



**HAL**  
open science

# Apport des récentes évolutions de la cryo-microscopie électronique et du traitement d'images dans l'étude structurale de virus de plantes

François Lecorre

## ► To cite this version:

François Lecorre. Apport des récentes évolutions de la cryo-microscopie électronique et du traitement d'images dans l'étude structurale de virus de plantes. Médecine humaine et pathologie. Université Montpellier, 2016. Français. NNT: 2016MONTT061 . tel-01530757

**HAL Id: tel-01530757**

**<https://theses.hal.science/tel-01530757>**

Submitted on 31 May 2017

**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.

# THÈSE

Pour obtenir le grade de  
**Docteur**

**Délivré par l'Université de Montpellier**

**Préparée au sein de l'école doctorale Sciences  
Chimiques et Biologiques pour la Santé  
Et de l'unité de recherche Centre de Biochimie  
Structurale CNRS UMR 5048 – INSERM U1054**

**Spécialité : Biologie Santé**

**Présentée par François LECORRE**

**Apport des récentes évolutions de la  
cryo-microscopie électronique et du  
traitement d'images dans l'étude  
structurale de virus de plantes**

**Soutenue le 14 décembre 2016 devant le jury composé de**

M. Olivier LAMBERT, DR, UMR 5248	Rapporteur
M. Guy SCHOEHN, DR, UMR 5075	Rapporteur
Mme Célia PLISSON-CHASTANG, CR, UMR 5099	Examineur
M. Stefano TRAPANI, MCU, UMR 5048 - U1054	Examineur
M. Christian ROUMESTAND, PR, UMR 5048 - U1054	Président du jury
M. Patrick BRON, DR, UMR 5048 - U1054	Directeur de thèse





# Remerciements

---

*Je remercie le Docteur Olivier Lambert et le Docteur Guy Schoehn de m'avoir fait l'honneur d'accepter d'être rapporteurs de ce travail de thèse, et pour leur lecture attentive de mon manuscrit, ainsi que la Docteure Célia Plisson-Chastang qui a eu la gentillesse d'examiner mon travail, et le Professeur Christian Roumestand, qui malgré son appartenance à la RMN, a accepté d'examiner mon travail et de présider ce jury de thèse.*

*Je remercie le gouvernement de la Nouvelle-Calédonie d'avoir financé ces travaux de thèse.*

*Le travail présenté dans ce manuscrit a été réalisé au Centre de Biochimie Structurale, dirigé successivement par Christian Roumestand et Pierre-Emmanuel Millhiet que je remercie chaleureusement pour leur accueil au sein du laboratoire.*

*Je remercie mon directeur de thèse, Patrick Bron, de m'avoir accepté dans son équipe et guidé, depuis mon stage de Master 2 jusqu'à la fin de ma thèse.*

*Une grande partie de mes travaux a été conduite en collaboration avec l'équipe du Docteur Stéphane Blanc du laboratoire Biologie et Génétique des Interactions Plante-Parasite que je remercie sincèrement pour son aide et sa bonne humeur. Je remercie Maryline pour ses précieux conseils et son dynamisme à toute épreuve, et surtout le Docteur Jean-Louis Zeddam pour son temps et son soutien tout au long de ma thèse.*

*Merci à mes premiers collègues de bureau pour leur bonne humeur et pour tous les gâteaux partagés, je vous dois mes premiers kilos : Aurélie F, Alice, Yannis et Angélique.*

*Merci à Joséphine, pour sa gentillesse, ses conseils et l'attention dont elle a fait preuve depuis 4 ans.*

*Merci à Stefano, jamais présent et pourtant toujours disponible, pour ta pédagogie et ta patience.*

*Merci aux autres membres de mon équipe pour leur soutien, François, Aurélie, Chloé.*

*Merci à Vanessa pour son écoute, ainsi qu'à Christine D pour son aide.*

*Merci à Ewelina et Mariya pour nos longues discussions et tous nos fous rires. Je ne compte plus les trop nombreux shots de vodka à 16h.*

*Merci à Selma pour sa gentillesse, à Zhanna pour son enthousiasme, et à Nath pour toutes nos discussions inutiles.*

*À tous les étudiants et postdocs passés et présents, sans vous le labo n'aurait pas été le même, Antoine, Anaïs, Désirée, Mélissa, Jad, Laura, Luca(s), Tiago, Ashley, Fatima, Baptiste, Annika, Andrès, Elise, Cao, Sarah et Anthony. Merci pour tous ces moments passés au labo et en dehors.*

*Merci à mes deux compères : à Jérôme, pour ta bonne humeur contagieuse et toutes ces soirées éthyliques passées chez toi. Et à Laurent, mon camarade de thèse, pessimiste malgré lui, toujours présent pour décourager les autres. Sans vous ces années n'auraient jamais été aussi mémorables.*

*Merci à tous mes amis calédoniens, paradoxalement tous rencontrés à Montpellier, que j'ai appris à apprécier sans vraiment le vouloir : Nico, Styvens, Olivier, Maé, Babin. À toutes nos galères inoubliables.*

*Je tiens à remercier ma famille, qui malgré la distance m'a soutenue moralement pendant tout mon parcours. Je pense à vous.*

*Mes derniers remerciements vont naturellement à Coralie, merci, simplement pour tout.*



## Résumé en français

---

La cryo-microscopie électronique a connu une révolution majeure ces dernières années, liée à des évolutions technologiques importantes, tant au niveau des microscopes, que des caméras ou des logiciels de traitement d'image. Avec l'arrivée de microscopes électroniques plus stables mécaniquement et électroniquement, il est possible d'enregistrer en quelques jours plusieurs milliers d'images de façon automatique, et obtenir ainsi un jeu initial d'images des complexes macromoléculaires étudiés pouvant contenir plusieurs centaines de milliers de particules. Sauf que maintenant, il ne s'agit plus d'images au sens strict, mais de films. En effet, les nouvelles caméras, dites à détection directe d'électrons qui ont une sensibilité 10 fois supérieure à celle des anciennes caméras, permettent d'enregistrer non pas une image, mais une série d'images au cours de l'exposition. Bien que piégées dans une mince pellicule de glace, l'analyse des séries d'images a permis de montrer que les particules macromoléculaires pouvaient bouger sous l'effet du faisceau d'électrons. Ces mouvements peuvent donc être maintenant corrigés. Ainsi, alors que pendant des dizaines d'années, les informations structurales extraites des images des microscopes électroniques étaient limitées à la moyenne résolution (entre 5 et 15 Å), nous voyons apparaître aujourd'hui de très nombreuses cartes de densité de complexes macromoléculaires issues de la microscopie électronique, à des résolutions meilleures que 4 Å, permettant de construire des modèles atomiques à l'aide d'outils jusqu'alors réservés à la cristallographie aux rayons X.

C'est dans ce contexte, que j'ai étudié par cryo-microscopie électronique et traitement d'image, l'organisation structurale de trois virus de plante :

- Le virus de la mosaïque de l'arabette (ArMV), un *Nepovirus* uniquement transmis par le nématode *Xiphinema diversicaudatum*, qui est responsable de la maladie du court-noué de la vigne ;
- Le virus de la tâche de la fève (BBSV), un *Comovirus* transmis par les coléoptères, responsable de la dégénération chez les légumineuses ;
- Le virus de la mosaïque du chou-fleur (CaMV), un *Caulimovirus* servant de virus modèle pour l'étude de la transmission des virus non circulants.

Ces virus sont des parasites endocellulaires obligatoires, dont l'efficacité de propagation dépend de leur capacité de réplication au sein de la cellule infectée et de leur transmission vers de nouveaux hôtes. En raison de l'immobilité des plantes, les phytovirus font souvent appel à des vecteurs pour la transmission plante à plante, qui sont principalement des insectes, des nématodes, des champignons ou des acariens. Les phytovirus sont généralement responsables d'une baisse importante de la croissance de la plante et des fruits, voire de la mort de l'hôte infecté. Les dégâts ainsi causés engendrent des pertes de rendement dans les cultures à travers le monde, se traduisant par d'énormes pertes économiques pour les cultivateurs. Ce travail de thèse présente les structures atomiques de l'ArMV et du BBSV obtenues par cryo-microscopie électronique, ainsi que les premières données structurales de la protéine virale P2 impliquée dans la transmission du CaMV par le puceron.

# Résumé en anglais

---

A significant revolution has occurred in the world of cryo-electron microscopy in the last five years, as the result of major upgrades of microscopes, cameras and image processing software. With the advent of high-end microscopes, mechanically and electronically more stable, it is now possible to record thousands of images in a few days through automated acquisition, and thereby obtain an initial data set of hundreds of thousands of particles. Also, since the latest direct electron detectors have a higher sensitivity than the previous generation cameras, single frame image acquisition has been superseded by multi-frame (movie) recording. Frame processing has revealed that the imaged particles, in spite of being trapped in a thin ice layer, very often undergo some motion during beam exposure, and correction of this beam-induced motion drastically improves the quality of data. Thus, while the resolution level of cryo-electron microscopy reconstructions was, until recently, limited to within 5 and 15 Å, numerous electron density maps of protein complexes with better than 4 Å resolution are being solved now, allowing the construction of atomic models using tools originally developed for X-ray crystallography.

In this context, I have studied the structural organization of three plant viruses, using cryo-electron microscopy and image processing:

- Arabis Mosaic Virus (ArMV), which is a *Nepovirus* transmitted exclusively by the nematode *Xiphinema diversicaudatum*, responsible for a disease of vineyards;
- Broad Bean Stain Virus (BBSV), which is a *Comovirus* transmitted by beetles, responsible for the degeneration of leguminous plants;
- Cauliflower Mosaic Virus (CaMV), a *Caulimovirus* which is a model virus studied to characterize the transmission of non circulative viruses.

Viruses are obligate intracellular parasites, whose efficiency is directly related to their replicative capacity inside the infected cell, and its transmission to new hosts. Plants are immobile, and therefore plant viruses often use vectors (mainly insects, nematodes, fungi or mites) for plant-to-plant transmission. Plant viruses are generally responsible for a significant decrease in plant and fruit growth, and even the death of the plant. Viruses are devastating fields worldwide, causing huge losses in crop yield each year. This study presents the atomic structures of both ArMV and BBSV capsids, as well as the first data about the CaMV capsid and its transmission protein P2.



# Table des matières

---

Résumé en français .....	3
Résumé en anglais.....	4
Table des figures .....	10
Liste des abréviations .....	14
Avant-propos .....	17
<b>Chapitre 1 : Introduction générale .....</b>	<b>20</b>
<b>1. Introduction à la microscopie électronique à transmission .....</b>	<b>20</b>
1.1. Anatomie d'un microscope électronique à transmission .....	21
1.1.1. Les sources d'électrons.....	21
1.1.2. La formation de l'image .....	22
1.1.2.1. L'origine du contraste de l'image .....	22
1.1.2.2. Les lentilles électromagnétiques .....	24
1.1.2.3. Relation entre diffraction et image .....	25
1.1.3. Les limites du système d'imagerie .....	25
1.1.3.1. Résolution maximale théorique .....	25
1.1.3.2. Les aberrations dues aux lentilles .....	26
1.1.3.3. La fonction de transfert de contraste (FTC) .....	27
1.1.3.4. La fonction d'enveloppe .....	29
1.2. L'étude d'un échantillon biologique .....	30
1.2.1. La coloration négative .....	30
1.2.2. La cryo-microscopie électronique : la porte vers la haute résolution .....	31
1.2.3. De l'image à la structure 3D de l'objet .....	34
1.2.3.1. Les conditions d'acquisition : le mode « faible-dose » .....	34
1.2.3.2. Principe de la reconstruction 3D en microscopie électronique .....	35
1.2.3.3. Obtention d'un modèle de référence .....	36
1.2.3.3.1. L'inclinaison conique aléatoire et la reconstruction inclinée orthogonale.....	37
1.2.3.3.2. Le théorème des lignes communes des projections.....	37
1.2.3.3.3. Tomographie électronique et addition de sous-volumes.....	38
1.3. La révolution de la microscopie électronique .....	39
1.3.1. Des échantillons plus stables et des conditions plus reproductibles.....	41
1.3.1.1. Les systèmes de congélation actuels.....	41
1.3.1.2. La nouvelle ruée vers l'or .....	42
1.3.1.3. Vers une technique entièrement automatisée ?.....	44
1.3.2. L'avènement du titan.....	45
1.3.3. Les caméras à détection directe .....	47
1.3.3.1. Un meilleur rapport signal/bruit .....	48
1.3.3.2. Le mode film.....	51
1.3.4. Le traitement des films .....	52
1.3.5. Les méthodes de validation de la résolution .....	53
1.3.6. Le dépôt des structures .....	53
1.3.7. Conclusion .....	54
<b>2. Objectifs de la thèse.....</b>	<b>55</b>
<b>Chapitre 2 : Étude structurale du virus du retard de croissance et de la nécrose de la fève</b>	<b>59</b>
<b>1. Introduction.....</b>	<b>59</b>
<b>2. Matériel et méthodes .....</b>	<b>63</b>
2.1. Expression de protéines recombinantes dans les bactéries <i>Escherichia coli</i> ( <i>E. coli</i> ).....	63

2.1.1.	Les différents plasmides .....	63
2.1.2.	Les différentes souches <i>E. coli</i> .....	64
2.1.3.	Les conditions de culture .....	65
2.1.4.	Purification de la protéine .....	65
2.2.	Expression de protéines recombinantes dans les cellules d'insectes <i>Spodoptera frugiperda 9 (Sf9)</i> .....	66
2.3.	Caractérisation de la protéine de capsid par cryo-ME .....	66
2.3.1.	Purification du virus <i>in planta</i> .....	66
2.3.2.	Congélation de l'échantillon .....	67
2.3.3.	Acquisition des images .....	67
<b>3.</b>	<b>Résultats</b> .....	<b>68</b>
3.1.	Expression de protéines recombinantes dans la bactérie <i>E. coli</i> .....	68
3.2.	Expression de protéines recombinantes dans les cellules d'insecte <i>Sf9</i> .....	69
3.3.	Caractérisation de la protéine de capsid par cryo-ME .....	71
<b>4.</b>	<b>Conclusion</b> .....	<b>72</b>
<b>Chapitre 3 : Étude structurale du virus de la mosaïque de l'arabette</b> .....		<b>74</b>
<b>1.</b>	<b>Introduction</b> .....	<b>74</b>
<b>2.</b>	<b>Matériel et méthodes</b> .....	<b>78</b>
2.1.	Congélation de l'échantillon .....	78
2.2.	Acquisitions des films .....	78
2.3.	Le traitement d'image et la reconstruction 3D des cartes de densité .....	78
2.3.1.	Le tri des films .....	78
2.3.2.	La sélection des particules des films .....	80
2.3.3.	Calcul de la FTC et des valeurs de défocalisation .....	80
2.3.4.	La reconstruction 3D du virus .....	81
2.3.4.1.	La classification 2D .....	81
2.3.4.2.	La classification 3D .....	82
2.3.4.3.	L'affinement 3D automatique .....	82
2.3.4.4.	Le post-traitement d'image .....	83
2.3.5.	Le traitement des vues .....	84
2.4.	La construction et le premier affinement du modèle atomique .....	84
<b>3.</b>	<b>Résultats et discussion</b> .....	<b>85</b>
3.1.	Reconstruction 3D de la carte de densité .....	85
3.1.1.	Classification 2D .....	85
3.1.2.	Classification 3D .....	87
3.1.3.	Affinement automatique et post-traitement d'image .....	89
3.1.4.	L'impact des vues sur la reconstruction 3D .....	91
3.1.5.	Structure de la capsid de l'ArMV .....	93
3.2.	Affinement du modèle atomique .....	96
3.2.1.	Positionnement de la carte de densités et du modèle .....	97
3.2.2.	Ajustement du monomère du GFLV dans la carte de densité .....	98
3.2.3.	Construction du modèle .....	98
3.2.4.	Ajustement des paramètres de la carte de densité .....	99
3.2.4.1.	Ajustement de l'aiguissement gaussien de la carte .....	99
3.2.4.2.	Ajustement de la taille de pixels de la carte .....	100
3.2.4.3.	Ajustement des paramètres liés à l'affinement du modèle atomique .....	101
3.2.4.3.1.	Création des restrictions géométriques supplémentaires .....	101
3.2.4.3.2.	Optimisation du poids des données expérimentales .....	101
3.2.4.3.3.	Création d'un octamère .....	102
3.2.4.3.4.	Création des restrictions NCS .....	103
3.2.5.	Affinement du modèle atomique .....	103
3.2.6.	Vérification de la géométrie du modèle .....	104
3.2.7.	Estimation du surrecalage (« overfitting ») .....	105



