. A ce pic est associé un pic d'absorbance à 650 nm. Comme pour le VHB, celui-ci est décalé vers 15 mL, indiquant également la présence de deux populations de dimères de protéines, une marquée et une non-marquée. Les maximum d'absorbance à 280 nm et à 650 nm,

du pic des dimères, sont de 115 mUA et 11 mUA, respectivement, ce qui indique qu'environ 4 % des dimères de protéines sont marqués.

Le marquage et le désassemblage ont été plus efficaces avec le CCMV qu'avec le VHB. Des améliorations sont tout de même encore possibles :

- Réduire le débit lors de la SEC.
- Utiliser une colonne avec un volume plus important, par ex. la S200 16/600 de 120 mL.

Dans les deux cas, l'idée serait de séparer les dimères marqués, des non-marqués, lors de la SEC. De cette façon, le pourcentage de dimères marqués en solution pourrait être accru.

# L'ARN pré-génomique

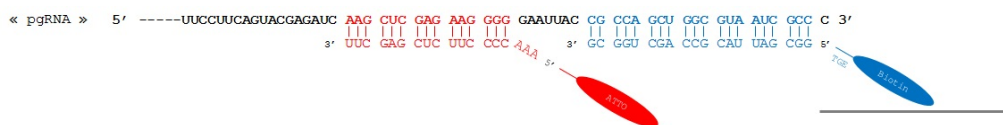
Dans cette partie, nous verrons la préparation du génome en vue de l'étude *in vitro* de l'auto-assemblage des capsides par TIRFM. Pour ce faire, je commencerai par présenter succinctement le cycle viral du VHB ce qui introduira naturellement l'ARN pré-génomique. La section suivante est consacrée à la présentation des mutations réalisées sur celui-ci afin de pouvoir réaliser les expériences de TIRFM. Je terminerai cette partie avec les résultats préliminaires obtenus en TIRFM.

## 3.1 Préparation de l'ARN pré-génomique

Comme nous avons pu le voir dans le chapitre 3.2, le cycle de réplication du VHB est complexe et encore méconnu. La réplication, divisible en 7 phases, comporte 5 états différents du génome. Un point important, pour l'étude par TIRFM de l'auto-assemblage de la capside du VHB, est de savoir que seule la nucléo-capside associée à la polymérase est susceptible d'être infectieuse. De ce fait, notre étude qui s'intéresse à la phase 5 du cycle présenté sur la figure 2.14, ne nécessite pas de conditions de sécurité particulière. En effet, pour étudier cette étape de la réplication, nous nous concentrons sur deux éléments, nécessaires et suffisants : les protéines et l'ARN pré-génomique. Ce dernier, matrice pour la rétro-transcription en brin d'ADN négatif, mesure 3,5 kb, ce qui est légèrement plus grand que le génome des virions matures. Un point important est qu'il contient une région terminale, redondante, d'une centaine de nucléotides [125][32].

### 3.1.1 Les mutations effectuées sur le ARN-pg

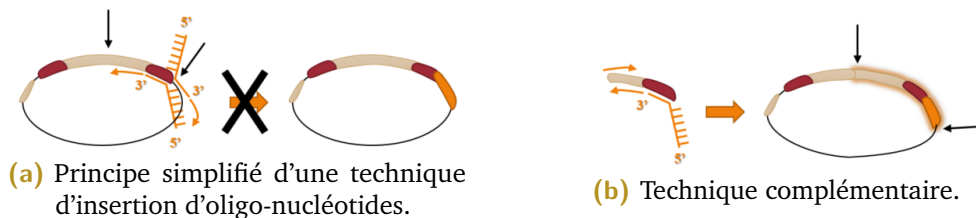
Afin de pouvoir greffer l'ARN-pg sur la surface de la cellule d'observation, il faut ajouter des séquences de nucléotides. Celles-ci sont présentées sur la figure 6.28. La séquence permettant l'hybridation avec un ARN fluorescent est montrée en



**Figure 6.28:** Illustration schématique des mutations par insertion réalisées sur l'ARN-pg.

rouge. La séquence, ajoutée en bleu, permet l'hybridation avec un ARN ancré sur la cellule d'observation via une biotine (voir section 1.1.3). Entre ces deux ajouts, il y a un connecteur, présenté en noir.

Plusieurs techniques existent pour réaliser cette insertion, effectuée par Sonia Fieulaine de l'I2BC. L'une d'elles est résumée de manière schématique sur la figure 6.29a). Sur celle-ci, à gauche, le plasmide portant l'information nécessaire à la transcription de l'ARN-pg est symbolisé par un cercle noir et les régions codantes pour celui-ci sont mises en avant en beige et rouge. Les séquences de nucléotides à insérer, sous forme de simple brin, sont symbolisées par les traits oranges. Celles-ci sont fixées, sur le plasmide, via des sites de reconnaissance. Des polymérases viennent ensuite synthétiser la nouvelle chaîne dans le sens 5' à 3'.



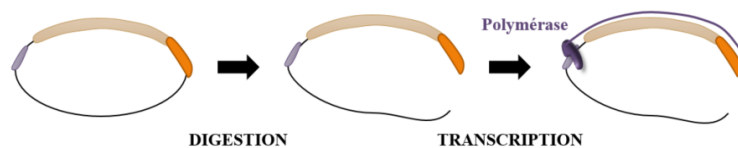
**Figure 6.29:** Illustrations schématiques du processus d'insertion des mutations dans l'ARN pré-génomique.

Dans notre cas, ce plasmide possède une région redondante, symbolisée en rouge sur la figure. De ce fait, l'une des séquences de nucléotides pouvait s'insérer à deux endroits dans le plasmide. Pour pallier ce problème, le plasmide a été découpé en deux endroits, symbolisés par les flèches noires sur cette même figure. L'un des deux fragments obtenus est visible sur la figure 6.29b, à gauche. C'est sur celui-ci, uniquement, que la séquence de nucléotides a été insérée. Ce fragment a ensuite été réintroduit pour former le plasmide portant les mutations nécessaires. Le plasmide final est schématisé à droite de la figure et l'insertion est symbolisée par un cylindre orange.

### 3.1.2 Production de l'ARN-pg avec insertion

A l'issue de ces modifications, nous avons produit l'ARN-pg avec insertion en quantité et pureté suffisantes pour pouvoir réaliser des expériences d'assemblage par TIRFM. Les protocoles de production nous ont été fournis par Lauriane Lecoq du MMSB de Lyon.

Succinctement, la production se passe en plusieurs étapes, schématisées sur la figure 6.30. La première étape est une phase de linéarisation du plasmide codant pour



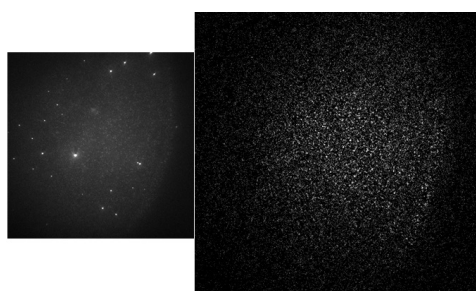
**Figure 6.30:** Illustrations schématique du processus de production de l'ARN pré-génomique

l'ARN pré-génomique modifié. Cette étape passe par une digestion du plasmide à l'aide d'une enzyme qui va couper le plasmide au niveau d'un site de restriction spécifique. Cette étape est nécessaire car le plasmide ne possède pas de site signalant aux enzymes d'arrêter la transcription. La deuxième étape sert à purifier l'ADN linéarisé. Cette étape est très importante car la transcription du plasmide nécessite un haut niveau de pureté. Il faut donc séparer les molécules d'ADN, des enzymes présentes en solution du fait de l'étape de linéarisation. L'étape suivante est celle de transcription. Lors de cette étape, les polymérases synthétisent l'ARN-pg à partir d'un promoteur spécifique correspondant, symbolisé en violet sur le schéma.

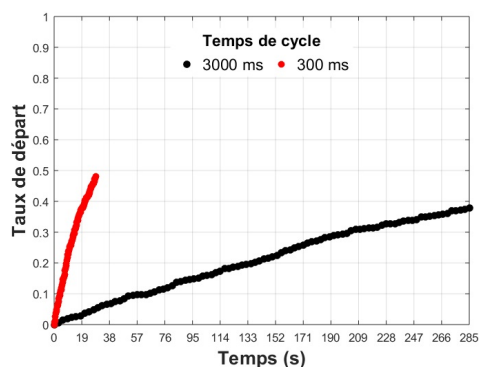
## 3.2 Résultats préliminaires en TIRFM

Dans cette section, je présenterai les résultats préliminaires obtenus lors de nos expériences de TIRFM, avec l'ARN pré-génomique uniquement. La préparation de la surface de la cellule et l'hybridation de l'ARN-pg nécessitent une demi-journée et 15-20 minutes, respectivement.