3.2.8.	Modèle atomique de l'ArMV .....	106
3.2.9.	Interaction ARN/protéine de capsid .....	108
3.2.10.	Première comparaison GFLV/ArMV.....	110
<b>4.</b>	<b>Conclusion .....</b>	<b>111</b>
<b>Chapitre 4 :</b>	<b>Etude structurale du virus de la tache de la fève .....</b>	<b>113</b>
<b>1.</b>	<b>Introduction .....</b>	<b>113</b>
<b>2.</b>	<b>Article scientifique .....</b>	<b>114</b>
<b>3.</b>	<b>Conclusion .....</b>	<b>140</b>
<b>Chapitre 5 :</b>	<b>Etude structurale du virus de la mosaïque du chou-fleur .....</b>	<b>142</b>
<b>1.</b>	<b>Introduction .....</b>	<b>142</b>
1.1.	Génome du virus .....	142
1.2.	Multiplication virale dans la cellule de plante .....	143
1.3.	Étude de la transmission plante-vecteur du CaMV .....	145
<b>2.</b>	<b>Matériel et Méthodes .....</b>	<b>153</b>
2.1.	Étude de la protéine de capsid du CaMV .....	153
2.1.1.	Purification du virus du CaMV à partir de plantes infectées .....	153
2.1.2.	Préparation des grilles .....	153
2.1.3.	Congélation de l'échantillon .....	154
2.1.4.	Acquisition des films .....	154
2.1.5.	Traitement d'image .....	154
2.2.	Étude de la protéine P2 exprimée dans <i>E. coli</i> .....	154
2.3.	Étude du complexe 157m: microtubules .....	155
2.3.1.	Purification du complexe à partir de cellules d'insectes <i>Sf9</i> infectées par des baculovirus .....	155
2.3.2.	Congélation de l'échantillon .....	155
2.3.3.	Acquisition des micrographies .....	155
2.3.4.	Traitement d'image .....	156
<b>3.</b>	<b>Résultats .....</b>	<b>156</b>
3.1.	Étude de la protéine de capsid du CaMV .....	156
3.2.	Étude de la protéine P2 exprimée dans <i>E.coli</i> .....	159
3.3.	Étude du complexe 157m: microtubules .....	162
<b>4.</b>	<b>Conclusion .....</b>	<b>165</b>
<b>Conclusion générale .....</b>		<b>169</b>
<b>Bibliographie .....</b>		<b>176</b>



# Table des figures

---

Figure 1 : Représentation schématique d'un microscope électronique à transmission. ....	21
Figure 2 : Les sources d'électrons.....	22
Figure 3 : Schéma de l'interaction d'un faisceau d'électrons sur un échantillon .....	23
Figure 4 : La lentille magnétique .....	24
Figure 5 : Théorie de la formation de l'image. ....	25
Figure 6 : Oscillation de la FTC en fonction d'une valeur de défocalisation donnée. ....	28
Figure 7 : Superposition de la FTC d'images à différentes valeurs de défocalisation. ....	28
Figure 8 : Superposition de la FTC d'images enregistrées à différentes valeurs de défocalisation. ....	29
Figure 9 : Principe de la coloration négative. ....	31
Figure 10 : Étapes de congélation d'une grille de cryo-microscopie électronique. ....	32
Figure 11 : Transfert de la grille congelée dans le microscope .....	33
Figure 12 : Observation du virus de la mosaïque du tabac par microscopie électronique à transmission. ....	34
Figure 13 : Illustration de l'inclinaison conique aléatoire et de la reconstruction inclinée orthogonale. ....	37
Figure 14 : Schéma d'une ligne commune (C) de deux projections d'une même particule dans l'espace de Fourier.....	38
Figure 15 : Principe de l'addition de sous-volumes acquis par cryo-TE .....	39
Figure 16 : Les différents systèmes de congélation actuels.....	42
Figure 17 : Mouvement de l'échantillon sous l'effet du faisceau d'électrons.....	43
Figure 18 : Effet des grilles en or de microscopie électronique .....	43
Figure 19 : Photo du robot Spotiton.....	44
Figure 20 : Formation des nanofibres après traitement à l'ammonium persulfate sur une grille de cuivre au cours du temps. ....	45
Figure 21 : Acquisitions réalisées sur le microscope Titan Krios.....	46
Figure 22 : Évolution des caméras équipant les microscopes électroniques à transmission .....	47
Figure 23 : Impact de l'affinement du détecteur sur la fonction de transfert modulaire. ....	49
Figure 24 : Les différents modes d'acquisition proposée par la caméra Gatan K2 Summit. ....	50
Figure 25 : Comparaison de la variation de DQE différent modes de détection. ....	50
Figure 26 : Comparaison de la variation de DQE des différentes caméras à détection directe. ....	51
Figure 27 : Effet de l'alignement des vues sur la qualité du signal. ....	52
Figure 28 : Dessin simplifié du mode de transmission circulant non propagatif.....	60
Figure 29 : Organisation génomique du FBNSV .....	61
Figure 30 : Réplication des virus à ADN simple brin selon le mécanisme du cercle roulant. ....	61
Figure 31 : Carte du plasmide pETM-41 utilisé pour la production des protéines recombinantes en <i>E. coli</i> ....	63
Figure 32 : Système de congélation Gatan CP3.....	67

Figure 33 : Expression de la protéine de fusion His-MBP-S produite chez la souche BL21 en milieu d'auto-induction .....	68
Figure 34 : Expression de la protéine de fusion His-GST-N produite chez la souche C41 en milieu d'auto-induction .....	69
Figure 35 : Composition des 12 tampons utilisés pour l'expression des segments viraux du FBNSV .....	70
Figure 36 : Expression de la protéine de fusion His-S produite chez les cellules <i>Sf9</i> .....	70
Figure 37 : Expression de la protéine de fusion His-N produite chez les cellules <i>Sf9</i> .....	71
Figure 38 : Images de microscopie électronique à transmission du virus FBNSV .....	71
Figure 39 : L'impact du virus de l'ArMV .....	75
Figure 40 : Structure de la protéine de capsid du GFLV obtenue par homologie de structure .....	76
Figure 41 : Représentation du potentiel électrostatique de la capsid du GFLV.....	77
Figure 42 : Interface du programme e2evalimage.py permettant de visualiser les films.....	79
Figure 43 : Interface du programme e2boxer.py. ....	80
Figure 44 : Représentation schématique de l'application du masque gaussien.....	84
Figure 45 : Évolution des classes 2D du virus de l'ArMV au fil des 30 itérations.....	86
Figure 46 : Visuel des classes 2D après 30 itérations .....	87
Figure 47 : Classification 3D du virus de l'ArMV à partir des 15889 particules issues de la classification 2D... ..	88
Figure 48 : Affinement automatique des 15350 particules issues de classes 3D 2 et 3. ....	89
Figure 49 : Carte de densité des particules contenues dans les classes 2 et 3, avant et après post-traitement d'image.....	90
Figure 50 : Carte de densité des classes 2 et 3 avant et après post-traitement d'image.....	91
Figure 51 : Impact du filtre de dose proposée par Unblur sur la résolution. ....	92
Figure 52 : Impact du nombre de vues alignés par Unblur sur la résolution.....	92
Figure 53 : Impact du traitement des vues par Relion sur la résolution. ....	92
Figure 54 : Reconstruction du virus de l'ArMV à 6,5 Å.....	94
Figure 55 : Reconstruction du virus de l'ArMV à 3,8 Å.....	95
Figure 56 : Figure : Estimation de la résolution locale par Resmap. Cette carte correspond à une carte de différence entre les deux demi-cartes. ....	96
Figure 57 : Diagramme récapitulatif du protocole utilisé pour l'affinement du modèle atomique. ....	97
Figure 58 : Affinement de corps rigides par « Jiggle Fit ».....	98
Figure 59 : Construction du modèle à partir d'un alignement de séquence. ....	99
Figure 60 : Graphique du nombre de facteurs de structure par rapport à la résolution. ....	100
Figure 61 : Variation de la taille du pixel et l'accord entre le modèle et la carte. ....	101
Figure 62 : Variation du poids des données expérimentales avec la géométrie du modèle.....	102
Figure 63 : Création de l'octamère .....	103
Figure 64 : Graphique de la fonction d'optimisation du modèle atomique dans la carte de l'ArMV par rapport au cycle d'affinement.....	104

Figure 65 : Tableau statistique de la validation de la géométrie du modèle de l'ArMV par Molprobit. ....	105
Figure 66 : Corrélation du modèle atomique avec les demi-cartes .....	106
Figure 67 : Modèle atomique de la protéine de capsid de l'ArMV .....	107
Figure 68 : Interaction ARN/protéine de capsid de l'ArMV .....	109
Figure 69 : Première comparaison des protéines de capsid de l'ArMV et du GFLV .....	110
Figure 70 : Diagramme schématique du génome du CaMV .....	142
Figure 71 : Cycle de multiplication du CaMV à l'intérieur d'une cellule de plante .....	144
Figure 72 : Séquestration des particules virales .....	145
Figure 73 : Le complexe transmissible P4:P3 .....	146
Figure 74 : Modèle d'acquisition du CaMV .....	147
Figure 75 : Visualisation de la carte de différence du CaMV obtenue par cryo-ME .....	148
Figure 76 : Modèle de l'interaction de la protéine P3 avec la capsid du CaMV .....	148
Figure 77 : Structure de la partie N-terminale de la protéine P3 obtenue par cristallographie aux rayons X	149
Figure 78 : Les différentes formes de la protéine P3 .....	150
Figure 79 : Observation des paracristsaux de P2 dans des cellules <i>Sf9</i> par cryo-ME en coloration négative ..	151
Figure 80 : Observation du mutant 157m décorant les microtubules de cellules <i>Sf9</i> par cryo-ME en coloration négative .....	152
Figure 81 : Prédiction de structure de la protéine P2 .....	152
Figure 82 : Reconstruction 3D de la capsid du CaMV .....	157
Figure 83 : Classification 2D et 3D de la capsid du CaMV .....	158
Figure 84 : Gels de polyacrylamide des étapes de purification de la protéine mutante P2C. Av=avant, Ap=après .....	161
Figure 85 : Classification 2D du complexe 157m:microtubule par Imagic .....	163
Figure 86 : Reconstruction 3D du complexe 157m:microtubule .....	164
Figure 87 : Figure : Carte de différence entre le microtubule décoré et le microtubule nu .....	165
Figure 88 : Étude des paracristsaux de P2 par cryo-tomographie électronique. ....	167
Figure 89 : Reconstructions 3D de l'ArMV obtenues par Relion (bleu) et Imagic (rose) .....	171



# Liste des abréviations

---

2D : 2-dimensions

3D : 3-dimensions

AD : domaine d'ancrage

ADN : Acide désoxyribonucléique

ARN : Acide ribonucléique

ArMV : Arabis Mosaic Virus ou virus de la mosaïque de l'arabette

BBSV : Broad Bean Stain Virus ou virus de la tâche de la fève

BBTV : Banana Bunchy Top Virus

CaMV : Cauliflower Mosaic Virus ou virus de la mosaïque du chou-fleur

CBS : Centre de Biochimie Structurale

Cc : aberration chromatique

CDD : Charge Coupled Device

CIRAD : Centre de coopération Internationale en Recherche Agronomique pour le Développement

CMOS : Complementary Meta Oxide Semiconductor

CPC : Cystéine-Proline-Cystéine

CPMV : Cowpea Mosaic Virus ou virus de la mosaïque du niébé

Cryo-ME : Cryo-Microscopie Electronique

Cryo-TE : Cryo-Tomographie Electronique

Cs : aberration sphérique

DQE : Detective Quantum Efficiency ou efficacité quantique de détection

EMBL : European Molecular Biology Laboratory ou Laboratoire de Biologie Moléculaire Européen

*E. coli* : *Escherichia coli*

EMDB : Electron Microscopy Data Bank ou banque de données de microscopie électronique

EMPIAR : Electron Microscopy Pilot Image Archive ou Archive des images pilotes de microscopie électronique

FBNSV : Faba Bean Necrotic Stunt Virus ou virus du retard de croissance et de la nécrose de la fève

FEG : Field Emission Gun ou canon à émission de champ

FSC : Fourier Shell Correlation

FTC : Fonction de Transfert de Contraste

GDH : Glutamate déhydrogénase

GFLV : Grapevine Fanleaf Virus

GST : Glutathione S-Transferase

His : Histidine

IBS : Institut de Biologie Structurale  
IDH : Isocitrate déshydrogénase  
IGBMC : Institut de Génétique et de Biologie Moléculaire et Cellulaire  
IPTG : Isopropyl  $\beta$ -D-1-Thiogalactopyranoside  
LB : Lysogeny Broth  
MBP : Maltose Binding Protein  
MET : Microscopie Electronique à Transmission  
MP : Protéine de mouvement  
MRA : Multi-Reference Alignment ou alignement contre de multiples références  
MSA : Multivariate Statistical Analysis  
MTF : Fonction de transfert modulaire  
NCS : Non-Crystallographic Symmetry  
NeCEN : Netherlands Centre for Electron Nanoscopy  
ORF : Open Reading frame ou cadres de lectures ouverts  
PDB : Protein Data Bank ou banque de données de protéines  
PDBe : Protein Data Bank in Europe ou banque de données de protéines européenne  
PES : Polyethersulfone  
PSF : Point Spread Function  
RMN : Résonance Magnétique Nucléaire  
*Sf9* : *Spodoptera frugiperda*  
TEV : Tobacco Etch Virus ou protéase du virus de la gravure du tabac  
TF : Transformée de Fourier  
TLS : Translation, Libration and Screw  
TRPV1 : Transient Receptor Potential Channel sub family V member 1  
Trx : Thioredoxin  
*X. diversicaudatum* : *Xiphinema diversicaudatum*  
*X. index* : *Xiphinema index*





# Avant-propos

---

Le Gouvernement de la Nouvelle-Calédonie m'a attribué une « Bourse d'Encouragement à la Recherche Universitaire » destinée aux jeunes étudiants calédoniens qui souhaitent réaliser des travaux de recherche sur un sujet présentant un intérêt pour la Nouvelle-Calédonie. Nous avons proposé initialement d'effectuer une étude structurale d'un virus de la famille des *Nanoviridae*, qui infecte les plantes à fleurs et utilise le puceron comme vecteur de transmission. En effet, l'île est gravement touchée par un virus appartenant aux *Nanoviridae*, le virus du sommet touffu du bananier (Banana Bunchy Top Virus) qui provoque l'une des maladies virales les plus graves affectant les bananiers, et contre laquelle les seuls moyens de luttés sont l'utilisation de pesticides et l'éradication de la plante infectée et de toutes ses voisines. Dans ce contexte, nous nous sommes focalisés sur l'étude structurale d'un virus modèle de la famille des *Nanoviridae* : le virus du retard de croissance et de la nécrose de la fève (Faba Bean Necrotic Stunt Virus ou FBNSV), afin de mieux comprendre le mode de transmission plante-vecteur de cette famille de virus. Cette thèse s'est déroulée dans l'équipe du Dr Patrick Bron, au Centre de Biochimie Structurale, qui utilise des approches intégrées de biologie structurale pour caractériser la structure et le mode de fonctionnement de différents virus humains et de plantes. De plus, cette étude a été réalisée en collaboration avec l'équipe du Dr Stéphane Blanc du laboratoire de Biologie et Génétique des Interactions Plante-Parasite, qui s'intéresse à l'impact des mécanismes de la transmission par vecteur sur l'évolution des virus de plantes, et notamment des *Nanoviridae*. L'étude structurale d'un virus de la famille des *Nanoviridae* étant un sujet d'étude difficile, très peu documenté, et donc hautement risqué, le cadre de la thèse a été élargi à l'étude structurale de la transmission du virus de la mosaïque du chou-fleur (Cauliflower Mosaic Virus ou CaMV), virus modèle permettant d'étudier la transmission non circulante des virus par les pucerons. Ainsi, l'étude du virus du retard de croissance et de la nécrose de la fève permet d'aborder le mode de transmission circulant où le virus traverse l'appareil intestinal du puceron, tandis que l'étude du virus de la mosaïque du chou-fleur permet d'aborder le mode de transmission non circulant où le virus reste localisé dans les pièces buccales du puceron.

Par conséquent, la thèse a porté initialement sur l'étude structurale de deux virus de plantes, le FBNSV et le CaMV ainsi que sur la caractérisation structurale des protéines virales impliquées dans leur transmission par les pucerons, et ce, par des approches de cryo-microscopie électronique à transmission et de traitement d'image.

La cryo-microscopie électronique à transmission vit une véritable révolution depuis ces cinq dernières années, puisque la résolution qu'il est possible d'atteindre après traitement d'image est

maintenant comparable à celle obtenue par cristallographie aux rayons X ou par résonance magnétique nucléaire. La cryo-microscopie électronique devient une technique à part entière de la biologie structurale. Mais cette révolution de la « résolution » s'accompagne d'un bouleversement dans la manière d'acquérir et de traiter les images de microscopie électronique. De même, l'interprétation des cartes de densité issues de la cryo-microscopie et la construction de modèles atomiques à partir de ces cartes nécessitent de définir de nouvelles stratégies d'étude.

# CHAPITRE 1

# Chapitre 1 : Introduction générale

---

## 1. Introduction à la microscopie électronique à transmission

La microscopie électronique est une technique d'imagerie utilisant un faisceau d'électrons accélérés sous une différence de potentiel donnée pour produire une image agrandie de la matière. La prouesse technologique qui m'a permis de présenter cette thèse aujourd'hui est le résultat d'une aventure scientifique qui a débuté au XX<sup>e</sup> siècle.

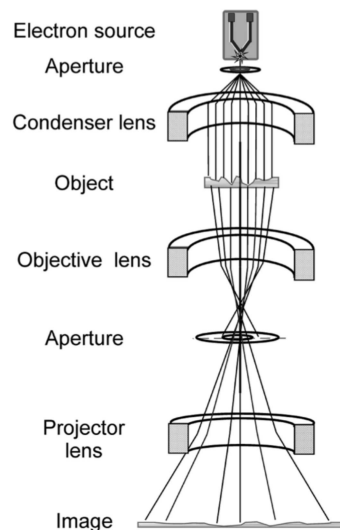
La découverte des propriétés des électrons est attribuée à J.J. Thomson qui détermina le rapport entre la charge et la masse de ces charges élémentaires d'électricité négative, et Millikan, qui en 1909 mesura leur charge individuelle égale à  $1,6 \cdot 10^{-19} \text{C}$ . Plus tard, les travaux menés par Louis de Broglie sur l'aspect ondulatoire des électrons, et par Hans Busch sur l'optique électronique qui démontra le potentiel de focalisation d'un faisceau d'électrons d'une bobine, ont permis à Max Knoll, un ingénieur en électricité allemand, et son étudiant Ernst Ruska de construire le premier microscope électronique. Les premières images furent dévoilées le 4 juin 1931 à l'École Technique de Berlin, avec une résolution de quelques dizaines de nanomètres obtenues à l'aide d'un microscope à deux lentilles fonctionnant sous une tension de quelques milliers de volts. On assiste alors aux prémices de la microscopie électronique telle qu'on la connaît aujourd'hui. Pour des raisons liées à la Seconde Guerre mondiale, Ernst Ruska n'obtiendra le prix de Nobel de Physique qu'en 1986 pour ses travaux.

L'effervescence provoquée en Allemagne va alors se répandre en Europe et outre-Atlantique. À Berlin, un groupe de recherche dirigé par Ernst Ruska est créé par la société Siemens. Celui-ci, avec l'aide de Bodo von Borries, un électrotechnicien allemand, a construit le premier « ultramicroscope » commercial. Les techniques de préparation d'échantillon se développent alors, et des premières images de bactéries, virus, particules de carbone et diatomées sont réalisées. Un développement industriel important se met alors en place à la fin de la 2<sup>de</sup> guerre mondiale, à la fois en Europe avec les entreprises Zeiss et Philips, mais également au Japon avec JEOL et Hitachi. Jusqu'à aujourd'hui, les progrès techniques ont continué, et plusieurs compagnies sont maintenant implantées sur le marché, comme Zeiss et Leica en Europe, JEOL au Japon et FEI (anciennement Philips) aux États-Unis. En moins d'un siècle, la microscopie électronique à transmission (MET) a évolué, permettant de passer

d'une résolution de quelques dizaines de nanomètres à ses débuts, à une résolution atomique aujourd'hui.

## 1.1. Anatomie d'un microscope électronique à transmission

L'utilisation de la microscopie électronique à transmission en biologie permet l'étude d'un échantillon biologique par irradiation avec un faisceau d'électrons cohérent. L'anatomie d'un microscope électronique à transmission est comparable à celui d'un microscope optique, avec une source d'électrons qui remplace la lumière, des lentilles magnétiques qui vont dévier les électrons, et des diaphragmes destinés à bloquer les électrons diffusant aux très grands angles (**Figure 1**).

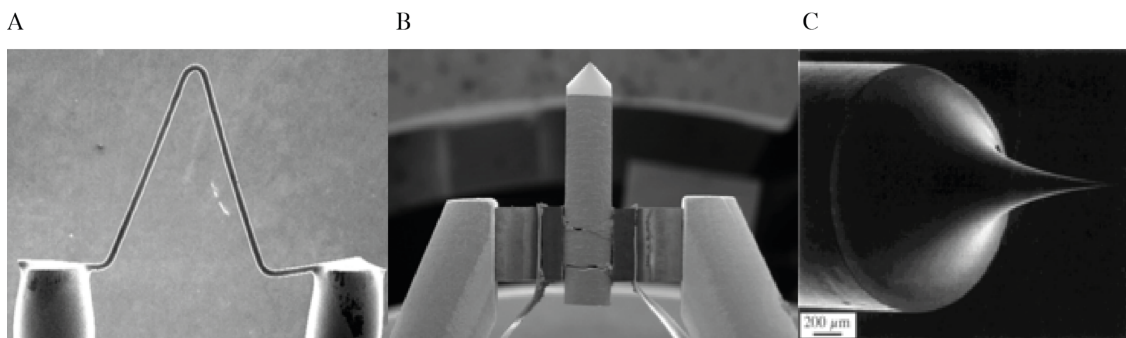


**Figure 1** : Représentation schématique d'un microscope électronique à transmission. Reproduit de (Orlova & Saibil 2011).

### 1.1.1. Les sources d'électrons

Dans la microscopie électronique à transmission, la source d'électrons est obtenue à partir d'un métal à l'intérieur duquel circule un courant électrique. Ce courant va provoquer des collisions entre électrons voisins. Si un électron possède assez d'énergie, il peut être extrait sous l'action d'un potentiel électrique dont la valeur dépend du métal utilisé. Historiquement, la 1<sup>re</sup> source d'électrons utilisée était un filament de tungstène en forme de V (**Figure 2A**), les électrons étaient émis par la pointe du V lorsque le filament était chauffé à plus de 2500°C. Par la suite, la cathode LaB<sub>6</sub> (**Figure**

**2B)** qui est un cristal d'hexaborure de lanthane a remplacé le filament de tungstène, possédant une brillance supérieure, une zone d'émission des électrons réduite, et ne devait être chauffée qu'à 1500°C. Il existe une 3<sup>e</sup> source d'électrons : le canon à émission de champ (ou FEG) (**Figure 2C**) qui consiste en un cristal de tungstène chauffé à plus de 2000°C qui va émettre des électrons par une pointe aiguisée. Par rapport aux précédentes méthodes, l'utilisation d'un canon de type FEG permet une meilleure cohérence spatiale, une meilleure cohérence temporelle et une meilleure brillance du faisceau d'électrons émis (Spence, 2009).



**Figure 2 : Les sources d'électrons. (A)** Filament de tungstène en V. **(B)** Cristal LaB6. **(C)** Canon à émission de champ (FEG).

Les électrons émis sont accélérés sous un vide poussé par une tension habituellement comprise entre 100 et 300 kV et traversent ensuite un système de lentilles électromagnétiques avant de former l'image.

## 1.1.2. La formation de l'image

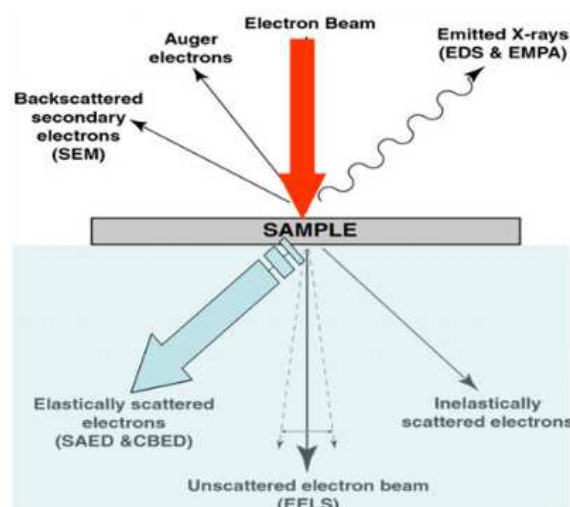
### 1.1.2.1. L'origine du contraste de l'image

Les électrons peuvent être considérés comme des particules chargées négativement pouvant interagir avec d'autres particules, et être diffusées. Dans un microscope électronique à transmission, on considère un faisceau d'électrons comme une somme d'électrons possédant une longueur d'onde et une phase, avançant parallèlement par rapport à l'axe optique. Le faisceau d'électrons est donc considéré comme une vague plane qui vient traverser l'échantillon. À la suite du contact entre le faisceau d'électrons et l'échantillon, différents phénomènes physiques se produisent. Dans le cas de

la microscopie électronique à transmission, trois phénomènes sont responsables de la formation de l'image. Ces trois phénomènes sont schématisés dans la zone bleue de la **Figure 3** :

- Phénomène 1 : l'électron n'interagit avec aucun atome de l'échantillon, il ne sera pas dévié de sa trajectoire et ne subira pas de perte d'énergie.
- Phénomène 2 : l'électron subit une diffusion élastique, il interagit avec un atome de l'échantillon et dévie de sa trajectoire, sans perte d'énergie.
- Phénomène 3 : l'électron subit une diffusion inélastique, il interagit avec un atome de l'échantillon, dévie de sa trajectoire, mais perd de l'énergie. L'énergie perdue va notamment induire la formation de radicaux libres conduisant à la dégradation de l'échantillon.

Dans le cas d'un échantillon de 500 Å d'épaisseur et composé principalement d'atomes de carbone, ce qui est le cas pour un échantillon biologique, sous un faisceau d'électrons ayant une tension accélératrice de 300 kV, la probabilité qu'un électron n'interagisse pas avec l'atome de carbone est de 80 %, celle d'avoir une diffusion élastique est de 5 % et celle d'avoir une diffusion inélastique est de 15 %.



**Figure 3** : Schéma de l'interaction d'un faisceau d'électrons sur un échantillon. Reproduit de *ammrf.org.au*.

Les interactions diverses des électrons avec l'échantillon sont à l'origine du contraste des images, on distingue :

- Le contraste d'amplitude, résultant de l'absorption d'une partie du faisceau dans l'échantillon. Dans le cas d'un échantillon biologique composé d'atomes légers (H, O, N et C),



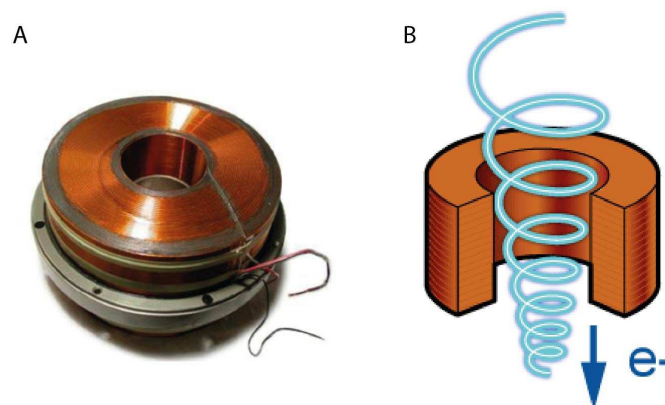
le contraste d'amplitude est très faible. Il est augmenté avec le diaphragme objectif qui stoppe les électrons ayant diffusé aux grands angles.

- Le contraste de phase résultant du décalage de phases entre les électrons non déviés et les électrons déviés. Ce décalage dépend de la longueur du trajet des électrons, qui est modulé par la défocalisation de la lentille objective.

Les électrons inélastiques ayant perdu de l'énergie, seront focalisés dans un plan différent de celui des électrons élastiques, et vont donc diminuer le contraste des images.

### 1.1.2.2. Les lentilles électromagnétiques

La formation de l'image en microscopie électronique à transmission reste similaire à celle d'un système optique, mais elle intègre des particularités qu'il est important de noter afin de mieux comprendre les limites de cette imagerie. Dans un microscope électronique à transmission, les lentilles sont appelées « lentilles électromagnétiques », et correspondent à des bobines de cuivre dans lesquelles passe un courant, générant un champ magnétique qui va permettre de modifier la trajectoire du faisceau d'électrons (**Figure 4**). Il est donc possible de réaliser des combinaisons de champs exerçant sur les électrons une action identique à celle des lentilles de verre sur les rayons lumineux.



**Figure 4 : La lentille magnétique.** (A) Photo d'une lentille magnétique de microscope électronique à transmission. (B) Schéma représentatif du trajet des électrons dans une lentille magnétique de microscope électronique à transmission. Reproduit de ([ammrf.org.au](http://ammrf.org.au)).

### 1.1.2.3. Relation entre diffraction et image

La formation de l'image d'un objet dépend donc de la diffusion des électrons sur l'échantillon et de la capacité des lentilles électromagnétiques à dévier ces électrons. La lentille la plus importante est la lentille objective qui va faire converger le faisceau d'électrons après sa diffusion sur l'échantillon. La formation de l'image peut se décomposer en deux étapes (**Figure 5**) :

- Les électrons sont tout d'abord rassemblés au niveau du plan focal de la lentille, formant le spectre de diffraction de l'objet (pouvant être décrit mathématiquement par la transformée de Fourier de l'objet).
- Ensuite, les électrons sont recombinaés pour former une image agrandie de l'objet. Le passage du spectre de diffraction à l'image peut être décrit mathématiquement par la transformée de Fourier inverse de l'objet.

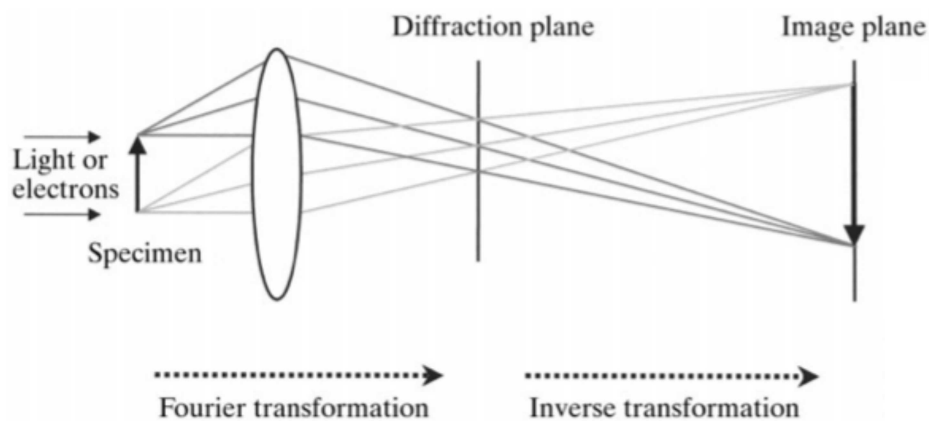


Figure 5 : Théorie de la formation de l'image. Reproduit de (Saibil 2000).

## 1.1.3. Les limites du système d'imagerie

### 1.1.3.1. Résolution maximale théorique

La résolution maximale d'un microscope peut être résumée comme la distance minimale permettant de distinguer deux points. La résolution maximale théorique qu'il est possible d'atteindre est dépendante de la tension accélératrice du microscope, de la qualité du système d'imagerie, et de la caméra installée. Tout d'abord, la résolution dépend de la longueur d'onde  $\lambda$ , de l'indice de réfraction du milieu traversé  $n$  et de l'angle d'incidence du faisceau  $\alpha$ , décrit par le critère de Rayleigh (**équation 1**), qui peut être simplifiée à l'**équation 2** :

$$d = \frac{0,61 \times \lambda}{n \times \sin \alpha} \quad (1)$$

$$d = \frac{\lambda}{2} \quad (2)$$

Dans le cas d'un microscope électronique, la longueur d'onde  $\lambda$  est dépendante de la tension accélératrice  $V$  du microscope telle que décrit dans l'équation 3 :

$$\lambda (\text{Å}) = \frac{12,3}{\sqrt{V}} \quad (3)$$

Ainsi, un microscope avec une tension accélératrice de 300 kV aura une longueur d'onde de  $\lambda = 0,022$  Å, donc une résolution théorique de 0,011 Å.

Puisque l'on parle de résolution, il faut également considérer la résolution maximale qu'il sera possible d'atteindre à partir des images enregistrées. Cette résolution sera liée à la taille du pixel du détecteur installé dans le microscope. Cette limitation est expliquée par le théorème d'échantillonnage (ou théorème de Nyquist-Shannon), qui énonce que l'échantillonnage d'un signal exige une fréquence d'échantillonnage supérieure au double de la fréquence maximale. En microscopie électronique, ce théorème signifie que la résolution maximale théorique sera égale au double de (taille du pixel du détecteur) / (facteur de grossissement du microscope). Par exemple, pour un grossissement effectif de 65 000 X au niveau de la caméra, et sachant que la caméra à un capteur dont l'unité CCD est de 15  $\mu\text{m}$ , le théorème de Shannon nous dit que la résolution maximale qu'il sera possible d'atteindre sera de  $(2 \times (150\,000 \text{ Å} / 65\,000))$  soit 4.6 Å.

### 1.1.3.2. Les aberrations dues aux lentilles

Les lentilles électromagnétiques présentent plusieurs types d'aberrations qui provoquent une diminution significative de la résolution :

- L'aberration sphérique ( $C_s$ ) : la déviation des électrons sera plus forte pour les électrons éloignés de l'axe optique, par rapport aux électrons proches de l'axe optique. L'image d'un point ne sera pas un point, mais un disque dit disque de confusion (le rayon de ce disque est proportionnel au coefficient d'aberration sphérique donné par le constructeur) ;

- L'aberration chromatique (Cc) : la déviation des électrons dépend de la longueur d'onde du rayon incident, si les électrons n'ont pas tous la même énergie, ils ne convergeront pas au même point, produisant également un disque de confusion ;
- L'astigmatisme : le grossissement n'est pas le même dans toutes les directions.