Sur l'image 6.31a sont montrées des photographies de deux des zones des cellules que nous avons observées. Sur celles-ci, les deux cellules d'observation ont été préparées avec des solutions d'ARN-pg différentes. Les concentrations en ARN sont de 1,75 et 1,15  $\text{ng}\cdot\mu\text{L}^{-1}$ , à gauche et à droite, respectivement. Sur l'image de gauche, nous pouvons remarquer, qu'au sein de la préparation, des objets/agrégats semblent fixer particulièrement le marqueur fluorescent. En effet, contrairement à l'image de droite qui montre un tapis très homogène, il y a des taches sensiblement plus lumineuses. Pour les deux échantillons d'ARN-pg utilisés, le rapport



(a) Photos des ARN-pg fluorescents. Les concentrations en ARN sont de 1,75 et 1,15  $\text{ng}\cdot\mu\text{L}^{-1}$ , à gauche et à droite, respectivement.



(b) Tests de stabilité des ARN-pg fluorescents.

**Figure 6.31:** Résultats préliminaires en TIRFM

$A_{260}/A_{230}$ <sup>1</sup> et le rapport  $A_{260}/A_{280}$ , indicateur de pureté, sont équivalents à 2,2. Ce rapport est correct et ne semble pas être à l'origine de cette différence. De plus, les concentrations de stockage des ARN-pg étaient de 175 et 575  $\text{ng}\cdot\mu\text{L}^{-1}$ , à gauche et à droite, respectivement. C'est donc la solution la moins concentrée qui présente des agrégats. Cette différence ne semble donc pas non plus imputable à ce paramètre. Ces agrégats étaient présents sur toutes les cellules d'observation préparées avec le même échantillon d'ARN-pg. Or, ceux-ci rendant la cellule d'observation inutilisable, un test similaire devra donc être réalisé pour valider l'utilisation de chaque solution d'ARN-pg.

De plus, malgré la présence de la biotine, l'adhérence des ARNs à la surface de la cellule d'observation est temporaire. Celle-ci, de même que le photo-blanchiment des marqueurs au cours du temps, doit être évaluée dans les conditions réelles de l'expérience pour un marqueur fluorescent donné. En effet, ces valeurs dépendent du fluorophore, des conditions d'excitation et de l'environnement du dit fluorophore. Lors de ces tests, le temps d'exposition était de 100 ms, la puissance de 20 mW et le gain d'acquisition de 15 sur la caméra EMCCD. Le temps de cycle, défini comme la somme des temps d'exposition et de pause, est variable. Les résultats sont présentés sur la figure 6.31b. Le graphique montre l'évolution du taux de disparition en fonction du temps, i.e. l'ensemble des décrochages et des marqueurs photo-blanchis. Sur le graphique, la courbe rouge correspond à un cycle dit rapide, i.e. 300 ms (100 ms d'exposition et 200 ms de pause), tandis que la courbe noire correspond à un cycle 10 fois plus lent (3000 ms). Ces tests ont permis d'estimer

1. Plusieurs contaminants absorbant à 230 nm proviennent de l'échantillon ou de la purification. Par ex. le phénol, l'urée, l'EDTA, etc. Leur présence est détectable grâce à ce ratio.

que 650 à 700 particules sont exploitables par film. Dans le premier cas, 50% des marqueurs ont disparu en 30 secondes environ, tandis que, dans le second, 60 % des marqueurs sont encore présents au bout d'un peu moins de 5 minutes.

Si le cycle rapide semble pouvoir permettre une observation détaillée des premières secondes de l'assemblage, l'évolution rapide du taux de départ, i.e. décrochement et photo-blanchiment, ne permet pas d'envisager un suivi sur plusieurs dizaines de minutes. A contrario, le cycle lent ne permet pas une observation précise des premiers instants. Cependant, nous pouvons noter que, si les deux cycles ont un temps total d'exposition équivalent, le taux de départ de la cinétique lente est réduit de 10% du fait de l'augmentation du temps de pause. De ce fait, celui-ci permet une observation plus longue, voire de l'entièreté, de la cinétique. En conclusion, il faudra composer avec, au moins, ces deux types de cycles pour gérer le compromis entre précision et endurance.

Dans ce chapitre, nous avons testé différentes stratégies pour obtenir des dimères de protéines fluorescents et préparé le génome en vue de l'étude, par TIRFM, de l'auto-assemblage des capsides du VHB.

Tout d'abord, nous avons essayé de répondre à cette problématique via un mutant par insertion d'une protéine fluorescente, le mutant GFP-CP183. Nous avons pu observer l'impact de la eGFP sur les premières étapes de production et modifier en conséquence les différents protocoles. En effet, le plasmide et la protéine induisent un stress pour les bactéries. De plus, la formation de la structure finale du chromophore nécessite une étape d'oxydation d'une durée de plusieurs heures. En conséquence, les températures et temps de production ont dû être modifiés et des contrôles réguliers ont dû être ajoutés. Nous avons également pu constater l'impact du protocole de désassemblage sur la fluorescence des dimères de GFP-CP183. En effet, le chlorure de guanidinium dénature la structure tridimensionnelle du chromophore qui perd sa capacité à émettre de la fluorescence. Après divers tests, nous avons pu obtenir des dimères fluorescents du mutant. Cependant, le signal de fluorescence de ces derniers n'était pas suffisamment stable pour être utilisé en l'état. Plusieurs possibilités d'amélioration existent. En effet, la photostabilité des protéines pourrait être due à la composition du milieu nutritif utilisé pendant l'étape d'expression des protéines. L'expression pourrait être réalisée dans un milieu nutritif différent, tel que le milieu minium M9 amélioré avec oligo-éléments (cadmium, cobalt, etc) et vitamines (sans vitamine B2, i.e. riboflavine). Une autre possibilité serait d'utiliser une autre protéine fluorescente comme la Mcherry. En effet, les protéines dérivées de la DsRed sont plus stables devant les dénaturants et le pH que la eGFP.

Toujours dans l'optique d'obtenir des dimères de protéines fluorescents, nous avons également investigué la possibilité de produire des dimères de protéines du VHB covalamment liées avec des marqueurs chimiques fluorescents. Dans ce cas, les protéines marquées dissociées montraient des problèmes d'agrégaions du fait du DMSO, des protéines ou des marqueurs. En conséquence, dans les meilleures conditions de notre étude, nous n'avons pas pu obtenir plus que 1,7 % de protéines



marquées. De ce fait, nous avons également testé notre protocole de marquage sur les protéines d'un autre virus, le CCMV. Dans ce cas, nous avons pu marquer 4 % des dimères de protéines. Ainsi, le marquage et le désassemblage se sont révélés plus efficaces avec le CCMV qu'avec le VHB. Plusieurs possibilités d'amélioration sont tout de même encore possibles. En effet, en réduisant le débit, lors de la chromatographie ou en utilisant une colonne avec un volume plus important (par ex. la S200 16/600 de 120 mL), les dimères marqués et non-marqués pourraient être séparés plus efficacement. De cette façon, le pourcentage final de dimères marqués, en solution, pourrait être accru.

Enfin, nous avons muté et produit de l'ARN pré-génomique avec insertion, i.e. des séquences supplémentaires permettant l'hybridation avec un fluorophore et un ancrage sur la cellule d'observation. Si les résultats préliminaires ont été concluants, puisque nous avons estimé qu'entre 650 et 700 particules sont exploitables par film, ceux-ci ont tout de même mis en évidence un taux de photo-blanchiment et de départ non-négligeables. En conséquence, il faudra composer avec, au moins, deux types de cycles différents pour pouvoir observer avec précision l'ensemble des cinétiques d'auto-assemblage.

Ces résultats sont prometteurs et permettent d'envisager une étude, par TIRFM, de l'auto-assemblage des capsides du VHB à court ou moyen terme.

# VII

---

Conclusion et perspectives

Les travaux réalisés pendant ces trois années de thèse au LPS, à l'I2BC, mais également aux synchrotrons ESRF et SOLEIL ainsi qu'à la source de neutrons LLB, et présentés dans ce manuscrit, ont permis de mieux comprendre les chemins cinétiques d'assemblage et de désassemblage des capsides de virus, vides ou pleines. Dans un premier temps, la cinétique d'auto-assemblage des capsides pleines du CCMV a été étudiée par diffusion de rayons X aux petits angles résolue en temps (TR-SAXS). Ensuite, les cinétiques d'auto-assemblage et de désassemblage des capsides vides du VHB ont été examinées par la même technique. Enfin, des protocoles de biochimie ont été développés en vue d'une étude, par microscopie de fluorescence par réflexion interne totale (TIRFM), de l'auto-assemblage des capsides pleines du VHB.

Dans une première partie du manuscrit, quelques applications possibles de ces nanostructures auto-assemblées ont été présentées. De plus, certaines parties du manuscrit comportent un état de l'art qui leur est propre. Ainsi, l'ensemble met en évidence les nombreuses études réalisées pour comprendre ces phénomènes. Cependant, celles-ci ont été principalement réalisées à l'équilibre ou avec des résolutions spatiales et temporelles insuffisantes. Le TR-SAXS, technique qui permet de suivre des processus impliquant des espèces de taille nanométrique avec une résolution temporelle de l'ordre de quelques millisecondes, a été couplée à la décomposition en valeurs singulières (SVD), une méthode d'analyse indépendante de modèles, afin de répondre à ces problématiques.