### 1.1.3.3. La fonction de transfert de contraste (FTC)

La transmission de l'information par le microscope électronique à transmission n'est pas parfaite, le microscope introduit des modifications dans l'information qu'il est nécessaire de corriger. Ces modifications sont appelées fonction de transfert de l'instrument. La notion de fonction de transfert de contraste (FTC) a été introduite pour modéliser mathématiquement les effets du microscope électronique sur le contraste de l'image. Ainsi, chaque projection bidimensionnelle d'un objet sera affectée par la FTC du microscope, de telle manière que la transformée de Fourier d'une image (TF(Image)) correspond à la transformée de Fourier de l'objet (TF(Objet)) multipliée par la FTC (équation 4).

$$TF(Image) = TF(Objet) \times FTC \quad (4)$$

La définition mathématique de la FTC est décrite dans l'équation 5 :

$$FTC = -2\sin\left[\pi\left(\Delta z\lambda q^2 - \frac{C_s\lambda^3 q^4}{2}\right)\right] \quad (5)$$

$C_s$  – coefficient d'aberration sphérique

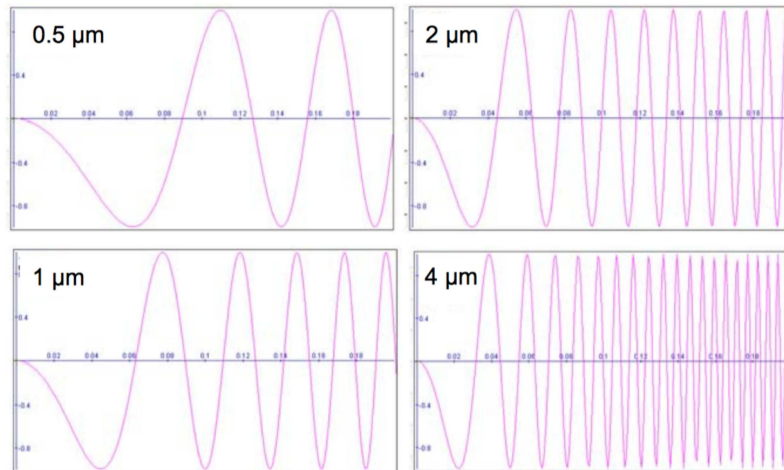
$\Delta z$  – valeur de la défocalisation

$q$  – fréquence spatiale (1/résolution(Å))

$\lambda$  – la longueur d'onde des électrons

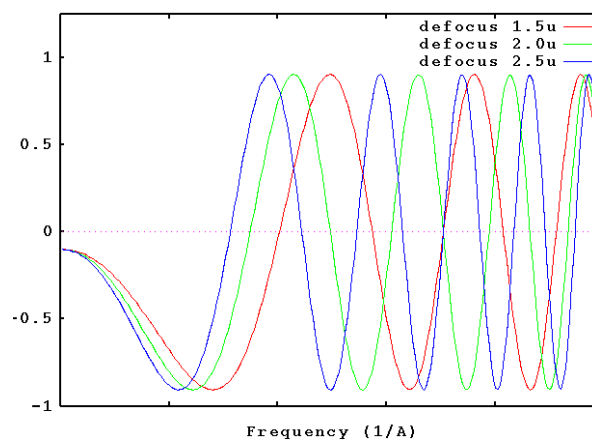
La fonction de transfert de contraste est une fonction avec une amplitude oscillante entre 1 et -1, correspondant à une inversion de contraste provoquée par le microscope. Un changement de la

valeur de défocalisation de la lentille objective induit une oscillation différente. Une forte défocalisation provoquera une oscillation plus importante, *a contrario* une faible défocalisation provoquera une oscillation moins importante (**Figure 6**).



**Figure 6 : Oscillation de la FTC en fonction d'une valeur de défocalisation donnée.** Modélisations réalisées avec le logiciel FTC simulation (Jiang & Chiu 2001).

Une défocalisation forte entraînera un signal plus fort dans les fréquences spatiales basses, mais l'oscillation traversera plus de zéros. Lorsque l'amplitude est à zéro, le signal est perdu, car il n'y a pas de différence d'intensité entre les électrons déviés et non déviés. De plus, si l'amplitude est négative, le contraste de l'image est inversé : ce qui est blanc devient noir et ce qui est noir devient blanc. Ainsi, il est à la fois nécessaire de corriger l'inversion de contraste et de cumuler des images à différentes valeurs de défocalisation pour retrouver du signal à chaque fréquence spatiale (**Figure 7**).



**Figure 7 : Superposition de la FTC d'images à différentes valeurs de défocalisation.**

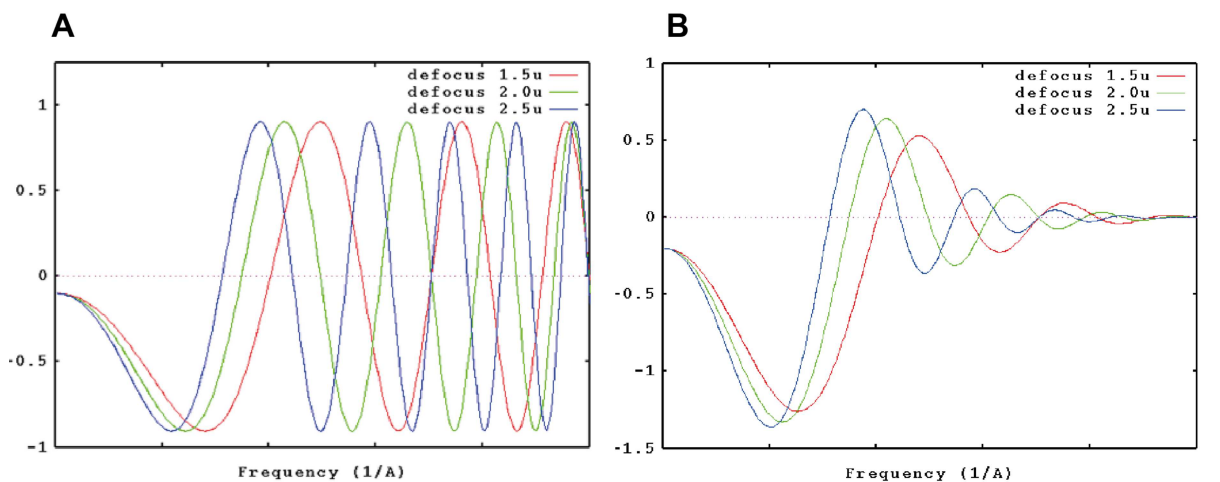
### 1.1.3.4. La fonction d'enveloppe

L'ensemble des aberrations de lentilles, combinées avec d'autres facteurs tels que la stabilité du microscope (vibration et fluctuation de la tension accélératrice), va affecter la FTC et induire une décroissance du signal en fonction des fréquences spatiales. Ce phénomène, appelé fonction d'enveloppe  $E_{(k)}$ , est traduit mathématiquement par l'équation 6 :

$$E_{(k)} = e^{[-\pi^2 Q^2 (k^2 C_s \lambda^3 - \Delta Z k \lambda)^2]} \times e^{[k^4 \lambda^2 D z^2 - \frac{\pi^2}{16 \ln(2)} \times \sqrt{(C_s \lambda)}]} \times e^{[-(\frac{k}{E_g})^2]} \quad (6)$$

La fonction d'enveloppe correspond à la multiplication de trois exponentielles. La première exponentielle permet de décrire la taille de la source (Q), le deuxième terme tient compte de l'instabilité de la défocalisation Ds (« defocus spread ») et le troisième terme rend compte des variations de tous les autres paramètres, avec  $E_g$  qui est la largeur à mi-hauteur de la fonction de type Gaussien.

Ainsi, cette fonction d'enveloppe va grandement atténuée dans les hautes fréquences spatiales contenant les informations à haute résolution (**Figure 8**).



**Figure 8 : Superposition de la FTC d'images enregistrées à différentes valeurs de défocalisation. (A) FTC non atténuée par une fonction d'enveloppe. (B) FTC atténuée par une fonction d'enveloppe.**

## 1.2. L'étude d'un échantillon biologique

L'utilisation de la microscopie électronique à transmission nous permet de caractériser la structure d'un échantillon biologique. L'utilisation d'une telle technique soulève trois questions fondamentales qui seront traitées au cours de ce sous-chapitre :

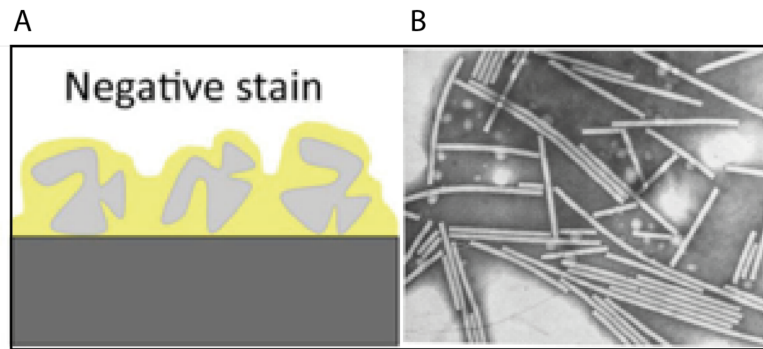
- Comment préserver la forme native de la protéine d'intérêt ?
- Comment obtenir des informations en 3-dimensions (3D) à partir d'images en 2-dimensions (2D) ?
- Quelles sont les limites de la MET sur l'étude d'échantillons biologiques ?

Deux techniques majeures sont utilisées actuellement pour l'observation des échantillons biologiques : la coloration négative et la cryo-microscopie électronique.

### 1.2.1. La coloration négative

La coloration négative est la méthode la plus simple pour l'observation rapide de particules isolées. Elle a été établie en 1959 par Brenner et Horne qui utilisaient de l'acide phosphotungstique comme agent colorant (Brenner & Horne 1959). Leur méthode a d'abord été utilisée sur les virus de la mosaïque du tabac et de la mosaïque jaune du navet, puis rapidement utilisée pour l'étude d'autres virus de plantes, d'animaux ou de bactéries. Cette technique consiste à enrober l'échantillon dans une solution de sels d'atomes lourds comme l'uranium ou le molybdène, qui permettront de préserver sa structure tout en augmentant fortement son contraste.

Une goutte des particules biologiques en solution est déposée sur une grille de microscopie électronique, généralement en cuivre, et recouverte d'un film de carbone. La solution est ensuite éliminée par absorption sur papier filtre puis remplacée par une solution saline de métaux lourds, habituellement de l'acétate d'uranyle ou du molybdate d'ammonium. La solution de métaux lourds est ensuite éliminée par absorption, puis la grille est séchée à température ambiante. De cette manière, les particules biologiques ne sont plus en solution, mais maintenant recouvertes d'un film fin de métaux lourds. On n'observe pas directement l'échantillon, mais la répartition du contrastant autour et sur celui-ci, qui se localise préférentiellement dans les zones hydrophiles accessibles (**Figure 9**).



**Figure 9 : Principe de la coloration négative.** (A) Schéma de l'échantillon biologique sur film de carbone, séché dans une solution de métaux lourds (Durand et al. 2013). (B) Coloration négative appliquée au virus de la mosaïque du tabac piégé dans de l'acide phosphotungstique (x 59 000) (Brenner & Horne 1959).

La coloration négative présente ainsi deux atouts majeurs :

- L'échantillon étant piégé dans la solution de métaux lourds, il n'y a plus de molécules d'eau et le vide du microscope affecte moins l'intégrité de l'échantillon ;
- Les métaux générant un très haut contraste d'amplitude, permettent une meilleure visualisation de l'échantillon ;

La coloration négative présente néanmoins de nombreuses limites. En effet, des artefacts peuvent apparaître, à cause d'une coloration partielle de l'échantillon ou de l'effondrement de celui-ci quand la grille est séchée. De par la taille des agrégats de métaux lourds qui se forment autour de l'échantillon, la résolution est typiquement limitée à 15 Å. Bien que limitées, des reconstructions tridimensionnelles de l'échantillon observé peuvent être réalisées à partir d'images prises en coloration négative. De nos jours, la coloration négative est maintenant principalement utilisée pour contrôler la qualité d'un échantillon, ou pour générer un premier modèle 3D de l'échantillon par les techniques de « Random conical tilt » ou « Orthogonal tilt reconstruction » que nous verrons par la suite.

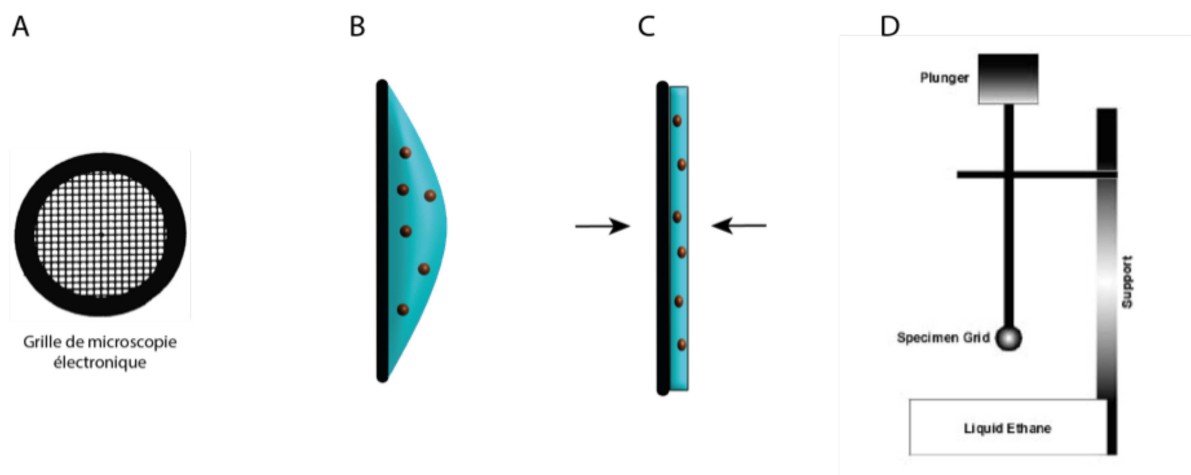
## 1.2.2. La cryo-microscopie électronique : la porte vers la haute résolution

La cryo-microscopie électronique (cryo-ME) consiste à piéger l'échantillon biologique dans une fine couche de glace amorphe, dans le but de préserver l'état natif de l'échantillon biologique. Cette technique a été développée par Jacques Dubochet et ses collègues Marc Adrian et Jean Lepault, à



l'EMBL de Heidelberg (Allemagne) au début des années 1980 (Adrian et al. 1984). Pour cette approche, des grilles à trous sont utilisées. Il s'agit de grilles recouvertes d'un film de carbone qui présentent des trous qui peuvent être réguliers ou non.

D'un point de vue pratique, l'échantillon est déposé sur une grille, puis l'excès de solution est éliminé par adsorption sur papier filtre, afin de ne laisser qu'une fine couche de solution dans les trous du film de carbone. La grille est ensuite immédiatement plongée dans un gaz liquéfié (éthane, propane, ou un mélange des deux) à  $-172^{\circ}\text{C}$  qui est lui-même refroidi par de l'azote liquide ( $-196^{\circ}\text{C}$ ) (**Figure 10**). La vitesse de congélation et le choix du cryogène sont primordiaux pendant l'étape de congélation, afin d'obtenir une congélation suffisamment rapide pour obtenir une glace amorphe dans les trous, c'est à dire une glace dans laquelle les molécules d'eau se figent et ne peuvent donc pas adopter une organisation cristalline. On parle alors d'échantillons hydratés-congelés. L'échantillon n'est pas plongé directement dans l'azote liquide, car il y aurait apparition d'une couche isolante de gaz qui diminuera la conductivité thermique et entraînera une cristallisation de la glace.



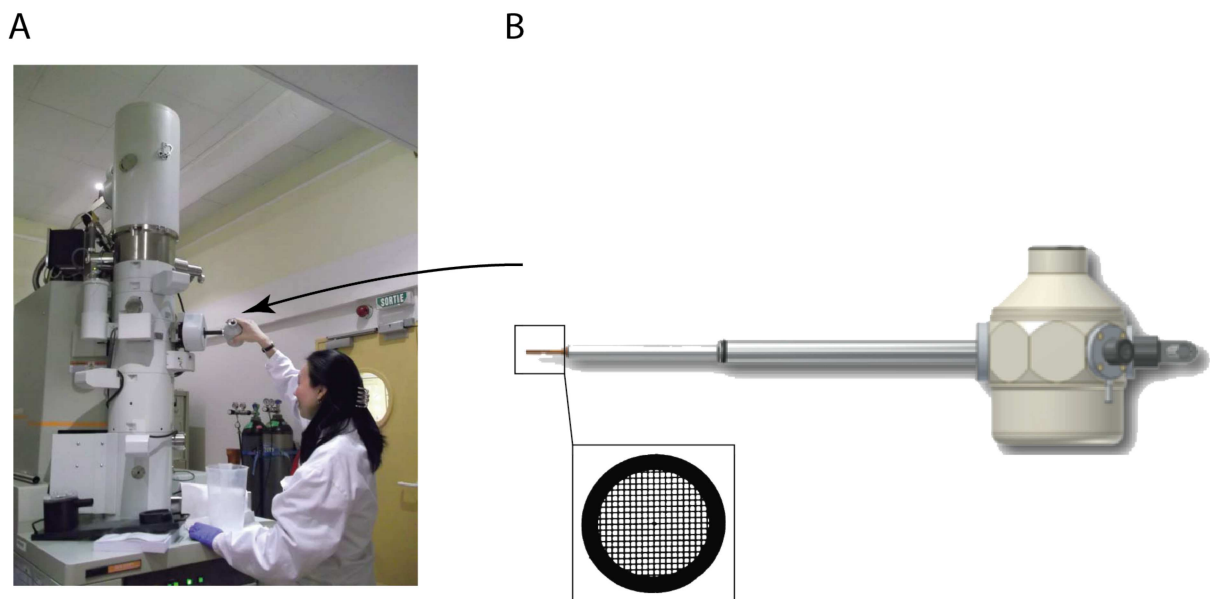
**Figure 10 : Étapes de congélation d'une grille de cryo-microscopie électronique.** (A) Schéma d'une grille de microscopie électronique. (B) Dépôt de l'échantillon sur la grille. (C) Retrait de l'excès de solution par absorption à l'aide de papiers Whatman® appliqués des deux côtés de la grille. (D) Schéma d'une guillotine qui va plonger la grille dans l'éthane liquide.

Cette méthode de congélation permet non seulement la conservation de l'échantillon dans son état natif, mais permet également une protection contre le vide du microscope.

Trois facteurs sont cruciaux pendant l'étape de congélation :

- L'épaisseur de la glace : selon l'échantillon, des optimisations peuvent être nécessaires concernant le temps de blot durant la congélation. En effet, une glace trop épaisse diminuera le contraste de l'image tandis qu'une glace trop fine pourra entraîner une orientation préférentielle des particules ;
- La solution dans laquelle est piégée la particule doit respecter certaines conditions : le contraste étant très faible, le glycérol, le sucrose et les détergents sont déconseillés, car ils réduisent le contraste ;
- Une concentration de protéines généralement 10 fois supérieure à celle utilisée pour la coloration négative est nécessaire afin d'avoir une densité homogène de particules dans la glace.

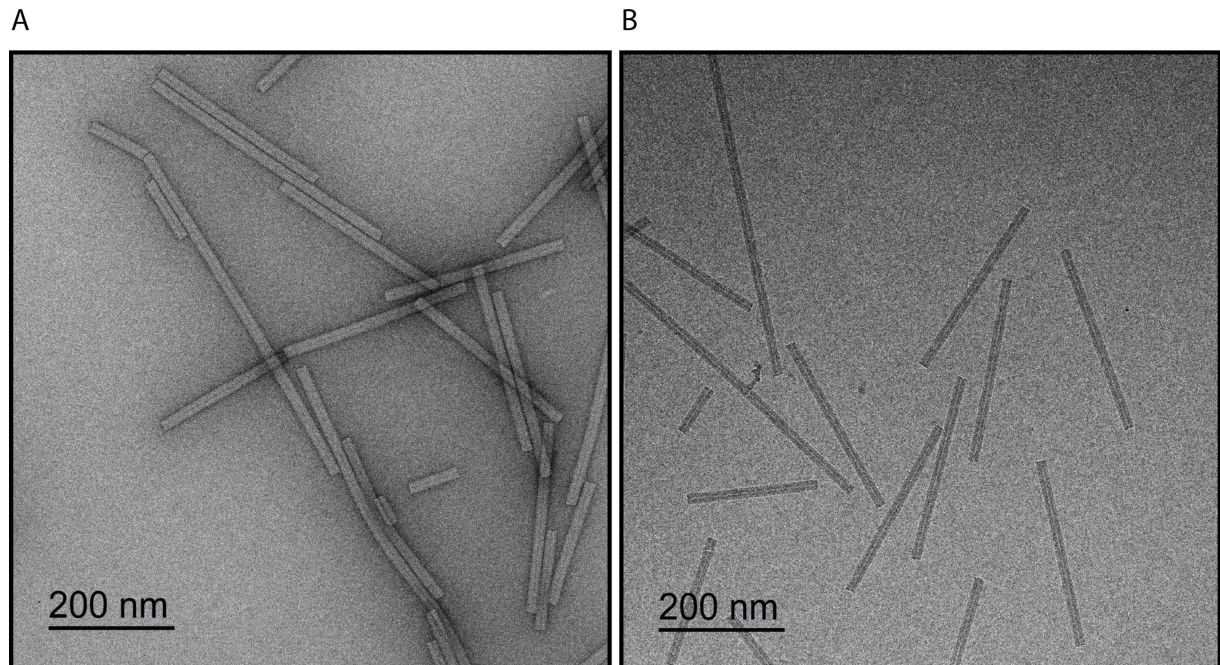
Après l'étape de congélation, les grilles sont maintenues à la température de l'azote liquide jusqu'à l'observation. Au sein de notre laboratoire, un porte-objet Gatan 626 est utilisé pour l'observation des échantillons hydratés-congelés (**Figure 11**).



**Figure 11 : Transfert de la grille congelée dans le microscope. (A)** Insertion du porte-objet refroidi à l'azote liquide dans le microscope JEOL 2200FS. **(B)** Porte-objet refroidi à l'azote liquide avec une grille positionnée à l'extrémité et refroidi par conductivité.

Outre une préparation d'échantillon qui nécessite plus de temps et d'optimisation, la cryo-ME rend possible l'étude d'une particule dans son état natif. Au prix d'un contraste beaucoup plus faible, elle

permet de s'affranchir de la limite de résolution imposée par la coloration négative, et ouvre la porte vers la haute résolution. L'absence de coloration diminue alors grandement le contraste de l'image (**Figure 12**).



**Figure 12 : Observation du virus de la mosaïque du tabac par microscopie électronique à transmission. (A)** Préparation de l'échantillon en coloration négative. **(B)** Virus de la mosaïque du tabac observé après congélation.

### **1.2.3. De l'image à la structure 3D de l'objet**

Les images enregistrées en MET correspondent à des projections 2D de l'échantillon observé. Ainsi un jeu de donnée en MET correspond à des projections très bruitées d'un objet prises dans des orientations variées. Les principes de la reconstruction et de l'obtention d'un modèle de référence sont abordés ici, mais les détails des étapes utilisées pendant la reconstruction d'une structure 3D seront décrits dans le sous-chapitre « Matériels et méthodes » du Chapitre 3.

#### **1.2.3.1. Les conditions d'acquisition : le mode « faible-dose »**

Les échantillons biologiques sont très sensibles aux dommages d'irradiations causés par les électrons. Une exposition longue aux électrons induit la formation de radicaux libres qui détruisent

l'échantillon, les acquisitions doivent alors s'effectuer en mode « faible-dose » qui permet de prendre des photos en contrôlant et limitant la dose totale d'électrons que subit l'échantillon. Une séquence spécifique, composée de 3 différents modes, est utilisée pour ne pas irradier la zone d'intérêt avant sa prise de vue :

- Le mode « recherche » : Ce mode permet de parcourir la grille et de repérer notre zone d'intérêt à photographier. Sur notre microscope JEOL 2200FS, ce mode est réglé de la manière suivante : à x 50000 de grossissement, le faisceau est étalé jusqu'à son maximum pour diminuer la dose d'électrons. Le mode diffraction est alors activé, une zone plus large de la grille est éclairée, équivalent à un grossissement x 3500 mais avec une dose d'électrons bien plus faible ;
- Le mode « focalisation » : une fois la zone d'intérêt trouvée, une région voisine va être utilisée pour faire la mise au point du faisceau, afin de ne pas irradier la zone d'intérêt avant sa prise de vue. La zone de focalisation est située à +/- 2  $\mu\text{m}$  de la zone d'intérêt et le microscope est réglé à un grossissement x 20000 ;
- Le mode « photo » : la zone est finalement photographiée au grossissement et une dose d'électrons est choisie (typiquement entre 10 et 20  $\text{e}^-/\text{\AA}^2$ ).

La faible dose d'électrons utilisée devient la cause du faible ratio signal/bruit présent dans les images.

### **1.2.3.2. Principe de la reconstruction 3D en microscopie électronique**

Lorsque l'on effectue la projection d'un objet tridimensionnel, chaque projection correspond à la somme de ses densités le long de l'axe de projection, donnant ainsi une image 2D. En réalisant l'opération inverse de la projection, appelée rétroprojection, sur un ensemble de projections 2D obtenues selon des directions de projection différentes, il est possible d'obtenir une reconstruction de l'objet en trois dimensions. C'est ce principe qui est utilisé en microscopie électronique puisque les images enregistrées correspondent à des projections de l'objet qui se trouve généralement dans différentes orientations que se soit sur un film de carbone ou dans une mince pellicule de glace amorphe. Le problème majeur de cette méthode tient du fait qu'il est donc nécessaire de connaître pour chaque projection 2D, l'orientation respective de l'objet. Cette étape s'effectue à l'aide d'un modèle 3D de référence, qui correspond à une structure basse résolution de l'objet étudié. L'obtention de ce modèle de référence est un problème en soi et sera décrite dans la partie 1.2.3.3.

Le modèle de référence sera projeté dans toutes les directions de l'espace, générant un ensemble de projections 2D. Les images expérimentales seront comparées avec ces projections 2D. Ainsi chaque image sera associée par corrélation à une projection 2D du modèle de référence, et prendra ses paramètres de projection, c'est-à-dire les paramètres d'orientation de l'objet qui ont permis de donner cette projection. L'attribution d'orientations des images permet de générer par rétro-projection un nouveau modèle basé cette fois-ci donc sur les données expérimentales. Ce nouveau modèle sera utilisé pour réaliser une nouvelle attribution des paramètres d'orientation des images.

Ainsi, de manière itérative, les paramètres d'orientation pourront être affinés jusqu'à convergence de la reconstruction. En réalité, les images expérimentales sont divisées en deux lots équivalents et deux procédures d'affinement itératif se font en parallèle, produisant chacun une reconstruction. A chaque itération, la corrélation entre les deux reconstructions est calculée dans l'espace de Fourier (« Fourier Shell Correlation » ou FSC), permettant ainsi de suivre l'amélioration de la résolution des reconstructions. Lorsque la FSC ne varie plus, les données sont fusionnées et une reconstruction est réalisée avec l'ensemble des images.

### **1.2.3.3. Obtention d'un modèle de référence**

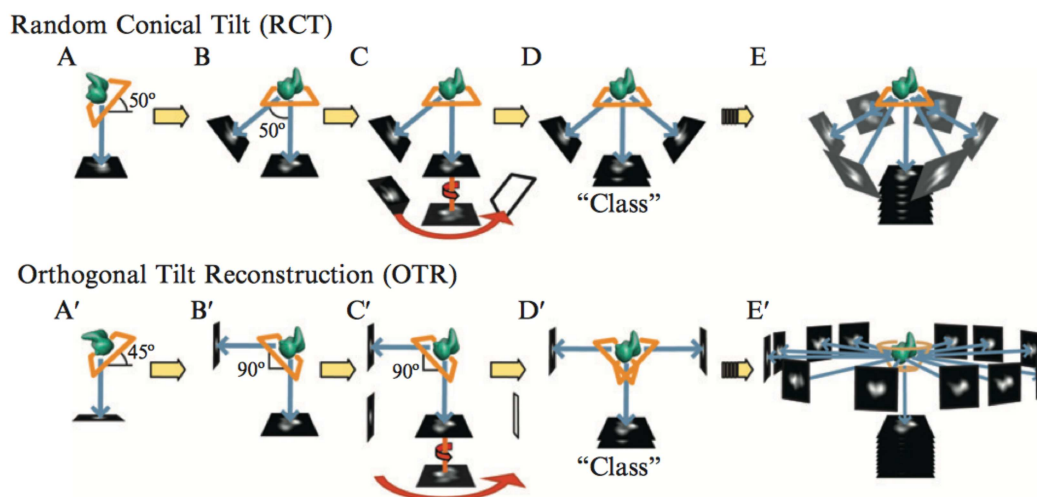
Un modèle de référence est nécessaire pour pouvoir déterminer pour chaque image, les paramètres d'orientation de l'objet. Les images en MET étant acquises en faible dose d'électrons, le rapport signal sur bruit est donc très petit. En d'autres termes, les images ne sont pas enregistrées avec suffisamment d'électrons pour donner une image nette et précise et présentent donc beaucoup de bruit qui vient masquer le signal. La présence du bruit dans les projections rend le choix du modèle de référence très important. En effet, au cours du traitement, les images expérimentales vont toujours s'aligner par rapport aux projections du modèle de référence, et ce quelque soit le modèle initial choisi (Henderson 2013). On aura l'illusion que la reconstruction progresse au cours des cycles itératifs alors que le signal ne proviendra que du bruit contenu dans les images et du modèle de référence. Il convient donc d'être extrêmement prudent lors de l'obtention du modèle initial de référence et d'être critique par rapport au résultat obtenu.

Le modèle de référence peut correspondre à une structure homologue déjà caractérisée, ou bien il peut avoir été obtenu *ab initio* à partir des données expérimentales. Différentes méthodes existent pour créer un premier modèle de référence à partir des données expérimentales.



### 1.2.3.3.1. L'inclinaison conique aléatoire et la reconstruction inclinée orthogonale

Le principe de ces deux méthodes est le même : un champ de particules est photographié successivement à deux angles distincts, afin d'obtenir des « paires inclinées ». Dans le cas de l'inclinaison conique aléatoire ou « Random conical tilt » (Radermacher et al. 1987), les angles sont typiquement de  $50^\circ$ , pour l'inclinaison orthogonale ou « Orthogonal tilt reconstruction » (Leschziner & Nogales 2006), les angles sont de  $90^\circ$  (la première image est enregistrée à  $-45^\circ$  et la seconde à  $+45^\circ$ ). La reconstruction avec une inclinaison orthogonale est plus récente, et présente l'avantage de ne pas avoir de coins manquants, contrairement à l'inclinaison conique aléatoire. Les projections 2D issues de la photographie avant inclinaison sont ensuite alignées entre elles. Ainsi, en connaissant l'angle d'inclinaison, les projections 2D issues de la photographie après inclinaison sont alors orientées dans l'espace, et le modèle est ainsi reconstruit (**Figure 13**).

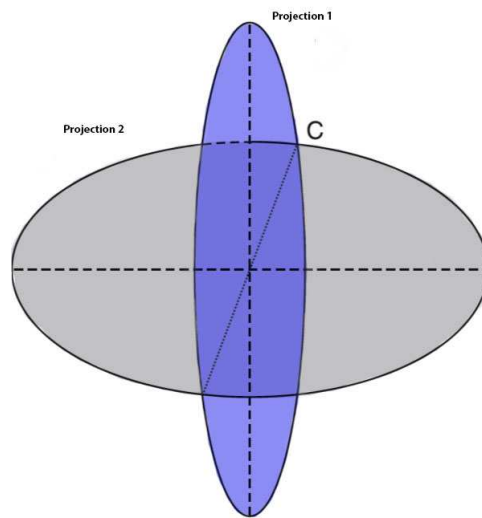


**Figure 13 : Illustration de l'inclinaison conique aléatoire et de la reconstruction inclinée orthogonale.** (A, A') Acquisition d'une première image à un angle choisi. (B, B') Acquisition d'une deuxième image après inclinaison, permettant de créer des paires inclinées. (C, C') Les images issues de l'acquisition avant inclinaison sont alignées entre elles, si elles représentent la même vue. Les images issues de l'acquisition après inclinaison sont donc repositionnées en fonction de la rotation que les premières particules ont subie pendant l'alignement. (D, D') Les images alignées sont maintenant regroupées en classes de particules ayant la même orientation. (E, E') Le modèle est reconstruit si le nombre d'images est suffisant. Reproduit de (Leschziner 2010).

### 1.2.3.3.2. Le théorème des lignes communes des projections

Dans l'espace de Fourier, chaque projection 2D d'un objet 3D correspond à une section centrale de la transformée de Fourier (TF) 3D de l'objet. L'orientation de cette section est directement liée à

l'orientation de l'objet 3D qui a donné la projection 2D considérée. Par conséquent, deux projections 2D d'un objet, donc deux images de l'objet, correspondront à deux sections de la TF 3D de l'objet. Il y aura donc une ligne commune correspondant à l'intersection des deux sections. Le logiciel Imagic (van Heel & Keegstra 1981) tire profit de ce théorème pour trouver des relations géométriques entre des images données et ainsi pouvoir leur attribuer des orientations respectives. Pour cela, le programme va aligner et classer les images des particules, puis additionner entre-elles les images afin d'obtenir des moyennes de classe, dans lesquelles le rapport signal sur bruit est donc fortement augmenté. Dans l'espace réciproque, les sections centrales des moyennes de classes sont alors comparées pour obtenir des informations sur leur orientation relative (**Figure 14**).



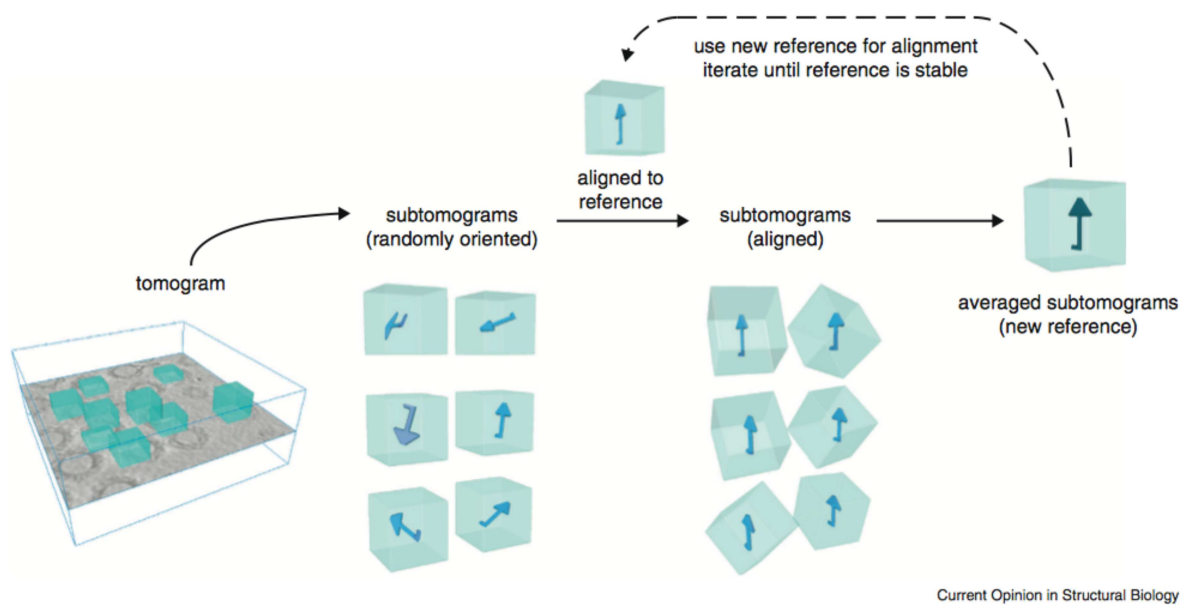
**Figure 14** : Schéma d'une ligne commune (C) de deux projections d'une même particule dans l'espace de Fourier.

En connaissant la relation angulaire entre les particules, il est possible de combiner entre elles les images afin de générer un modèle 3D de référence. Cette étape est typiquement réalisée par le logiciel Imagic dans l'espace réel à partir de sinogrammes, tandis que le logiciel EMAN2 utilise cette approche dans l'espace réciproque.