Ainsi, dans le cas du CCMV, les expériences réalisées par TR-SAXS ont permis de révéler les chemins cinétiques conduisant à la formation des virions, *in vitro*. Ces expériences ont également permis d'estimer les énergies de liaison des dimères de protéine avec le génome et mis en évidence la présence d'une barrière d'énergie pouvant être modulée. Ensuite, les expériences réalisées par TR-SAXS sur les capsides vides du VHB ont également permis d'identifier un chemin cinétique de désassemblage probable ainsi que la nature de l'espèce intermédiaire. Dans le cas de l'assemblage, les expériences ont permis de révéler un chemin cinétique en trois phases. De plus, certaines expériences ont également montré que la dernière phase pouvait être facilement inhibée. Dans la dernière partie du manuscrit, différentes stratégies pour obtenir des dimères de protéines fluorescents ont été mises en oeuvre afin d'étudier l'assemblage des capsides pleines du VHB et du CCMV, en molécule unique. Si les mutants de la protéine de la capsidite du VHB se sont révélés

difficiles à produire, la protéine de la capsidie du CCMV s'est révélée plus maniable, permettant d'envisager une étude par TIRFM.

Cependant, certaines questions restent en suspens. Dans le cas du VHB, notamment, l'assemblage de la capsidie pleine n'a pas été résolu. Si l'étude par TIRFM de l'auto-assemblage du VHB n'est pas encore mature, une étude par TR-SAXS permettrait d'identifier des chemins ou des mécanismes cinétiques probables de l'assemblage. En effet, comme nous avons pu le voir avec la partie consacrée à l'assemblage de la capsidie pleine du CCMV, cette technique est parfaitement adaptée à l'étude de systèmes complexes. De plus, l'utilisation de la protéine native, non-phosphorylée lors de la production en *E. Coli.*, couplée à l'utilisation du mutant CP183-EEE, qui imite les propriétés physico-chimiques de la protéine au moment de l'emballage, permettrait de mieux comprendre les interactions entre les protéines et le génome. Dans le cas de l'assemblage de la capsidie vide du VHB, nous n'avons pas pu confirmer la présence d'intermédiaire précoce ou tardif, du fait d'un mauvais rapport signal/bruit. Des expériences à plus fortes concentrations en protéines pourraient répondre à ces questions. Dans le cas du désassemblage de la capsidie vide du VHB, nous avons pu déterminer la nature de l'espèce intermédiaire, observée auparavant. Cependant, celle-ci semble être générée par la présence de l'agent chaotrope utilisé pour accélérer le désassemblage. Une résolution temporelle de quelques millisecondes n'étant pas requise, puisque le désassemblage s'effectue sur plusieurs jours, à pH9 en l'absence d'urée, une étude par diffusion statique de la lumière utilisant le système du goniomètre pourrait être envisagée pour confirmer ou infirmer le modèle de dépolymérisation proposé. Dans le cas du CCMV, l'assemblage avec le génome a pu être réalisé par un saut de pH mais également en augmentant la concentration en protéines et en génome. Cependant, l'ensemble de ces études a été réalisé en milieu non-encombré, ce qui n'est pas le cas *in vivo*. Ainsi, des études en présence de PEG, qui mimerait l'encombrement cellulaire, pourraient également permettre de mieux comprendre la dynamique d'assemblage des virus. Dans ce cas, le remplacement du mélangeur classique, situé dans la tête d'observation du stopped-flow, au profit du mélangeur HDS (High Density Solution) devra être envisagé afin d'éviter des problèmes similaires à ceux observés lors des expériences de désassemblage de la capsidie vide du VHB.

De manière plus générale, les expériences présentées ci-avant démontrent la puissance de la diffusion de rayons X aux petits angles résolue en temps pour l'étude

de systèmes complexes hors équilibre tels que l'assemblage ou le désassemblage des particules virales. En effet, cette technique a permis, de façon claire et non-biaisée, de répondre à des questions restées jusqu'ici sans réponse ou de remettre en cause certains modèles établis précédemment à partir d'observations indirectes. Le cas le plus frappant est le modèle d'assemblage de CCMV, qui concluait à des NPC sur-stoechiométriques évoluant en capsides par relargage des dimères surnuméraires : nos données établissent sans ambiguïté que les NPC sont sous-stoechiométriques et sont converties en capsides par capture des dimères manquants et structuration synchrone.

# Bibliographie

- [1]N. Philippe, M. Legendre, G. Doutré et al. „Pandoraviruses : Amoeba Viruses with Genomes Up to 2.5 Mb Reaching That of Parasitic Eukaryotes“. In : *Science* 341.6143 (2013), p. 281-286. DOI : [10.1126/science.1239181](https://doi.org/10.1126/science.1239181) (cf. p. 6).
- [2]Patrick Forterre. „To be or not to be alive : How recent discoveries challenge the traditional definitions of viruses and life“. In : *Studies in History and Philosophy of Science Part C : Studies in History and Philosophy of Biological and Biomedical Sciences* 59 (2016), p. 100-108. DOI : [10.1016/j.shpsc.2016.02.013](https://doi.org/10.1016/j.shpsc.2016.02.013) (cf. p. 6).
- [3]Matthieu Galtier, Luisa De Sordi, Adeline Sivignon et al. „Bacteriophages targeting adherent invasive Escherichia coli strains as a promising new treatment for Crohn's disease“. In : *Journal of Crohn's and Colitis* (2017), jjw224. DOI : [10.1093/ecco-jcc/jjw224](https://doi.org/10.1093/ecco-jcc/jjw224) (cf. p. 6).
- [4]Yinon M. Bar-On, Rob Phillips et Ron Milo. „The biomass distribution on Earth“. In : *Proceedings of the National Academy of Sciences* 115.25 (2018), p. 6506-6511. DOI : [10.1073/pnas.1711842115](https://doi.org/10.1073/pnas.1711842115) (cf. p. 10).
- [5]Mark Woolhouse, Fiona Scott, Zoe Hudson, Richard Howey et Margo Chase-Topping. „Human viruses : discovery and emergence“. In : *Philosophical Transactions of the Royal Society B : Biological Sciences* 367.1604 (2012), p. 2864-2871. DOI : [10.1098/rstb.2011.0354](https://doi.org/10.1098/rstb.2011.0354) (cf. p. 10).
- [6]D. Baltimore. „Expression of animal virus genomes.“ In : *Bacteriological reviews* 35 (3 sept. 1971), p. 235-241 (cf. p. 11).
- [7]B. V. Venkataram Prasad et Michael F. Schmid. „Principles of Virus Structural Organization“. In : *Viral Molecular Machines*. Springer US, 2011, p. 17-47. DOI : [10.1007/978-1-4614-0980-9\\_3](https://doi.org/10.1007/978-1-4614-0980-9_3) (cf. p. 12).
- [8]T. Tubiana. „Dynamique d'assemblage de la capsid des norovirus“. Thèse de doct. Innovation thérapeutique, 2017 (cf. p. 12).
- [9]Patrick Forterre et David Prangishvili. „The Great Billion-year War between Ribosome- and Capsid-encoding Organisms (Cells and Viruses) as the Major Source of Evolutionary Novelties“. In : *Annals of the New York Academy of Sciences* 1178.1 (2009), p. 65-77. DOI : [10.1111/j.1749-6632.2009.04993.x](https://doi.org/10.1111/j.1749-6632.2009.04993.x) (cf. p. 13).