### 1.2.3.3.3. Tomographie électronique et addition de sous-volumes

La cryo-tomographie électronique (cryo-TE), est une technique de microscopie électronique où plusieurs images d'une même zone sont prises à des angles différents (typiquement de  $-60^\circ$  à  $+60^\circ$ ). Cette série d'images est ensuite alignée et combinée pour obtenir un volume tridimensionnel. À cause des contraintes physiques du porte-objet, la tomographie ne permet pas d'obtenir un jeu

d'information complet puisque certains angles ne sont pas imagés. Cependant, à partir du tomogramme il est possible d'extraire différentes zones (appelées sous-volumes) contenant des particules. Ces sous-volumes, correspondant aux particules adoptant différentes orientations, peuvent ensuite être additionnés pour produire un modèle initial complet (**Figure 15**). Du fait du nombre d'images enregistrées, la dose d'électrons doit être diminuée au maximum afin d'éviter les dommages d'irradiation. Il est donc important de maintenir un équilibre entre le nombre d'images collectées et la dose d'électrons afin d'avoir le contraste suffisant pour l'alignement du tomogramme.



**Figure 15 : Principe de l'addition de sous-volumes acquis par cryo-TE.** Les sous-volumes sont extraits du tomogramme et alignés contre une référence de manière itérative pour créer un nouveau modèle. La référence utilisée pour l'alignement des sous-volumes peut avoir été générée ultérieurement ou peut être créée directement à partir de tous les sous-volumes moyennés. Reproduit de (Briggs 2013).

### 1.3. La révolution de la microscopie électronique

Au cours de ces cinq dernières années, de nombreux développements ont permis l'avènement d'une nouvelle ère en microscopie électronique. On parle même de « révolution de la résolution » de la microscopie, tant le bond technologique est important (Nature Methods 2016). Avant ce saut technologique, la microscopie électronique à transmission était plutôt décrite comme une approche structurale « basse résolution » par rapport aux deux autres techniques de biologie structurale que sont la cristallographie aux rayons X et la résonance magnétique nucléaire. La résolution des



structures était comprise entre 5-6 et 10 Å ne permettant pas d'accéder à l'organisation atomique des objets étudiés. La microscopie électronique permettait cependant d'apporter des informations structurales sur les complexes moléculaires, à l'intérieur desquels étaient intégrées des données hautes résolutions provenant de la cristallographie et de la RMN.

Le passage de l'enregistrement d'images sur film photographique vers la numérisation directe grâce aux détecteurs CCD (*Charge Coupled Device*) avait également représenté une avancée majeure. Ces détecteurs ont permis un gain de temps significatif en supprimant les contraintes de développements des films et leur numérisation. Néanmoins, de nombreux microscopistes ont continué à préférer les films photographiques, qui présentaient une plus grande sensibilité et une plus grande surface d'acquisition que les détecteurs CCD.

Ladite « révolution » est la conjonction de l'évolution de plusieurs paramètres distincts (préparation de l'échantillon, acquisition automatique, traitement d'image) mais a réellement pris toute sa mesure en 2012, lorsque de nouveaux détecteurs d'électrons ont vu le jour. Les grands noms de la microscopie électronique tels que Richard Henderson ou David Agard ont travaillé avec les industriels pour mettre au point cette nouvelle génération de détecteurs (Callaway 2015). Des structures à des résolutions inférieures à 5 Å ont alors été publiées. Les virus symétriques (Campbell et al. 2012; Veesler et al. 2013), le ribosome (Bai et al. 2013) et le protéasome (Li et al. 2013) ont été les premières structures hautes résolutions obtenues par cryo-ME. Des protéines membranaires ont également été résolues à des résolutions inférieure à 4 Å sans avoir recours à la cristallisation, telles que TRPV1 à 3,4 Å (Transient receptor potential channel sub family V member 1) (Liao et al. 2013) et la  $\gamma$ -sécrétase humaine à 3,4 Å également (X. Bai et al. 2015).

En 2016, de nouvelles barrières ont été franchies. Ainsi, il a été possible d'obtenir la structure de la GDH (glutamate déshydrogénase) à 1,8 Å de résolution, ce qui représente à l'heure actuelle la meilleure résolution jamais atteinte. De même, la structure de l'IDH (Isocitrate déshydrogénase) qui a un poids moléculaire de 93 kDa a été calculée à 3,8 Å de résolution, ce qui est la plus petite structure actuellement résolue à un niveau atomique par cryo-ME en particule isolée (Merk et al. 2016).

Dans ce sous chapitre sera présenté les grandes avancées que la cryo-ME a connu, ainsi que celles toujours en cours de développements, depuis la préparation de l'échantillon jusqu'au traitement des données.

### 1.3.1. Des échantillons plus stables et des conditions plus reproductibles

La plupart des industriels ont développé différents systèmes semi-automatiques permettant de contrôler au mieux les paramètres de congélation des échantillons et d'améliorer la reproductibilité de cette étape critique. De nombreux développements sont actuellement en cours afin d'une part d'améliorer la stabilité de l'échantillon et d'autre part d'améliorer les conditions de congélation des échantillons, afin notamment d'obtenir une glace fine en diminuant la quantité d'échantillon utilisée.

#### 1.3.1.1. Les systèmes de congélation actuels

Les systèmes de congélation, autrefois appelés « guillotines » ou « plongeurs », sont composés d'un récipient contenant de l'éthane liquide, entouré par une cuve remplie d'azote liquide. La grille de microscopie électronique est maintenue par une pince au-dessus du récipient d'éthane liquide, à l'intérieur d'une chambre humide. Entre 2 et 3  $\mu\text{l}$  d'échantillon sont nécessaires pour une grille. L'excès de liquide est ensuite adsorbé sur papier filtre, et la grille est instantanément plongée dans le récipient contenant l'éthane liquide. Les systèmes de congélation les plus utilisés sont distribués par les compagnies industrielles FEI (Vitrobot), Gatan (CP3) et Leica (EM GP), et permettent de contrôler presque tous les paramètres de congélation de l'échantillon (**Figure 16**). Selon les modèles, il est possible de contrôler la température et le taux d'humidité de la chambre dans laquelle la grille est placée. L'adsorption de l'excès de liquide se fait de manière automatique par adsorption sur papier filtre sur une face ou sur les deux faces de la grille. Il est également possible de contrôler le temps de contact ou la pression du contact du papier filtre avec la grille. Certains systèmes peuvent également contrôler précisément la température du pot d'éthane liquide.

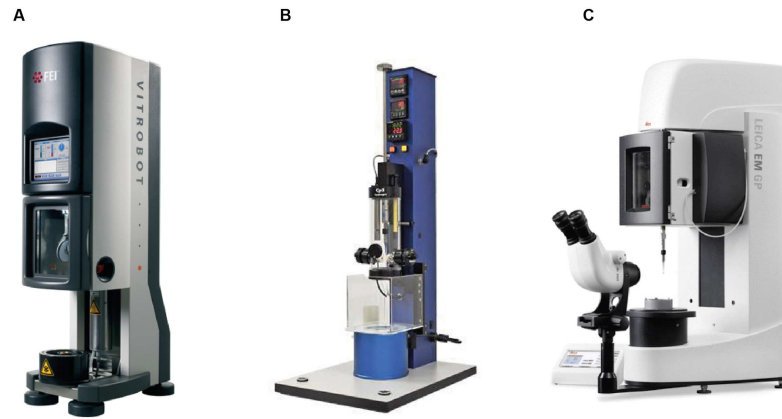
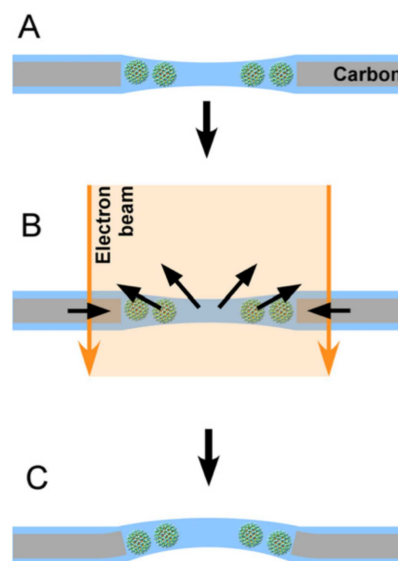


Figure 16 : Les différents systèmes de congélation actuels. (A) Vitrobot (FEI). (B) CP3 (Gatan). (C) EM GP (Leica).

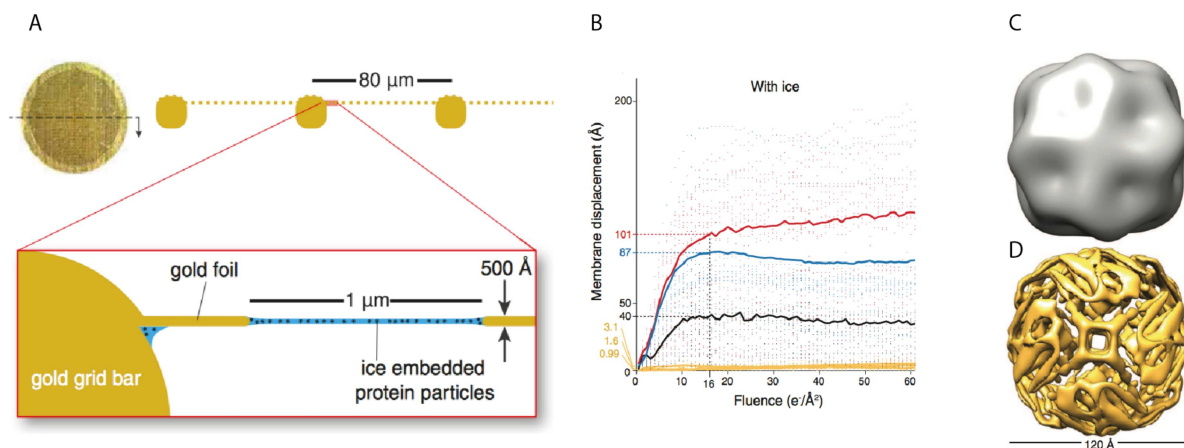
### 1.3.1.2. La nouvelle ruée vers l'or

Il a été montré que lors de l'illumination de l'échantillon par le faisceau d'électrons, les particules ne restent pas figées, mais bougent. Ce mouvement est à la fois dû au film de carbone de la grille qui se contracte et se déforme, ainsi qu'à la formation de radicaux libres qui augmente la pression dans la couche de glace (Figure 17). L'idée d'une pré-irradiation de la zone dans le but de ne plus provoquer de mouvement pendant l'acquisition n'est pas envisageable puisqu'une seconde illumination sur la même zone entraîne une répétition du processus avec un mouvement aussi important. Néanmoins, une exposition longue stabilise la zone, suggérant que le faisceau d'électrons provoque des mouvements plus importants durant les premières irradiations (Brilot et al. 2012).



**Figure 17 : Mouvement de l'échantillon sous l'effet du faisceau d'électrons.** (A) Particules en suspension dans un trou de glace vitreuse. (B) Irradiation de l'échantillon par les électrons et mouvement des particules à cause de la déformation du carbone et de la formation de radicaux libres. (C) Les changements persistent après le passage du faisceau d'électrons. Reproduit de (Brilot et al. 2012).

En 2014, un nouveau type de support constitué entièrement d'or a été développé (**Figure 18A**) (Russo & Passmore 2014). Il a été montré que ces grilles d'or sont plus adaptées pour la haute résolution, car elles diminuent la déformation de la grille induite par le passage du faisceau (**Figure 18B**). Une étude comparative a été réalisée avec des particules d'apoferritine formant un complexe de 440 kDa qui ont été congelées sur des grilles de carbone et d'or. Des images acquises sur les deux types de grilles dans les mêmes conditions d'acquisition ont permis de reconstruire deux modèles à des résolutions différentes. Les particules sur les grilles d'or ont subi moins de mouvement du faisceau durant l'acquisition, permettant de résoudre la structure de l'apoferritine à plus haute résolution (**Figure 18C et D**).

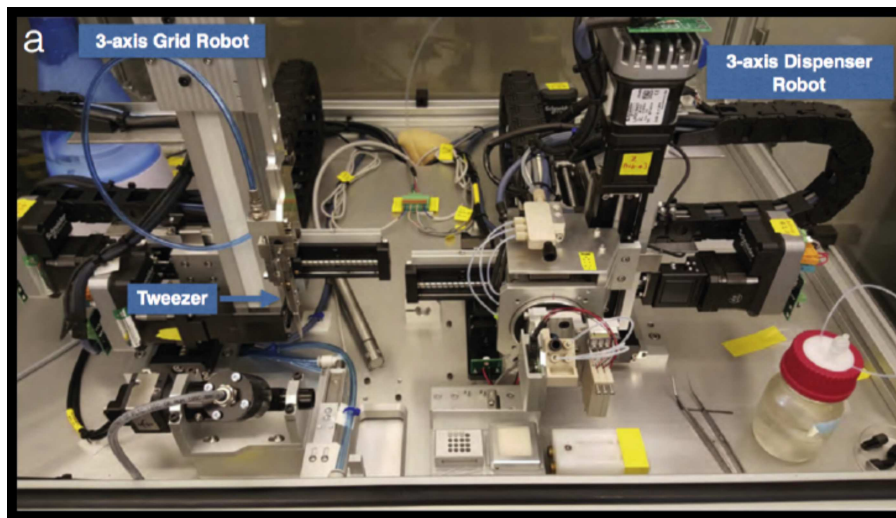


**Figure 18 : Effet des grilles en or de microscopie électronique.** (A) Schéma de la structure d'une grille d'or. (B) Moyenne quadratique du mouvement vertical de l'extrémité des trous mesurée sur 3 grilles de carbones (rouge, bleu et noir) et 3 grilles en or congelées. (C) Structure finale de l'apoferritine à 25 Å avec acquisition sur une grille de carbone. (D) Structure finale de l'apoferritine à 8 Å avec acquisition sur une grille d'or. Reproduit de (Russo & Passmore 2014).

L'utilisation des grilles d'or donne la possibilité d'obtenir des micrographies contenant plus d'informations. De cette façon, il deviendrait envisageable d'obtenir une structure atomique avec un total de micrographies inférieur à ce qui se fait aujourd'hui. Ces grilles sont maintenant brevetées par le constructeur Quantifoil sous le nom UltrAuFoil, mais la méthode de fabrication est décrite dans (Russo & Passmore 2016).

### 1.3.1.3. Vers une technique entièrement automatisée ?

Outre-Atlantique, l'équipe de Bridget Carragher a réalisé le développement d'un robot dédié à la congélation automatique d'échantillon : Spotiton 1.0. Spotiton est un robot prototype destiné à préparer l'échantillon pour la cryo-ME (**Figure 19**). Dans le même principe qu'une guillotine, Spotiton dépose l'échantillon sur la grille, la plonge dans le cryogène et transfère la grille dans une boîte remplie à l'azote liquide.



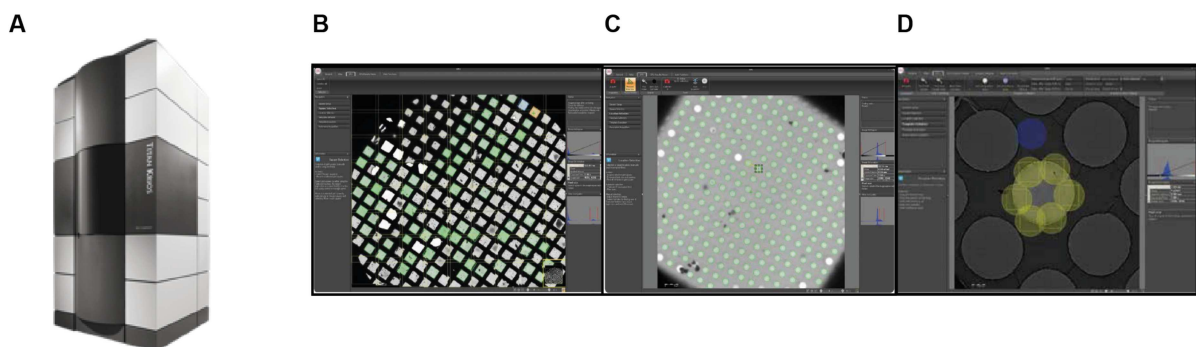
**Figure 19 : Photo du robot Spotiton.** À droite, le robot bouge dans 3 directions pour positionner la grille (maintenue à la verticale au bout du « tweezer »). À gauche, le robot bouge également dans 3 directions pour venir injecter l'échantillon sur la grille. Reproduit de (Razinkov et al. 2016) .

L'automatisation permet de travailler avec très peu de volume (entre 2,5 et 16 nL) et le blot papier est remplacé par un système automatique d'adsorption. En effet, les grilles sont recouvertes de nanofibres par traitement des grilles avec l'ammonium persulfate. Ces nanofibres vont aspirer l'excès d'eau de la même manière qu'une adsorption sur papier filtre (**Figure 20**), c'est à dire par capillarité.



- La stabilité du microscope est telle que l'alignement du microscope permet de collecter sans interruption des images durant 24h. Il faut ensuite procéder à un nouvel alignement du correcteur de Cs et refaire la normalisation de la caméra. Il est fort probable que ces étapes seront automatisées prochainement.

Pour la cryo-ME où des grilles trouées sont traditionnellement utilisées, il est nécessaire d'avoir une régularité des trous (diamètre et espace entre chaque trou) pour faire de l'acquisition automatique. Des grilles à trous réguliers produites par lithographie ont vu le jour ces dernières années, commercialisées actuellement par les sociétés Quantifoil ou Protochips. Dans le cas d'un microscope Titan Krios équipé d'une caméra Falcon II (FEI), l'acquisition automatique se fait avec le logiciel propriétaire EPU (FEI). L'utilisateur doit seulement sélectionner manuellement les zones d'acquisition, c'est-à-dire les carreaux à partir de l'atlas de la grille, puis les trous et les zones dans chaque trou. Il doit ensuite entrer les valeurs de défocalisations désirées, et l'acquisition se fait ensuite automatiquement. Le logiciel EPU va réaliser un déplacement du faisceau pour que plusieurs micrographies soient enregistrées par trous, ainsi qu'un déplacement mécanique de la grille pour bouger de trou en trou (**Figure 21B-D**).