- [10]D. L. D. Caspar et A. Klug. „Physical Principles in the Construction of Regular Viruses“. In : *Cold Spring Harbor Symposia on Quantitative Biology* 27.0 (1962), p. 1-24. DOI : [10.1101/sqb.1962.027.001.005](https://doi.org/10.1101/sqb.1962.027.001.005) (cf. p. 13, 14).
- [11]Jeffrey A Speir, Sanjeev Munshi, Guoji Wang, Timothy S Baker et John E Johnson. „Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy“. In : *Structure* 3.1 (1995), p. 63-78. DOI : [10.1016/s0969-2126\(01\)00135-6](https://doi.org/10.1016/s0969-2126(01)00135-6) (cf. p. 14, 16-18, 79, 102).
- [12]Owen Pornillos, Barbie K. Ganser-Pornillos et Mark Yeager. „Atomic-level modelling of the HIV capsid“. In : *Nature* 469.7330 (2011), p. 424-427. DOI : [10.1038/nature09640](https://doi.org/10.1038/nature09640) (cf. p. 14).
- [13]J.B. Bancroft, G.J. Hills et Rou Markham. „A study of the self-assembly process in a small spherical virus formation of organized structures from protein subunits in vitro“. In : *Virology* 31.2 (1967), p. 354-379. DOI : [10.1016/0042-6822\(67\)90180-8](https://doi.org/10.1016/0042-6822(67)90180-8) (cf. p. 16).
- [14]J B Bancroft et E Hiebert. „Formation of an infectious nucleoprotein from protein and nucleic acid isolated from a small spherical virus.“ In : *Virology* 32 (2 juin 1967), p. 354-356. DOI : [10.1016/0042-6822\(67\)90284-x](https://doi.org/10.1016/0042-6822(67)90284-x) (cf. p. 16).
- [15]Adriana Lucía-Sanz et Susanna Manrubia. „Multipartite viruses : adaptive trick or evolutionary treat?“ In : *npj Systems Biology and Applications* 3.1 (2017). DOI : [10.1038/s41540-017-0035-y](https://doi.org/10.1038/s41540-017-0035-y) (cf. p. 16).
- [16]J. Chidlow et J.H. Tremaine. „Limited hydrolysis of cowpea chlorotic mottle virus by trypsin and chymotrypsin“. In : *Virology* 43.1 (1971), p. 267-278. DOI : [10.1016/0042-6822\(71\)90244-3](https://doi.org/10.1016/0042-6822(71)90244-3) (cf. p. 17).
- [17]G. Vriend, B.J.M. Verduin et M.A. Hemminga. „Role of the N-terminal part of the coat protein in the assembly of cowpea chlorotic mottle virus“. In : *Journal of Molecular Biology* 191.3 (1986), p. 453-460. DOI : [10.1016/0022-2836\(86\)90140-3](https://doi.org/10.1016/0022-2836(86)90140-3) (cf. p. 17).
- [18]Guillaume Tresset, Jingzhi Chen, Maelenn Chevreuil et al. „Two-Dimensional Phase Transition of Viral Capsid Gives Insights into Subunit Interactions“. In : *Physical Review Applied* 7.1 (2017). DOI : [10.1103/physrevapplied.7.014005](https://doi.org/10.1103/physrevapplied.7.014005) (cf. p. 18, 75).
- [19]L. Lavelle, M. Gingery, M. Phillips et al. „Phase Diagram of Self-assembled Viral Capsid Protein Polymorphs†“. In : *The Journal of Physical Chemistry B* 113.12 (2009), p. 3813-3819. DOI : [10.1021/jp8079765](https://doi.org/10.1021/jp8079765) (cf. p. 18).
- [20]Jean-Philippe Michel, Mari Gingery et Laurence Lavelle. „Efficient purification of bromoviruses by ultrafiltration“. In : *Journal of Virological Methods* 122.2 (2004), p. 195-198. DOI : [10.1016/j.jviomet.2004.09.005](https://doi.org/10.1016/j.jviomet.2004.09.005) (cf. p. 22).

- [21] B.J.M. Verduin. „The preparation of CCMV-protein in connection with its association into a spherical particle“. In : *FEBS Letters* 45.1-2 (1974), p. 50-54. DOI : [10.1016/0014-5793\(74\)80808-2](https://doi.org/10.1016/0014-5793(74)80808-2) (cf. p. 24).
- [22] A. Lurman. „Eine icterus epidemic.“ In : *Berl Klin Wochenschr* (1885) (cf. p. 25).
- [23] B. S. Blumberg. „Polymorphisms of the serum proteins and the development of iso-precipitins in transfused patients.“ In : *Bulletin of the New York Academy of Medicine* 40 (mai 1964), p. 377-386 (cf. p. 25).
- [24] D.S. Dane, C.H. Cameron et Moya Briggs. „VIRUS-LIKE PARTICLES IN SERUM OF PATIENTS WITH AUSTRALIA-ANTIGEN-ASSOCIATED HEPATITIS“. In : *The Lancet* 295.7649 (1970), p. 695-698. DOI : [10.1016/s0140-6736\(70\)90926-8](https://doi.org/10.1016/s0140-6736(70)90926-8) (cf. p. 25).
- [25] Francis Galibert, Elisabeth Mandart, Françoise Fitoussi, Pierre Tiollais et Patrick Charnay. „Nucleotide sequence of the hepatitis B virus genome (subtype ayw) cloned in *E. coli*“. In : *Nature* 281.5733 (1979), p. 646-650. DOI : [10.1038/281646a0](https://doi.org/10.1038/281646a0) (cf. p. 25).
- [26] Sushmita Singh et Adam Zlotnick. „Observed Hysteresis of Virus Capsid Disassembly Is Implicit in Kinetic Models of Assembly“. In : *Journal of Biological Chemistry* 278.20 (2003), p. 18249-18255. DOI : [10.1074/jbc.m211408200](https://doi.org/10.1074/jbc.m211408200) (cf. p. 25, 114, 154).
- [27] Jin-Seung Park, Moon Kyu Cho, Eun Jung Lee et al. „A highly sensitive and selective diagnostic assay based on virus nanoparticles“. In : *Nature Nanotechnology* 4.4 (2009), p. 259-264. DOI : [10.1038/nnano.2009.38](https://doi.org/10.1038/nnano.2009.38) (cf. p. 25).
- [28] Lina Yoo, Jin-Seung Park, Koo Chul Kwon et al. „Fluorescent viral nanoparticles with stable in vitro and in vivo activity“. In : *Biomaterials* 33.26 (2012), p. 6194-6200. DOI : [10.1016/j.biomaterials.2012.05.028](https://doi.org/10.1016/j.biomaterials.2012.05.028) (cf. p. 25).
- [29] D. Lavanchy. „Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures“. In : *Journal of Viral Hepatitis* 11.2 (2004), p. 97-107. DOI : [10.1046/j.1365-2893.2003.00487.x](https://doi.org/10.1046/j.1365-2893.2003.00487.x) (cf. p. 25).
- [30] Aparna Schweitzer, Johannes Horn, Rafael T Mikolajczyk, Gérard Krause et Jördis J Ott. „Estimations of worldwide prevalence of chronic hepatitis B virus infection : a systematic review of data published between 1965 and 2013“. In : *The Lancet* 386.10003 (2015), p. 1546-1555. DOI : [10.1016/s0140-6736\(15\)61412-x](https://doi.org/10.1016/s0140-6736(15)61412-x) (cf. p. 26).
- [31] W. W. Bond, M. S. Favero, N. J. Petersen et al. „Survival of hepatitis b virus after drying and storage for one week“. In : *The Lancet* 317.8219 (1981), p. 550-551. DOI : [10.1016/s0140-6736\(81\)92877-4](https://doi.org/10.1016/s0140-6736(81)92877-4) (cf. p. 26).
- [32] Juergen Beck. „Hepatitis B virus replication“. In : *World Journal of Gastroenterology* 13.1 (2007), p. 48. DOI : [10.3748/wjg.v13.i1.48](https://doi.org/10.3748/wjg.v13.i1.48) (cf. p. 26, 196).
- [33] Chiaho Shih, Shu-Fan Chou, Ching-Chun Yang et al. „Control and Eradication Strategies of Hepatitis B Virus“. In : *Trends in Microbiology* 24.9 (2016), p. 739-749. DOI : [10.1016/j.tim.2016.05.006](https://doi.org/10.1016/j.tim.2016.05.006) (cf. p. 27).



- [34]Mark Pasek, Tadaatsu Goto, Walter Gilbert et al. „Hepatitis B virus genes and their expression in *E. coli*“. In : *Nature* 282.5739 (1979), p. 575-579. DOI : [10.1038/282575a0](https://doi.org/10.1038/282575a0) (cf. p. 27).
- [35]F. Birnbaum et M. Nassal. „Hepatitis B virus nucleocapsid assembly : primary structure requirements in the core protein.“ In : *Journal of virology* 64 (7 juil. 1990), p. 3319-3330 (cf. p. 28, 29, 108).
- [36]S.A Wynne, R.A Crowther et A.G.W Leslie. „The Crystal Structure of the Human Hepatitis B Virus Capsid“. In : *Molecular Cell* 3.6 (1999), p. 771-780. DOI : [10.1016/s1097-2765\(01\)80009-5](https://doi.org/10.1016/s1097-2765(01)80009-5) (cf. p. 28, 140).
- [37]M Nassal. „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.“ In : *Journal of virology* 66 (7 juil. 1992), p. 4107-4116 (cf. p. 28).
- [38]David H. Nguyen, Laurie Ludgate et Jianming Hu. „Hepatitis B virus–cell interactions and pathogenesis“. In : *Journal of Cellular Physiology* 216.2 (2008), p. 289-294. DOI : [10.1002/jcp.21416](https://doi.org/10.1002/jcp.21416) (cf. p. 29).
- [39]Balasubramanian Venkatakrishnan et Adam Zlotnick. „The Structural Biology of Hepatitis B Virus : Form and Function“. In : *Annual Review of Virology* 3.1 (2016), p. 429-451. DOI : [10.1146/annurev-virology-110615-042238](https://doi.org/10.1146/annurev-virology-110615-042238) (cf. p. 29).
- [40]T. Friedmann et R. Roblin. „Gene Therapy for Human Genetic Disease ?“ In : *Science* 175.4025 (1972), p. 949-955. DOI : [10.1126/science.175.4025.949](https://doi.org/10.1126/science.175.4025.949) (cf. p. 30).
- [41]Mohube Betty Mowa, Carol Crowther, Abdullah Ely et Patrick Arbuthnot. „Inhibition of Hepatitis B Virus Replication by Helper Dependent Adenoviral Vectors Expressing Artificial Anti-HBV Pri-miRs from a Liver-Specific Promoter“. In : *BioMed Research International* 2014 (2014), p. 1-12. DOI : [10.1155/2014/718743](https://doi.org/10.1155/2014/718743) (cf. p. 31).
- [42]Odisse Azizgolshani, Rees F. Garmann, Ruben Cadena-Nava, Charles M. Knobler et William M. Gelbart. „Reconstituted plant viral capsids can release genes to mammalian cells“. In : *Virology* 441.1 (2013), p. 12-17. DOI : [10.1016/j.virol.2013.03.001](https://doi.org/10.1016/j.virol.2013.03.001) (cf. p. 31).
- [43]P. Suci, S. Kang, R. Gmur, T. Douglas et M. Young. „Targeted Delivery of a Photosensitizer to *Aggregatibacter actinomycetemcomitans* Biofilm“. In : *Antimicrobial Agents and Chemotherapy* 54.6 (2010), p. 2489-2496. DOI : [10.1128/aac.00059-10](https://doi.org/10.1128/aac.00059-10) (cf. p. 31).
- [44]Daniel Luque, Andrés de la Escosura, Joost Snijder et al. „Self-assembly and characterization of small and monodisperse dye nanospheres in a protein cage“. In : *Chem. Sci.* 5.2 (2014), p. 575-581. DOI : [10.1039/c3sc52276h](https://doi.org/10.1039/c3sc52276h) (cf. p. 31).
- [45]Noëlle Carette, Hans Engelkamp, Eric Akpa et al. „A virus-based biocatalyst“. In : *Nature Nanotechnology* 2.4 (2007), p. 226-229. DOI : [10.1038/nnano.2007.76](https://doi.org/10.1038/nnano.2007.76) (cf. p. 31).

- [46]Marta Comellas-Aragonès, Hans Engelkamp, Victor I. Claessen et al. „A virus-based single-enzyme nanoreactor“. In : *Nature Nanotechnology* 2.10 (2007), p. 635-639. DOI : [10.1038/nnano.2007.299](https://doi.org/10.1038/nnano.2007.299) (cf. p. 32).
- [47]Stella E. Aniagyei, Chelsea J. Kennedy, Barry Stein et al. „Synergistic Effects of Mutations and Nanoparticle Templating in the Self-Assembly of Cowpea Chlorotic Mottle Virus Capsids“. In : *Nano Letters* 9.1 (2009), p. 393-398. DOI : [10.1021/nl8032476](https://doi.org/10.1021/nl8032476) (cf. p. 32).
- [48]Andrew Burke, Maelenn Chevreuil, Alisier Paris et al. „Nanoparticle-Templated Self-Assembly of Viral Capsids Probed by Time-Resolved Absorbance Spectroscopy and X-Ray Scattering“. In : *Physical Review Applied* 10.5 (2018). DOI : [10.1103/physrevapplied.10.054065](https://doi.org/10.1103/physrevapplied.10.054065) (cf. p. 32).
- [49]Erik C. Dreaden, Alaaldin M. Alkilany, Xiaohua Huang, Catherine J. Murphy et Mostafa A. El-Sayed. „The golden age : gold nanoparticles for biomedicine“. In : *Chem. Soc. Rev.* 41.7 (2012), p. 2740-2779. DOI : [10.1039/c1cs15237h](https://doi.org/10.1039/c1cs15237h) (cf. p. 32).
- [50]K. T. Nam. „Virus-Enabled Synthesis and Assembly of Nanowires for Lithium Ion Battery Electrodes“. In : *Science* 312.5775 (2006), p. 885-888. DOI : [10.1126/science.1122716](https://doi.org/10.1126/science.1122716) (cf. p. 32).
- [51]Trevor Douglas et Mark Young. „Host–guest encapsulation of materials by assembled virus protein cages“. In : *Nature* 393.6681 (1998), p. 152-155. DOI : [10.1038/30211](https://doi.org/10.1038/30211) (cf. p. 32).
- [52]Michel H. J. Koch, Patrice Vachette et Dmitri I. Svergun. „Small-angle scattering : a view on the properties, structures and structural changes of biological macromolecules in solution“. In : *Quarterly Reviews of Biophysics* 36.2 (2003), p. 147-227. DOI : [10.1017/s0033583503003871](https://doi.org/10.1017/s0033583503003871) (cf. p. 33).
- [53]L. A. Feigin, D. I. Svergun et George W. Taylor. „Principles of the Theory of X-Ray and Neutron Scattering“. In : *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*. Springer US, 1987, p. 3-24. DOI : [10.1007/978-1-4757-6624-0\\_1](https://doi.org/10.1007/978-1-4757-6624-0_1) (cf. p. 34).
- [54]Dmitri I Svergun et Michel H J Koch. „Small-angle scattering studies of biological macromolecules in solution“. In : *Reports on Progress in Physics* 66.10 (2003), p. 1735-1782. DOI : [10.1088/0034-4885/66/10/r05](https://doi.org/10.1088/0034-4885/66/10/r05) (cf. p. 34, 42, 51, 85).
- [55]Jean Daillant et Alain Gibaud, éd. *X-ray and Neutron Reflectivity*. Springer Berlin Heidelberg, 2009. DOI : [10.1007/978-3-540-88588-7](https://doi.org/10.1007/978-3-540-88588-7) (cf. p. 34).
- [56]Cécile A. Dreiss, Kevin S. Jack et Andrew P. Parker. „On the absolute calibration of bench-top small-angle X-ray scattering instruments : a comparison of different standard methods“. In : *Journal of Applied Crystallography* 39.1 (2006), p. 32-38. DOI : [10.1107/s0021889805033091](https://doi.org/10.1107/s0021889805033091) (cf. p. 35).
- [57]Ian W. Hamley Hamley. *Introduction to Soft Matter Revised*. John Wiley & Sons, 11 oct. 2007. 342 p. (cf. p. 38).

- [58] Ü.Ö. Köylü, G.M. Faeth, T.L. Farias et M.G. Carvalho. „Fractal and projected structure properties of soot aggregates“. In : *Combustion and Flame* 100.4 (1995), p. 621-633. DOI : [10.1016/0010-2180\(94\)00147-k](https://doi.org/10.1016/0010-2180(94)00147-k) (cf. p. 38).
- [59] André Guinier, Gérard Fournet, C. B. Walker et George H. Vineyard. „Small-Angle Scattering of X-Rays“. In : *Physics Today* 9.8 (1956), p. 38-39. DOI : [10.1063/1.3060069](https://doi.org/10.1063/1.3060069) (cf. p. 42).
- [60] Maria Monica Castellanos, Arnold McAuley et Joseph E. Curtis. „Investigating Structure and Dynamics of Proteins in Amorphous Phases Using Neutron Scattering“. In : *Computational and Structural Biotechnology Journal* 15 (2017), p. 117-130. DOI : [10.1016/j.csbj.2016.12.004](https://doi.org/10.1016/j.csbj.2016.12.004) (cf. p. 43, 44).
- [61] J.A. Langer, D.M. Engelman et P.B. Moore. „Neutron-scattering studies of the ribosome of Escherichia coli : A provisional map of the locations of proteins S3, S4, S5, S7, S8 and S9 in the 30 S subunit“. In : *Journal of Molecular Biology* 119.4 (1978), p. 463-485. DOI : [10.1016/0022-2836\(78\)90197-3](https://doi.org/10.1016/0022-2836(78)90197-3) (cf. p. 44).
- [62] Anne T. Tuukkanen, Alessandro Spilotros et Dmitri I. Svergun. „Progress in small-angle scattering from biological solutions at high-brilliance synchrotrons“. In : *IUCrJ* 4.5 (2017), p. 518-528. DOI : [10.1107/s2052252517008740](https://doi.org/10.1107/s2052252517008740) (cf. p. 45, 56).
- [63] M. Levantino, A. Spilotros, M. Cammarata et al. „The Monod-Wyman-Changeux allosteric model accounts for the quaternary transition dynamics in wild type and a recombinant mutant human hemoglobin“. In : *Proceedings of the National Academy of Sciences* 109.37 (2012), p. 14894-14899. DOI : [10.1073/pnas.1205809109](https://doi.org/10.1073/pnas.1205809109) (cf. p. 56).
- [64] E. R. Henry et J. Hofrichter. „Singular value decomposition : Application to analysis of experimental data“. In : *Methods in Enzymology*. Elsevier, 1992, p. 129-192. DOI : [10.1016/0076-6879\(92\)10010-b](https://doi.org/10.1016/0076-6879(92)10010-b) (cf. p. 56, 57, 59).
- [65] Ernest Hiebert, J.B. Bancroft et C.E. Bracker. „The assembly in vitro of some small spherical viruses, hybrid viruses, and other nucleoproteins“. In : *Virology* 34.3 (1968), p. 492-508. DOI : [10.1016/0042-6822\(68\)90069-x](https://doi.org/10.1016/0042-6822(68)90069-x) (cf. p. 64).
- [66] María Medrano, Miguel Ángel Fuertes, Alejandro Valbuena et al. „Imaging and Quantitation of a Succession of Transient Intermediates Reveal the Reversible Self-Assembly Pathway of a Simple Icosahedral Virus Capsid“. In : *Journal of the American Chemical Society* 138.47 (2016), p. 15385-15396. DOI : [10.1021/jacs.6b07663](https://doi.org/10.1021/jacs.6b07663) (cf. p. 64).
- [67] Adam Zlotnick, Ryan Aldrich, Jennifer M. Johnson, Pablo Ceres et Mark J. Young. „Mechanism of Capsid Assembly for an Icosahedral Plant Virus“. In : *Virology* 277.2 (2000), p. 450-456. DOI : [10.1006/viro.2000.0619](https://doi.org/10.1006/viro.2000.0619) (cf. p. 64).
- [68] Jennifer M. Johnson, Jinghua Tang, Yaw Nyame et al. „Regulating Self-Assembly of Spherical Oligomers“. In : *Nano Letters* 5.4 (2005), p. 765-770. DOI : [10.1021/nl1050274q](https://doi.org/10.1021/nl1050274q) (cf. p. 64, 69, 72, 77).