**Figure 21 : Acquisitions réalisées sur le microscope Titan Krios. (A)** Photo du Titan Krios. **(B)** Photo de l'Atlas d'une grille qui permet de sélectionner les carreaux manuellement. **(C)** Photo d'un carreau où chaque trou est choisi manuellement. **(D)** Photo des positions du faisceau dans chaque trou (fei.com).

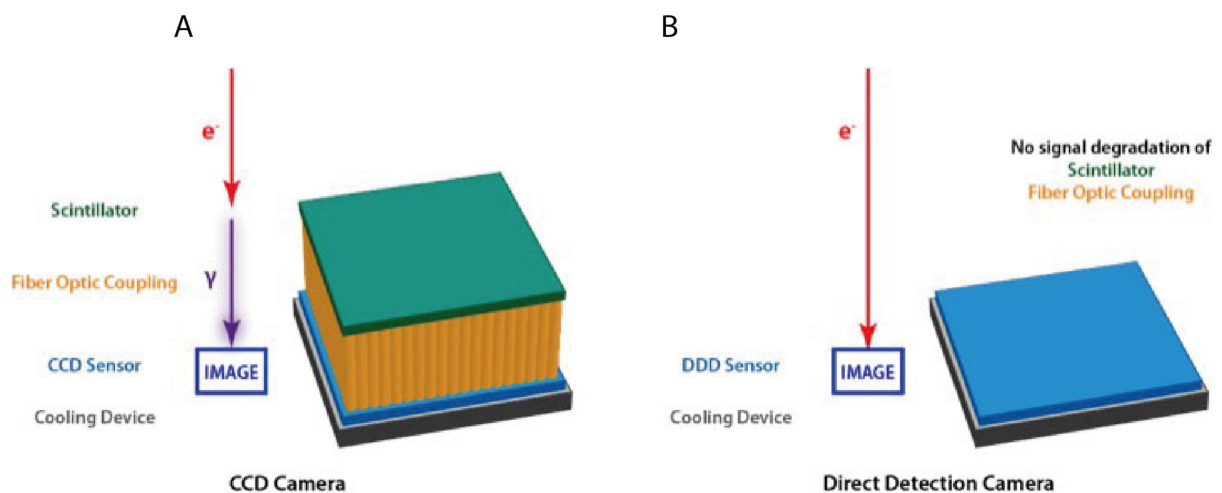
L'acquisition automatique permet un gain de temps significatif aussi bien au niveau du nombre de micrographes obtenus qu'au niveau du temps libéré pour le microscopiste qui n'intervient qu'au niveau du choix des zones et de l'alignement du microscope. Seul le logiciel EPU a été illustré dans ce sous-chapitre, et a été utilisé pour collecter automatiquement les micrographes durant ma thèse. Cependant, il existe aujourd'hui 11 logiciels différents (dont 3 commerciaux) adaptés pour



l'acquisition automatique d'image ou de séries inclinées pour la tomographie qui sont référencés dans la revue *Microscopy* (Tan et al. 2015).

### 1.3.3. Les caméras à détection directe

Les nouvelles caméras sont principalement responsables de la révolution de la résolution en microscopie électronique. L'apparition des caméras à détection directe sur le marché a donné l'accès à la très haute résolution et signé la montée en popularité de la cryo-ME. Traditionnellement, les caméras CCD utilisent un scintillateur pour convertir les électrons en photons et des fibres optiques pour transmettre ces photons (**Figure 22A**). L'avantage de ces caméras était un gain logistique par rapport aux films photographiques, les acquisitions étant réalisées en temps réel et directement digitalisées. Par contre, le confort logistique s'est fait au prix d'une sensibilité plus mauvaise que celle des films photographiques. Aujourd'hui, les nouvelles caméras utilisant la technologie CMOS (*Complementary Meta Oxide Semiconductor*) peuvent détecter les électrons directement, sans conversion en photons (**Figure 22B**).



**Figure 22 : Évolution des caméras équipant les microscopes électroniques à transmission. (A)** Schéma des composantes d'une caméra CCD. **(B)** Schéma des composantes d'une caméra à détection directe des électrons. Reproduit de *directelectron.com*.

Ces caméras de nouvelle génération sont aujourd'hui vendues par 3 industriels : FEI, Gatan et Direct Electron. Par rapport aux caméras CCD, elles offrent un meilleur rapport signal/bruit que les films photographiques ainsi qu'une meilleure vitesse d'acquisition, donnant l'opportunité d'acquérir non plus des photographies, mais des films.



### 1.3.3.1. Un meilleur rapport signal/bruit

Le meilleur moyen de mesurer la performance des caméras est l'Efficacité Quantique de Détection (Detective Quantum Efficiency ou DQE). La DQE est une indication de mesure permettant de connaître à quel point le signal détecté est élevé par rapport au bruit pour une fréquence spatiale donnée. La DQE d'une caméra est calculée avec trois composantes (**équation 7**) :

$$DQE(N, s) = \frac{MTF^2(s)}{NPS_{out}(N, s)/N} \quad (7)$$

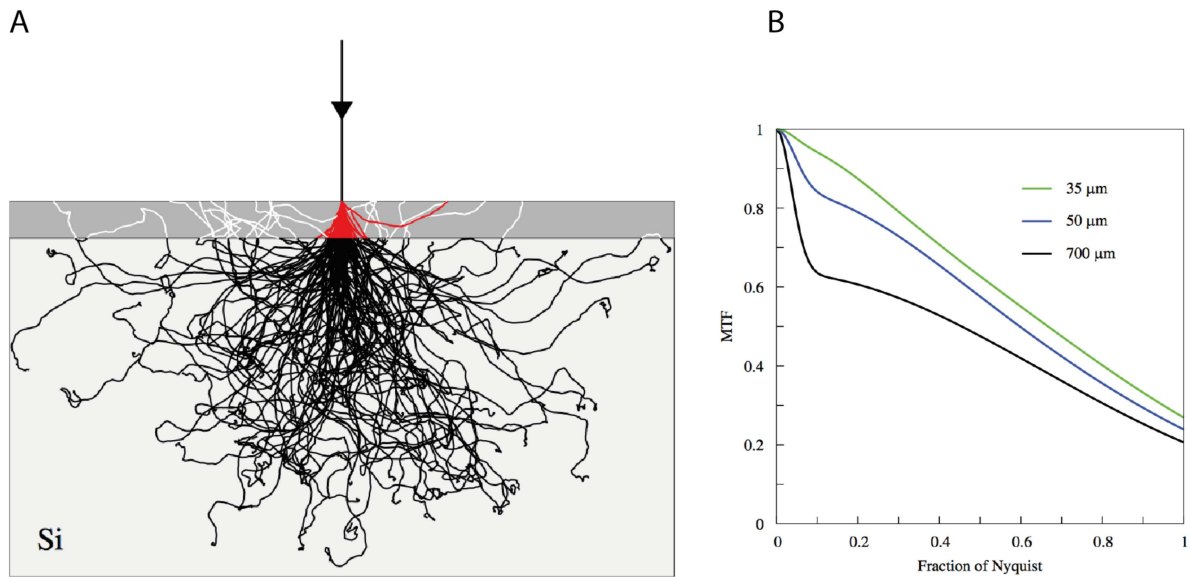
N = dose moyenne d'électrons par pixel

s = fréquence spatiale

$NPS_{out}$  = Spectre de puissance du bruit détecté

MTF = Fonction de transfert modulaire

La DQE dépend donc de la MTF qui est la réponse du détecteur à une fréquence spatiale donnée. Elle indique ainsi la performance du détecteur pour une fréquence spatiale donnée. La MTF est directement liée à l'épaisseur du détecteur (McMullan et al. 2014). Les causes d'une faible MTF sont le phénomène de diffusion des électrons dans l'épicoche du détecteur où le signal est détecté, ainsi que la diffusion inverse des électrons qui sont déjà dans le substrat du détecteur, mais reviennent vers l'épicoche. Pour pallier au problème de diffusion inverse des électrons, les industriels ont affiné le détecteur, qui empêche la diffusion inverse des électrons et contribue à améliorer la MTF de la caméra (**Figure 23**).

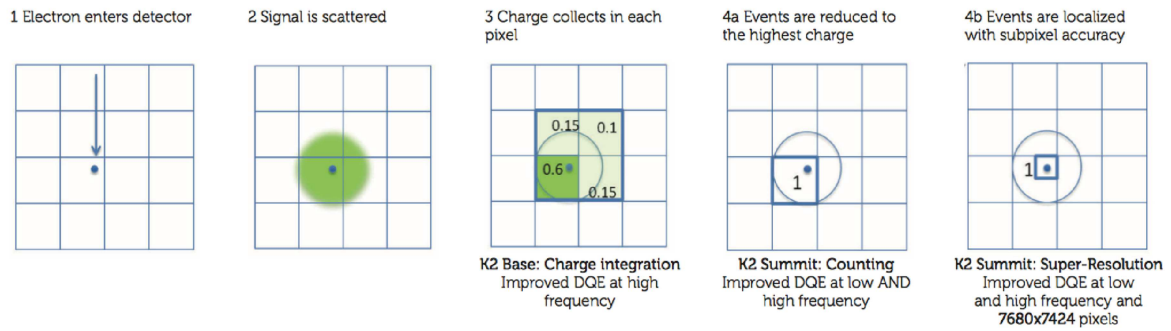


**Figure 23 : Impact de l'affinement du détecteur sur la fonction de transfert modulaire. (A)** Simulation Monte Carlo de la diffusion des électrons accélérés à 300 kV avant affinement du détecteur (rectangle gris foncé + gris clair, 350 μm d'épaisseur). Avant affinement du détecteur, les électrons détectés sont les lignes rouges et blanches. Après affinement, les électrons détectés sont seulement les lignes rouges. **(B)** Variation de la MTF de plusieurs détecteurs d'épaisseur différente en fonction de la fréquence spatiale. Reproduit de (McMullan et al. 2009).

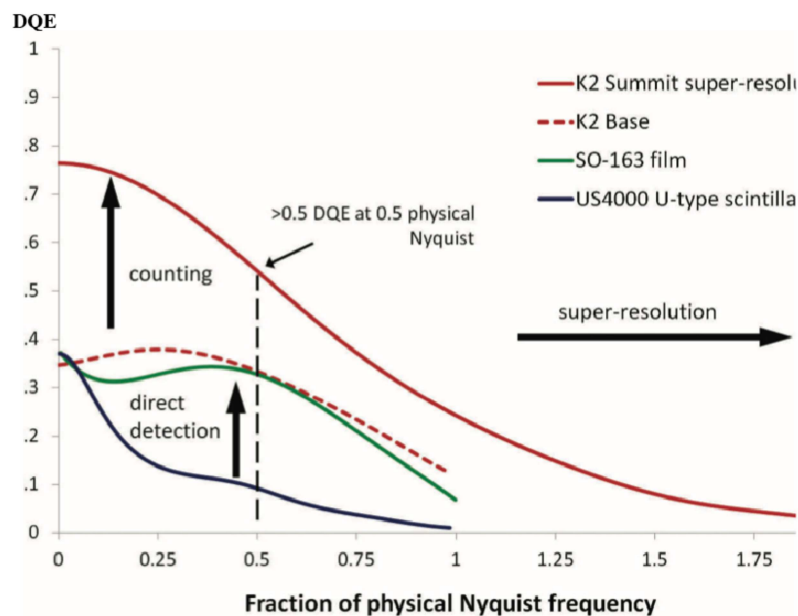
Pour aller plus loin, les derniers modèles de caméras proposent de nouveaux modes d'acquisition destinés à augmenter la DQE. Pour illustrer ces améliorations, les différents modes présentés ci-dessous sont proposés par l'industriel Gatan et sa caméra Gatan K2 Summit (Booth & Mooney 2013) qui est pour l'instant la seule caméra commercialisée proposant 3 modes.

En effet, la société Gatan propose 3 modes d'acquisition différents (**Figure 24**) permet d'obtenir une DQE grandement supérieure à ce qui se faisait auparavant (**Figure 25**) :

- Le mode intégré qui, lorsqu'une charge est diffusée sur plusieurs pixels, collecte les charges de chaque pixel ;
- Le mode comptage où la charge est collectée uniquement dans le pixel avec la charge la plus haute ;
- Le mode super-résolution où la plus haute charge est localisée à une précision sub-pixel.

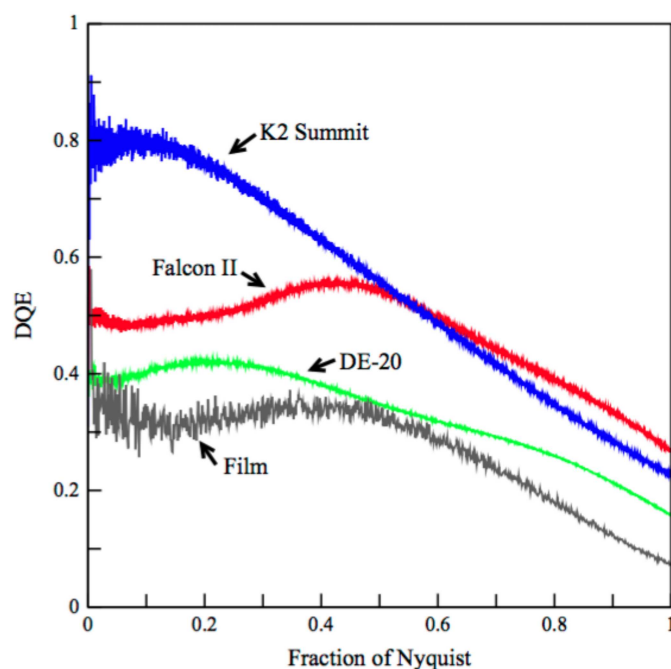


**Figure 24 : Les différents modes d'acquisition proposée par la caméra Gatan K2 Summit. (1) Arrivée de l'électron sur un pixel du détecteur. (2) Diffusion du signal. (3) Mode intégré : l'ensemble des pixels où la charge s'est déposée sont intégrés et lus par le détecteur. (4a) Mode comptage : seul le pixel qui a été frappé par la valeur de la charge la plus haute est lu par le détecteur. (4b) Mode super-résolution : par rapport au mode comptage, la précision est maintenant à un quart du pixel. Reproduit de (Booth & Mooney 2013).**



**Figure 25 : Comparaison de la variation de DQE différent modes de détection. Caméra CCD (ligne noire), film photographique (ligne verte), caméra à détection directe en mode intégré (pointillés rouges) et en mode comptage (ligne rouge). Reproduit de (Booth & Mooney 2013).**

Toutefois, le mode super-résolution n'est pas la condition *sine qua non* pour atteindre la haute résolution et de nombreuses structures quasi atomiques ont été caractérisées grâce à des caméras en mode intégré. Le mode super-résolution offre plus d'informations à basse résolution, ce qui est nécessaire à la bonne orientation des particules, mais demande un temps d'acquisition plus important et donc un échantillon relativement stable. De plus, la difficulté à connaître l'emplacement exact de l'arrivée de l'électron rend le mode intégré plus adapté pour obtenir des informations de haute résolution avec une DQE plus élevée dans les hautes fréquences (**Figure 26**).



**Figure 26 : Comparaison de la variation de DQE des différentes caméras à détection directe.** Film photographique (tracé gris), caméra à détection directe en mode intégré du constructeur DirectElectron (tracé vert), FEI (tracé rouge) et en mode comptage du constructeur Gatan (tracé bleu). Reproduit de (McMullan et al. 2014).

La caméra K2 de Gatan reste néanmoins la caméra la plus sensible du marché et a permis d'obtenir les premières structures à moins de 3 Å de résolution (Campbell et al. 2015; Jiang et al. 2015). Cependant, les autres constructeurs vont fortement développer l'ajout des modes comptages dans la prochaine génération de caméras.

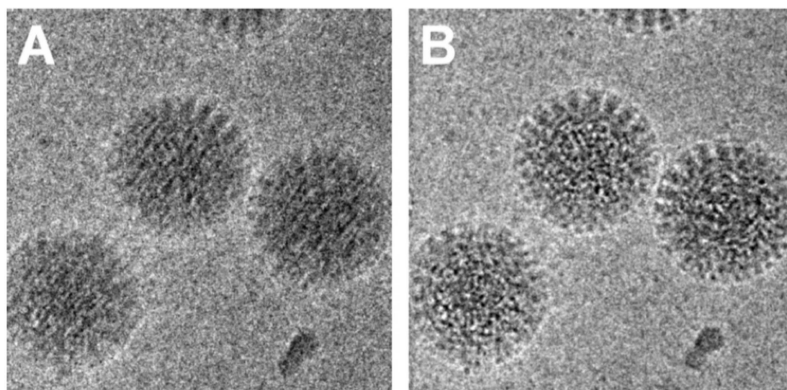
### 1.3.3.2. Le mode film

Tout comme l'invention du cinéma en 1895 par les frères Lumières a révolutionné le monde, les dernières caméras à détection directe ont révolutionné le monde de la microscopie électronique à transmission. L'amélioration majeure des caméras à détection directe est l'augmentation de la vitesse d'acquisition qui permet d'utiliser un mode appelé « mode film », permettant d'enregistrer une série de vues qui forment un film. L'enregistrement des films est le grand acteur de la révolution en cryo-ME, puisque couplé aux performances des dernières caméras, il permet d'améliorer significativement la qualité des données utilisée pendant le traitement d'image et d'accéder à des résolutions inférieures à 5 Å.