- [69]Guillaume Tresset, Clémence Le Coeur, Jean-François Bryche et al. „Norovirus Capsid Proteins Self-Assemble through Biphasic Kinetics via Long-Lived Stave-like Intermediates“. In : *Journal of the American Chemical Society* 135.41 (2013), p. 15373-15381. DOI : [10.1021/ja403550f](https://doi.org/10.1021/ja403550f) (cf. p. 64).
- [70]Roman Tuma, Hiro Tsuruta, Kenneth H. French et Peter E. Prevelige. „Detection of Intermediates and Kinetic Control during Assembly of Bacteriophage P22 Procapsid“. In : *Journal of Molecular Biology* 381.5 (2008), p. 1395-1406. DOI : [10.1016/j.jmb.2008.06.020](https://doi.org/10.1016/j.jmb.2008.06.020) (cf. p. 64).
- [71]Jason D. Perlmutter, Matthew R. Perkett et Michael F. Hagan. „Pathways for Virus Assembly around Nucleic Acids“. In : *Journal of Molecular Biology* 426.18 (2014), p. 3148-3165. DOI : [10.1016/j.jmb.2014.07.004](https://doi.org/10.1016/j.jmb.2014.07.004) (cf. p. 65, 103).
- [72]P.K. Sorger, P.G. Stockley et S.C. Harrison. „Structure and assembly of turnip crinkle virus“. In : *Journal of Molecular Biology* 191.4 (1986), p. 639-658. DOI : [10.1016/0022-2836\(86\)90451-1](https://doi.org/10.1016/0022-2836(86)90451-1) (cf. p. 66).
- [73]F. Qu et T.J. Morris. „Encapsidation of turnip crinkle virus is defined by a specific packaging signal and RNA size.“ In : *Journal of virology* 71 (2 fév. 1997), p. 1428-1435 (cf. p. 66).
- [74]Rees F. Garmann, Mauricio Comas-Garcia, Ajaykumar Gopal, Charles M. Knobler et William M. Gelbart. „The Assembly Pathway of an Icosahedral Single-Stranded RNA Virus Depends on the Strength of Inter-Subunit Attractions“. In : *Journal of Molecular Biology* 426.5 (2014), p. 1050-1060. DOI : [10.1016/j.jmb.2013.10.017](https://doi.org/10.1016/j.jmb.2013.10.017) (cf. p. 66, 67).
- [75]R. F. Garmann, M. Comas-Garcia, M. S. T. Koay et al. „Role of Electrostatics in the Assembly Pathway of a Single-Stranded RNA Virus“. In : *Journal of Virology* 88.18 (2014), p. 10472-10479. DOI : [10.1128/jvi.01044-14](https://doi.org/10.1128/jvi.01044-14) (cf. p. 67, 68, 72, 92).
- [76]Rees F. Garmann, Mauricio Comas-Garcia, Charles M. Knobler et William M. Gelbart. „Physical Principles in the Self-Assembly of a Simple Spherical Virus“. In : *Accounts of Chemical Research* 49.1 (2015), p. 48-55. DOI : [10.1021/acs.accounts.5b00350](https://doi.org/10.1021/acs.accounts.5b00350) (cf. p. 67, 72, 92).
- [77]R. D. Cadena-Nava, M. Comas-Garcia, R. F. Garmann et al. „Self-Assembly of Viral Capsid Protein and RNA Molecules of Different Sizes : Requirement for a Specific High Protein/RNA Mass Ratio“. In : *Journal of Virology* 86.6 (2011), p. 3318-3326. DOI : [10.1128/jvi.06566-11](https://doi.org/10.1128/jvi.06566-11) (cf. p. 67, 72).
- [78]Ruben D. Cadena-Nava, Mauricio Comas-Garcia, Rees F. Garmann et al. „Correction for Cadena-Nava et al., “Self-Assembly of Viral Capsid Protein and RNA Molecules of Different Sizes : Requirement for a Specific High Protein/RNA Mass Ratio”“. In : *Journal of Virology* 93.4 (2019). DOI : [10.1128/jvi.01939-18](https://doi.org/10.1128/jvi.01939-18) (cf. p. 68, 85).

- [79]M. Comas-Garcia, R. D. Cadena-Nava, A. L. N. Rao, C. M. Knobler et W. M. Gelbart. „In Vitro Quantification of the Relative Packaging Efficiencies of Single-Stranded RNA Molecules by Viral Capsid Protein“. In : *Journal of Virology* 86.22 (2012), p. 12271-12282. DOI : [10.1128/jvi.01695-12](https://doi.org/10.1128/jvi.01695-12) (cf. p. 70, 85, 92).
- [80]Morten K. Grøftehaug, Nelly R. Hajizadeh, Marcus J. Swann et Ehmke Pohl. „Protein–ligand interactions investigated by thermal shift assays (TSA) and dual polarization interferometry (DPI)“. In : *Acta Crystallographica Section D Biological Crystallography* 71.1 (2015), p. 36-44. DOI : [10.1107/s1399004714016617](https://doi.org/10.1107/s1399004714016617) (cf. p. 74).
- [81]G. V. Semisotnov, N. A. Rodionova, O. I. Razgulyaev et al. „Study of the ?molten globule? intermediate state in protein folding by a hydrophobic fluorescent probe“. In : *Biopolymers* 31.1 (1991), p. 119-128. DOI : [10.1002/bip.360310111](https://doi.org/10.1002/bip.360310111) (cf. p. 74).
- [82]Antonio Šiber et Rudolf Podgornik. „Role of electrostatic interactions in the assembly of empty spherical viral capsids“. In : *Physical Review E* 76.6 (2007). DOI : [10.1103/physreve.76.061906](https://doi.org/10.1103/physreve.76.061906) (cf. p. 75, 78, 117).
- [83]J. R. Vega-Acosta, R. D. Cadena-Nava, W. M. Gelbart, C. M. Knobler et J. Ruiz-García. „Electrophoretic Mobilities of a Viral Capsid, Its Capsid Protein, and Their Relation to Viral Assembly“. In : *The Journal of Physical Chemistry B* 118.8 (2014), p. 1984-1989. DOI : [10.1021/jp407379t](https://doi.org/10.1021/jp407379t) (cf. p. 78, 119).
- [84]Bodo D. Wilts, Iwan A.T. Schaap et Christoph F. Schmidt. „Swelling and Softening of the Cowpea Chlorotic Mottle Virus in Response to pH Shifts“. In : *Biophysical Journal* 108.10 (2015), p. 2541-2549. DOI : [10.1016/j.bpj.2015.04.019](https://doi.org/10.1016/j.bpj.2015.04.019) (cf. p. 81).
- [85]Guillaume Tresset, Mouna Tatou, Clémence Le Cœur et al. „Weighing Polyelectrolytes Packaged in Viruslike Particles“. In : *Physical Review Letters* 113.12 (2014). DOI : [10.1103/physrevlett.113.128305](https://doi.org/10.1103/physrevlett.113.128305) (cf. p. 82, 84, 99, 100).
- [86]Maxim V. Petoukhov, Peter V. Konarev, Alexey G. Kikhney et Dmitri I. Svergun. „ATSAS 2.1 – towards automated and web-supported small-angle scattering data analysis“. In : *Journal of Applied Crystallography* 40.s1 (2007), s223-s228. DOI : [10.1107/s0021889807002853](https://doi.org/10.1107/s0021889807002853) (cf. p. 89, 140, 144).
- [87]James M. Fox, Guoji Wang, Jeffrey A. Speir et al. „Comparison of the Native CCMV Virion within VitroAssembled CCMV Virions by Cryoelectron Microscopy and Image Reconstruction“. In : *Virology* 244.1 (1998), p. 212-218. DOI : [10.1006/viro.1998.9107](https://doi.org/10.1006/viro.1998.9107) (cf. p. 98).
- [88]Damien Baigl, Thomas A. P. Seery et Claudine E. Williams. „Preparation and Characterization of Hydrosoluble, Partially Charged Poly(styrenesulfonate)s of Various Controlled Charge Fractions and Chain Lengths“. In : *Macromolecules* 35.6 (2002), p. 2318-2326. DOI : [10.1021/ma011707o](https://doi.org/10.1021/ma011707o) (cf. p. 101).

- [89] Siyu Li, Gonca Erdemci-Tandogan, Paul van der Schoot et Roya Zandi. „The effect of RNA stiffness on the self-assembly of virus particles“. In : *Journal of Physics : Condensed Matter* 30.4 (2017), p. 044002. DOI : [10.1088/1361-648x/aaa159](https://doi.org/10.1088/1361-648x/aaa159) (cf. p. 101).
- [90] Izabela A. Rodenhuis-Zybert, Jan Wilschut et Jolanda M. Smit. „Partial maturation : an immune-evasion strategy of dengue virus?“ In : *Trends in Microbiology* 19.5 (2011), p. 248-254. DOI : [10.1016/j.tim.2011.02.002](https://doi.org/10.1016/j.tim.2011.02.002) (cf. p. 101).
- [91] Adam Zlotnick et Suchetana Mukhopadhyay. „Virus assembly, allostery and antivirals“. In : *Trends in Microbiology* 19.1 (2011), p. 14-23. DOI : [10.1016/j.tim.2010.11.003](https://doi.org/10.1016/j.tim.2010.11.003) (cf. p. 108).
- [92] Christopher C. DuFort et Bogdan Dragnea. „Bio-enabled synthesis of metamaterials“. In : *Annual Review of Physical Chemistry* 61.1 (2010), p. 323-344. DOI : [10.1146/annurev.physchem.012809.103300](https://doi.org/10.1146/annurev.physchem.012809.103300) (cf. p. 108).
- [93] T. Douglas. „Viruses : Making Friends with Old Foes“. In : *Science* 312.5775 (2006), p. 873-875. DOI : [10.1126/science.1123223](https://doi.org/10.1126/science.1123223) (cf. p. 108).
- [94] Michael F. HAGAN. „Modeling Viral Capsid Assembly“. In : *Advances in Chemical Physics*. John Wiley & Sons, Inc., 2014, p. 1-68. DOI : [10.1002/9781118755815.ch01](https://doi.org/10.1002/9781118755815.ch01) (cf. p. 109).
- [95] Adam Zlotnick, Jennifer M. Johnson, Paul W. Wingfield, Stephen J. Stahl et Dan Endres. „A Theoretical Model Successfully Identifies Features of Hepatitis B Virus Capsid Assembly“. In : *Biochemistry* 38.44 (1999), p. 14644-14652. DOI : [10.1021/bi991611a](https://doi.org/10.1021/bi991611a) (cf. p. 110, 113, 114, 147).
- [96] Kris Holmes, Dale A. Shepherd, Alison E. Ashcroft et al. „Assembly Pathway of Hepatitis B Core Virus-like Particles from Genetically Fused Dimers“. In : *Journal of Biological Chemistry* 290.26 (2015), p. 16238-16245. DOI : [10.1074/jbc.m114.622035](https://doi.org/10.1074/jbc.m114.622035) (cf. p. 110).
- [97] A. Zlotnick, N. Cheng, J. F. Conway et al. „Dimorphism of Hepatitis B Virus Capsids Is Strongly Influenced by the C-Terminus of the Capsid Protein“. In : *Biochemistry* 35.23 (1996), p. 7412-7421. DOI : [10.1021/bi9604800](https://doi.org/10.1021/bi9604800) (cf. p. 111).
- [98] Elizabeth E. Pierson, David Z. Keifer, Lisa Selzer et al. „Detection of Late Intermediates in Virus Capsid Assembly by Charge Detection Mass Spectrometry“. In : *Journal of the American Chemical Society* 136.9 (2014), p. 3536-3541. DOI : [10.1021/ja411460w](https://doi.org/10.1021/ja411460w) (cf. p. 111, 140, 141).
- [99] Panagiotis Kondylis, Christopher J. Schlicksup, Sarah P. Katen et al. „Evolution of Intermediates during Capsid Assembly of Hepatitis B Virus with Phenylpropenamide-Based Antivirals“. In : *ACS Infectious Diseases* 5.5 (2019), p. 769-777. DOI : [10.1021/acsinfectdis.8b00290](https://doi.org/10.1021/acsinfectdis.8b00290) (cf. p. 111, 112).

- [100] Sarah P. Katen, Srinivas Reddy Chirapu, M. G. Finn et Adam Zlotnick. „Trapping of Hepatitis B Virus Capsid Assembly Intermediates by Phenylpropenamide Assembly Accelerators“. In : *ACS Chemical Biology* 5.12 (2010), p. 1125-1136. DOI : [10.1021/cb100275b](https://doi.org/10.1021/cb100275b) (cf. p. 111).
- [101] Xinyu Sun, Dong Li, Zhaoshuai Wang et al. „A dimorphism shift of hepatitis B virus capsids in response to ionic conditions“. In : *Nanoscale* 10.36 (2018), p. 16984-16989. DOI : [10.1039/c8nr03370f](https://doi.org/10.1039/c8nr03370f) (cf. p. 112, 151, 152).
- [102] Pablo Ceres et Adam Zlotnick. „Weak Protein-Protein Interactions Are Sufficient To Drive Assembly of Hepatitis B Virus Capsids“. In : *Biochemistry* 41.39 (2002), p. 11525-11531. DOI : [10.1021/bi0261645](https://doi.org/10.1021/bi0261645) (cf. p. 113, 119, 136, 137, 147).
- [103] Willem K. Keghel et Paul van der Schoot. „Competing Hydrophobic and Screened-Coulomb Interactions in Hepatitis B Virus Capsid Assembly“. In : *Biophysical Journal* 86.6 (2004), p. 3905-3913. DOI : [10.1529/biophysj.104.040055](https://doi.org/10.1529/biophysj.104.040055) (cf. p. 117).
- [104] Xinyu Sun, Dong Li, Zhaoshuai Wang et al. „Role of Protein Charge Density on Hepatitis B Virus Capsid Formation“. In : *ACS Omega* 3.4 (2018), p. 4384-4391. DOI : [10.1021/acsomega.8b00021](https://doi.org/10.1021/acsomega.8b00021) (cf. p. 117, 119).
- [105] Jingzhi Chen, Maelenn Chevreuil, Sophie Combet, Yves Lansac et Guillaume Tresset. „Investigating the thermal dissociation of viral capsid by lattice model“. In : *Journal of Physics : Condensed Matter* 29.47 (2017), p. 474001. DOI : [10.1088/1361-648x/aa8d88](https://doi.org/10.1088/1361-648x/aa8d88) (cf. p. 118).
- [106] Brian M. Baker et Kenneth P. Murphy. „Prediction of binding energetics from structure using empirical parameterization“. In : *Methods in Enzymology*. Elsevier, 1998, p. 294-315. DOI : [10.1016/s0076-6879\(98\)95045-5](https://doi.org/10.1016/s0076-6879(98)95045-5) (cf. p. 119).
- [107] Ranjit Prasad Bahadur, Pinak Chakrabarti, Francis Rodier et Joel Janin. „Dissecting subunit interfaces in homodimeric proteins“. In : *Proteins : Structure, Function, and Genetics* 53.3 (2003), p. 708-719. DOI : [10.1002/prot.10461](https://doi.org/10.1002/prot.10461) (cf. p. 119).
- [108] Nancy Horton et Mitchell Lewis. „Calculation of the free energy of association for protein complexes“. In : *Protein Science* 1.1 (2008), p. 169-181. DOI : [10.1002/pro.5560010117](https://doi.org/10.1002/pro.5560010117) (cf. p. 119).
- [109] Alexander I. Norman, Derek L. Ho, Alamgir Karim et Eric J. Amis. „Phase behavior of block co-poly(ethylene oxide–butylene oxide), E18B9 in water, by small angle neutron scattering“. In : *Journal of Colloid and Interface Science* 288.1 (2005), p. 155-165. DOI : [10.1016/j.jcis.2005.02.091](https://doi.org/10.1016/j.jcis.2005.02.091) (cf. p. 130).
- [110] Tim E. Williamson, Bruce A. Craig, Elena Kondrashkina, Chris Bailey-Kellogg et Alan M. Friedman. „Analysis of self-associating proteins by singular value decomposition of solution scattering data“. In : *Biophysical Journal* 94.12 (2008), p. 4906-4923. DOI : [10.1529/biophysj.107.113167](https://doi.org/10.1529/biophysj.107.113167) (cf. p. 141).

- [111]Paul van der Schoot et Roya Zandi. „Kinetic theory of virus capsid assembly“. In : *Physical Biology* 4.4 (2007), p. 296-304. DOI : [10.1088/1478-3975/4/4/006](https://doi.org/10.1088/1478-3975/4/4/006) (cf. p. 143).
- [112]Sanaz Panahandeh, Siyu Li, Laurent Marichal et al. „How a virus circumvents energy barriers to form symmetric shells“. 2019 (cf. p. 150).
- [113]Nathalie Barbier. „Étude des erreurs programmées du ribosome mammifère par microscopie de fluorescence en molécule unique“. Thèse de doct. Onde et Matière, 2017 (cf. p. 161, 163).
- [114]Maren Vogel, Jolanta Vorreiter et Michael Nassal. „Quaternary structure is critical for protein display on capsid-like particles (CLPs) : Efficient generation of hepatitis B virus CLPs presenting monomeric but not dimeric and tetrameric fluorescent proteins“. In : *Proteins : Structure, Function, and Bioinformatics* 58.2 (2004), p. 478-488. DOI : [10.1002/prot.20312](https://doi.org/10.1002/prot.20312) (cf. p. 164).
- [115]Maren Vogel, Manuel Diez, Jochen Einfeld et Michael Nassal. „In vitro assembly of mosaic hepatitis B virus capsid-like particles (CLPs) : Rescue into CLPs of assembly-deficient core protein fusions and FRET-suited CLPs“. In : *FEBS Letters* 579.23 (2005), p. 5211-5216. DOI : [10.1016/j.febslet.2005.08.044](https://doi.org/10.1016/j.febslet.2005.08.044) (cf. p. 164).
- [116]J. Zachary Porterfield et Adam Zlotnick. „A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance“. In : *Virology* 407.2 (2010), p. 281-288. DOI : [10.1016/j.virol.2010.08.015](https://doi.org/10.1016/j.virol.2010.08.015) (cf. p. 173).
- [117]Olesia V. Stepanenko, Vladislav V. Verkhusha, Vasili I. Kazakov et al. „Comparative Studies on the Structure and Stability of fluorescent Proteins EGFP, zFP506, mRFP1, “dimer2”, and DsRed1†“. In : *Biochemistry* 43.47 (2004), p. 14913-14923. DOI : [10.1021/bi048725t](https://doi.org/10.1021/bi048725t) (cf. p. 181).
- [118]Stephen H. Bokman et William W. Ward. „Renaturation of aequorea green-fluorescent protein“. In : *Biochemical and Biophysical Research Communications* 101.4 (1981), p. 1372-1380. DOI : [10.1016/0006-291x\(81\)91599-0](https://doi.org/10.1016/0006-291x(81)91599-0) (cf. p. 181).
- [119]Mária PÁLFI, Erika KOVÁCS, László SZILÁGYI, Ildikó MIKLÓSSY and Beáta ÁBRAHÁM et Szabolcs LÁNYI. „fluorescence Quenching Analysis of Histidine-tagged Enhanced Green fluorescent Protein“. In : *UPB Scientific Bulletin B* (2010) (cf. p. 181).
- [120]Kichul Jung, Jaebok Park, Pil-Jae Maeng et Hackjin Kim. „Fluorescence quenching of green fluorescent protein during denaturation by guanidine“. In : *Bulletin of the Korean Chemical Society* 26 (2005), p. 413-417. DOI : [10.5012/bkcs.2005.26.3.413](https://doi.org/10.5012/bkcs.2005.26.3.413) (cf. p. 181).
- [121]R. Heim, D. C. Prasher et R. Y. Tsien. „Wavelength mutations and posttranslational autoxidation of green fluorescent protein.“ In : *Proceedings of the National Academy of Sciences* 91.26 (1994), p. 12501-12504. DOI : [10.1073/pnas.91.26.12501](https://doi.org/10.1073/pnas.91.26.12501) (cf. p. 183).



- [122] Sawako Enoki, Kimiko Saeki, Kosuke Maki et Kunihiro Kuwajima. „Acid denaturation and refolding of green fluorescent protein“. In : *Biochemistry* (2004). DOI : [10.1021/bi048733](https://doi.org/10.1021/bi048733) (cf. p. 185).
- [123] Anastasia V. Mamontova, Alexey M. Bogdanov et Konstantin A. Lukyanov. „Influence of cell growth conditions and medium composition on EGFP photostability in live cells“. In : *BioTechniques* 58.5 (2015). DOI : [10.2144/000114289](https://doi.org/10.2144/000114289) (cf. p. 185).
- [124] Alexey Bogdanov, Ekaterina Bogdanova, Dmitriy Chudakov et al. „Cell culture medium affects GFP photostability : a solution“. In : *Nature Methods* 6.12 (2009), p. 859-860. DOI : [10.1038/nmeth1209-859](https://doi.org/10.1038/nmeth1209-859) (cf. p. 185).
- [125] Ji A. Sohn, Samuel Litwin et Christoph Seeger. „Mechanism for CCC DNA Synthesis in Hepadnaviruses“. In : *PLoS ONE* 4.11 (2009). Sous la dir. de Wang-Shick Ryu, e8093. DOI : [10.1371/journal.pone.0008093](https://doi.org/10.1371/journal.pone.0008093) (cf. p. 196).



**Titre :** Phénomènes dynamiques d'auto-assemblage et de désassemblage dans des virus icosaédriques

**Mots clés :** Dynamique, auto-assemblage, désassemblage, virus, ARN, TR-SAXS

**Résumé :** L'auto-assemblage et le désassemblage de la capsid des virus, étapes critiques du cycle viral, est un sujet qui suscite beaucoup d'intérêt. Cependant, les mécanismes sous-jacents et, en particulier, les chemins dynamique d'assemblage et de désassemblage des capsides, vides ou pleines, des virus ne sont pas entièrement résolues. La diffusion des rayons X aux petits angles résolue en temps, combinée à la décomposition en valeur singulière, est une technique qui permet d'étudier des processus impliquant des espèces de taille nanométrique avec une résolution temporelle de l'ordre de la milliseconde. La technique de criblage de thermo-stabilité des protéines par fluorescence, couplée à un modèle théorique de champ moyen, permet d'estimer expérimentalement les énergies d'interactions entre les protéines et la charge effective de celles-ci. Dans le cas du virus de la marbrure chlorotique de la cornille, ces techniques nous ont permis d'étudier la dynamique d'auto-assemblage des protéines des capsides autour de leur

formation de complexes amorphes via un chemin cinétique appelé *en masse* tandis que leur relaxation en virions s'effectue via un chemin cinétique dit *synchrone*. Les énergies de liaison des protéines avec le génome se sont révélées modérées tandis qu'une barrière d'énergie élevée sépare les complexes des virions. Des expériences complémentaires ont également montré que cette barrière pouvait être abaissée, de sorte que des capsides pleines se forment directement. Dans le cas du virus de l'hépatite B, nous avons étudié les dynamiques d'assemblage et de désassemblage des capsides vides. Nous avons pu identifier un chemin cinétique probable de désassemblage avec la présence d'une espèce intermédiaire assimilable à une structure fractale-branchée. Les expériences d'assemblage ont révélé un chemin cinétique en trois phases, i.e. agglomération, croissance et relaxation, dirigé par l'attraction hydrophobe et modulé par la répulsion électrostatique. De plus, certaines expériences ont également montré que la dernière phase pouvait être facilement inhibée.

**Title :** Self-assembly and disassembly dynamical phenomena in icosahedral viruses

**Keywords :** Dynamics, self-assembly, disassembly, virus, ARN, TR-SAXS

**Abstract:** The self-assembly and disassembly of virus capsids, critical stages of the viral cycle, is a subject that arouses a lot of interest. However, the underlying mechanisms and, in particular, the kinetic pathways of assembly and disassembly, whether the capsid is empty or full, are not entirely solved. Time-resolved small-angle X-ray scattering is a technique for tracking processes involving nano-sized species with a time resolution of the millisecond order. Combined with the singular value decomposition, a model-independent method of analysis, the technique allows the investigation of the kinetics of capsid self-assembly and disassembly. In addition, the fluorescence thermal shift assays, coupled with a theoretical mean-field model, makes it possible to experimentally estimate the interaction energies between the proteins as well as their effective charge. Thus, in the case of the cowpea chlorotic mottle virus (CCMV), we have studied the self-assembly dynamics of capsid proteins with their genetic material. The experiments revealed the formation of amorphous

complexes via a kinetic pathway called *en masse* whereas their relaxation into virions occurs via a so-called synchronous kinetic path. The binding energies of proteins with the genome showed to be moderate, while the energetic barrier separating the complexes of virions is high. The experiments performed with a synthetic polyelectrolyte showed that this barrier could be lowered, so that full capsids form under conditions where virions cannot. In the case of the hepatitis B virus (HBV), we studied the self-assembly and disassembly dynamics of empty capsids. First, we were able to identify a probable disassembly kinetic path with the presence of an intermediate species organised like a fractal-branched structure. Finally, the assembly experiments revealed a kinetic pathway in three phases, i.e. agglomeration, growth and relaxation, directed by hydrophobic attraction and modulated by electrostatic repulsion. Due to the low effective charge of the protein, the last phase could be inhibited by raising the salinity, thus trapping the capsids in a conformation containing defects.