### 1.3.4. Le traitement des films

Les films ont ouvert la porte à la très haute résolution, nous ne parlons plus aujourd'hui seulement de traitement d'image, mais également de traitement de films. La plupart des logiciels de traitement d'image intègrent maintenant le traitement des films, et d'autres, uniquement dédiés à cela ont vu le jour ; par exemple MotionCorr (Li et al. 2013) ou Unblur (Grant & Grigorieff 2015b). Dans le but d'améliorer la qualité des données acquises, trois aspects sont traités pendant le traitement des films :

- La correction du mouvement causé par l'irradiation des électrons sur l'échantillon (**Figure 27**).
- Un contrôle du nombre de vues gardées pour la reconstruction. Il peut être judicieux d'enlever les 2-3 premières vues où le mouvement engendré par l'irradiation est le plus important (Russo & Passmore 2016), ainsi que les dernières vues où la dose d'irradiation peut être trop élevée.
- L'application d'un filtre qui va pondérer chaque vue par rapport à la dose d'électron reçue afin d'améliorer la qualité du signal. Ces filtres doivent être utilisés après correction du mouvement des particules et décrivent seulement la dégradation du signal au cours de l'irradiation par le faisceau d'électrons. À ce jour, seulement deux logiciels proposent de pondérer les vues : Unblur et Relion (Scheres 2014).



**Figure 27 : Effet de l'alignement des vues sur la qualité du signal.** 60 vues du rotavirus VP6 ont été moyennées avant réalignement (A) et après réalignement (B) (Brilot et al. 2012).

### 1.3.5. Les méthodes de validation de la résolution

L'estimation de la résolution d'une carte de densité obtenue par cryo-ME se fait avec la « Fourier Shell Correlation » (FSC) qui consiste à mesurer l'accord entre deux reconstructions 3D du même objet, qui ont été réalisées en parallèle après séparation aléatoire des particules en deux lots distincts.

Deux méthodes de validation sont maintenant acceptées par la communauté :

- La « gold standard FSC » : une méthode de reconstruction où les données sont séparées en deux moitiés et l'affinement de la carte de densité se fait indépendamment de l'autre moitié. (Scheres & Chen 2012).
- La substitution par du bruit dans les hautes résolutions : au-delà d'une résolution choisie, les phases de la reconstruction 3D sont remplacées par des phases aléatoires pour simuler du bruit (Chen et al. 2013). Une comparaison avec les phases est faite avec les phases originales pour vérifier la présence de bruit dans la reconstruction, au travers de l'équation 8 :

$$FSC_{réelle} = \frac{FSC_{données} - FSC_{bruit}}{1 - FSC_{bruit}} \quad (8)$$

### 1.3.6. Le dépôt des structures

Les modèles atomiques résolus par microscopie électronique sont déposés sur la banque de données de protéines (*Protein Data Bank* ou PDB) et les cartes de densités sur la banque de données de microscopie électronique (*Electron Microscopy Data Bank* ou EMDB) qui sont des serveurs publics pour le dépôt de structures. La PDB sert de dépôt pour tous les modèles obtenus par résonance magnétique nucléaire, cristallographie et microscopie électronique. L'EMDB sert de dépôt pour les cartes de densités électroniques obtenues par particules isolées, tomographie électronique, cristallographie 2D ou micro-diffraction électronique. Il existe maintenant un site de dépôt pour les données brutes de microscopie électronique appelée Archive des images pilotes de microscopie électronique (*Electron Microscopy Pilot Image Archive* ou EMPIAR), née de l'initiative de la banque européenne de données de protéines (*Protein Data Bank in Europe* ou PDBe) dont le but est de rendre accessibles les données ayant servi à la reconstruction des cartes de densités et de faciliter ainsi le développement des méthodes de validation.

### 1.3.7. Conclusion

Les grandes avancées de ces dernières années ont fait de la cryo-ME un acteur majeur de la biologie structurale. Les conventions de format et de validation commencent à apparaître, le nombre de structures à des résolutions supérieures à 4 Å continue d'augmenter et d'importants développements sont toujours en cours pour ouvrir le traitement d'image à un public moins averti. Le grand désavantage de la cryo-ME est le coût engendré, aussi bien pour l'achat du microscope et de la caméra que pour les consommables et le coût d'entretien. Par ailleurs, les méthodes de traitement d'image actuelles demandent un nombre élevé d'images en particulier pour les particules sans symétrie, ce qui se répercute indéniablement sur le temps d'utilisation du microscope, le stockage des données et l'infrastructure informatique. Malgré ces contraintes, la cryo-ME reste la technique phare de l'année qui se traduit par l'achat de nombreux Titan Krios partout à travers le monde. Chaque année les limites de la technique sont repoussées, des protéines de plus en plus petites sont maintenant caractérisées à des résolutions inférieures à 2 Å (Merk et al. 2016).

## 2. Objectifs de la thèse

Mon projet de thèse était initialement organisé autour de l'étude structurale des virus FBNSV et CaMV et des protéines virales impliquées dans leur transmission par les pucerons. La stratégie a été d'étudier la structure des particules virales par cryo-ME et traitement d'image soit à partir de virus natifs purifiés à partir de plantes infectées, soit à partir de capsides issues de l'auto assemblage de protéines de capsidite exprimées et purifiées en système homologue ou hétérologue. Une grande partie de ce travail de thèse a été de purifier diverses protéines virales (protéines de capsidite et/ou protéines de transmission) entières ou par domaine, nous avons alors envisagé de faire des études structurales de ces protéines recombinantes par cristallographie aux rayons X. La combinaison des données structurales issues de la cryo-ME et de la cristallographie nous permet ainsi d'appréhender au niveau moléculaire la formation des complexes de transmission, à savoir les complexes formés par la particule virale et la protéine de transmission.

Dans ce contexte, une part importante de mon travail a porté sur la caractérisation structurale de la capsidite du FBNSV et sur l'expression et la purification de la protéine de capsidite recombinante ainsi que de la protéine impliquée dans la transmission avec le puceron. La purification du FBNSV entier à partir de plantes infectées a été réalisée par le Dr Zeddam de l'équipe du Dr Blanc (Équipe « Interaction Virus Insecte Plante », Biologie et Génétique des Interactions Plante-Parasite ; UMR BGPI TA A54/K ; Campus International de Baillarguet, 34398 Montpellier Cedex 5, France). Après plusieurs tentatives de purification du FBNSV, Dr Zeddam a réussi à obtenir une fraction purifiée dont la concentration en virus était compatible avec une approche par cryo-ME et traitement d'image. J'ai donc congelé plusieurs grilles de ce virus afin de pouvoir effectuer une campagne d'acquisition d'images sur un microscope de type Titan Krios équipé d'une caméra à détection directe. Malheureusement, les premières reconstructions de ce virus nous ont laissé penser qu'il ne s'agissait pas du FBNSV, mais du virus de la mosaïque du niébé (Cowpea Mosaic Virus ou CPMV), un virus dont la structure avait déjà été caractérisée par cristallographie aux rayons X (Lin et al. 1999) et tout récemment par cryo-ME (Hesketh et al. 2015). Le virus purifié provenait en fait d'une contamination des graines de fèves. D'autre part, les approches par expression des protéines recombinantes de capsidite et de transmission n'ont pas été concluantes, que ce soit en système d'expression *E. coli* ou en système de cellules d'insecte *Sf9*. Par conséquent, afin de mettre à profit les approches structurales par cryo-ME et de traitement d'image initialement établies à partir du virus contaminant, j'ai élargi mon travail de thèse en m'investissant dans la résolution de la structure du virus de la mosaïque de l'arabette (Arabidopsis mosaic virus ou ArMV). J'ai exploré les différents paramètres qui pouvaient avoir une influence sur la résolution finale de la reconstruction de la



capside de l'ArMV, puis défini une stratégie utilisant des approches développées en cristallographie aux rayons X afin de calculer un modèle atomique de l'ArMV. L'équipe de Dr Patrick Bron avait calculé il y a quelques années une structure de la capsid de l'ArMV par cryo-ME et traitement d'image à 6.5 Å de résolution. À partir des images enregistrées sur un microscope Titan Krios équipé d'une caméra à détection directe d'électrons, j'ai pu obtenir une reconstruction à 3.8 Å de résolution.

En parallèle de ce travail, le séquençage du virus contaminant a permis de montrer qu'il ne s'agissait pas du CPMV, mais du virus de la tâche de la fève (Broad Bean Stain Virus ou BBSV), un virus appartenant également à la famille des *Comoviridae*, mais dont la structure atomique n'était pas connue. De par l'expertise acquise au travers de l'analyse de l'ArMV, j'ai donc repris le traitement des images du BBSV, puis calculé une carte de densité de la particule virale du BBSV à 3.8 Å de résolution. J'ai ainsi pu calculer un modèle atomique et explorer l'organisation structurale de ce virus.

Comme initialement prévue, je me suis également consacrée à la caractérisation structurale du CaMV et de la protéine virale P2 impliquée dans la transmission par le puceron. Les protéines virales P1, P2 et P3 de ce pararétrovirus contiennent des domaines en hélices super-enroulées leur permettant de s'auto assembler et de s'agréger. Ces caractéristiques physico-chimiques empêchent une approche structurale classique par cristallographie aux rayons X. Afin d'obtenir des informations sur l'organisation structurale de la protéine P2, il a donc fallu envisager une approche hybride combinant cryo-ME et cristallographie aux rayons X. La thèse initialement organisée autour de la caractérisation structurale de deux virus de plante (FBNSV et CaMV) et des protéines virales impliquées dans leur transmission par les pucerons a donc évolué au cours de ces trois ans, afin d'y intégrer l'étude de l'organisation structurale de la capsid des virus ArMV et BBSV. Dans ce contexte, mes résultats seront organisés en 4 chapitres :

- Dans un premier chapitre, je présenterai les travaux que j'ai effectués pour la caractérisation structurale de la capsid du FBNSV ainsi que les approches visant à exprimer et purifier les différentes protéines virales, notamment la protéine impliquée dans la transmission par le puceron ;
- Dans le chapitre suivant, j'aborderai le travail concernant la caractérisation structurale de l'ArMV, en insistant sur la méthodologie utilisée au cours de cette étude et qui m'a permis de calculer la structure atomique de ce virus ;

- Le troisième chapitre sera dédié aux travaux que j'ai réalisés sur le BBSV et qui m'ont permis de calculer un modèle atomique des deux protéines constituant la capsid de ce virus ;
- Le dernier chapitre présentera les travaux que j'ai effectués afin d'améliorer notre connaissance de l'organisation structurale de la capsid du CaMV et de la protéine virale P2 impliquée dans la transmission de ce virus par le puceron.

# CHAPITRE 2

# Chapitre 2 : Étude structurale du virus du retard de croissance et de la nécrose de la fève

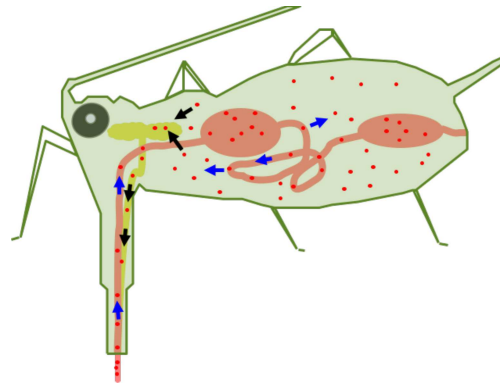
---

## 1. Introduction

Les *Nanoviridae* sont une famille de virus de plantes composée de deux genres, *Babuvirus* et *Nanovirus*. Ces deux genres infectent les Angiospermes, aussi appelées « plantes à fleurs », qui présentent le système reproductif le plus évolué du monde végétal : les fleurs et les fruits. Les Angiospermes sont divisées en deux groupes, les Monocotylédones (la plantule ne possède qu'un seul cotylédon) et les Dicotylédones (la plantule possède deux cotylédons), les cotylédons étant des structures de réserve de la graine des plantes utilisées pendant la germination de la graine. Le genre *Babuvirus* infecte les plantes monocotylédones (et touche en particulier les bananes) alors que le genre *Nanovirus* infecte les plantes dicotylédones (qui comprend les plantes légumineuses). Les *Nanoviridae* possèdent un génome circulaire simple brin qui est segmenté en plusieurs molécules d'ADN d'environ 1000 nucléotides. La particularité de leur matériel génétique leur a valu le nom de virus « multipartite », car à la différence des virus segmentés, ici chaque segment dispose de sa propre capsid virale (particule icosaédrique de 18-20 nm). De ce fait, chaque virus ne contiendra qu'un seul segment, et demandera la présence des autres segments pour se répliquer. Par ailleurs, un segment code pour une seule protéine, à l'exception du virus du sommet touffu du bananier (*Banana Bunchy Top Virus* ou BBTV), où un même segment code pour deux protéines différentes (Timchenko & Bernardi 2007). La catégorie des virus multipartites est très représentée chez les végétaux, puisqu'elle représente plus d'un tiers des virus décrits, mais n'a pas encore été observée dans le monde animal.

Le virus utilisé comme modèle d'étude des *Nanoviridae* est le virus du retard de croissance et de la nécrose de la fève (*Faba Bean Necrotic Stunt Virus* ou FBNSV), isolé pour la première fois en 2009 en Éthiopie sur l'espèce de fève *Vicia faba*. Tout comme les autres *Nanoviridae*, le FBNSV est transmis par les pucerons de façon circulante non propagatrice. Ce mode de transmission signifie que le virus circule à l'intérieur du vecteur sans s'y répliquer. Lorsque le vecteur acquiert le virus en se nourrissant du phloème d'une plante infectée, le virus est ingéré et fini par atteindre l'intestin. Le virus traverse alors la paroi intestinale puis diffuse dans l'hémolymphe pour revenir aux glandes

salivaires. Lors d'une nouvelle phase de nutrition par le puceron, la salivation va déclencher la libération des virus dans le phloème de la nouvelle plante (**Figure 28**).

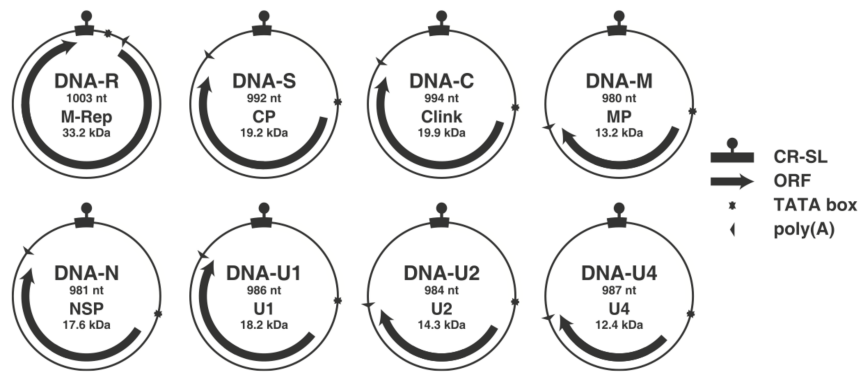


**Figure 28 : Dessin simplifié du mode de transmission circulant non propagatif.** Les ronds rouges sont les particules virales. Les flèches bleues sont le trajet des particules virales venant d'une plante infectée. Les flèches noires sont le trajet des particules virales lors de leur entrée dans les glandes salivaires, puis leur libération vers une nouvelle plante.

Le FBNSV peut être transmis par différentes espèces de pucerons (e.g *Myzus persicae*, *Aphis craccivora*) et infecte différentes plantes légumineuses en plus des fèves, comme le pois. Chez la plante infectée, il induit alors des chloroses, l'enroulement des feuilles, le nanisme et la mort prématurée de la plante.

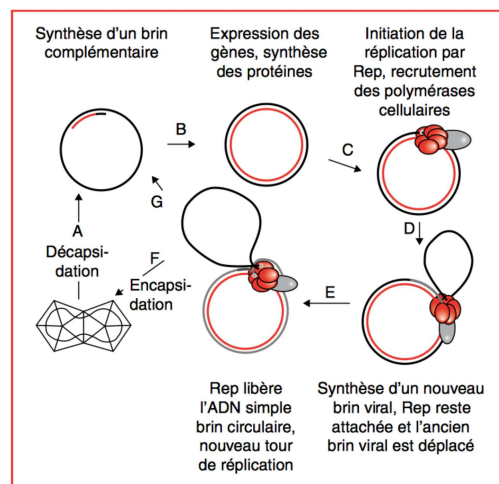
Le génome du FBNSV est divisé en huit segments distincts, dont six ont une fonction connue à ce jour (Grigoras et al. 2009) (**Figure 29**) :

- Le segment S code pour la protéine de capsid ;
- Le segment N code pour une protéine navette permettant le transport de l'ADN nouvellement synthétisé vers le cytoplasme, et est décrite comme la protéine de transmission puisqu'elle est nécessaire pour la transmission du virus à son vecteur ;
- Le segment C code pour la protéine interférant avec la régulation du cycle cellulaire ;
- Le segment M code pour la protéine de mouvement permettant le déplacement cellule à cellule ;
- Le segment R code pour la protéine initiant la réplication des segments d'ADN ;
- Les segments U1, U3 et U4 codent pour des protéines ayant des fonctions totalement inconnues.



**Figure 29 : Organisation génomique du FBNSV.** Les tailles de chaque segment d'ADN et de leur protéine correspondante sont indiquées respectivement en nombre de nucléotides et en kDa. Les flèches représentent les cadres ouverts de lecture et CR-SL les régions communes de la tige-boucle. Reproduit de (Grigoras et al. 2009).

La réplication du génome se fait selon le mécanisme du cercle roulant (**Figure 30**), dans le noyau des cellules du phloème où le virus est restreint. Le phloème correspond au tissu conducteur de la sève élaborée de la plante, qui est une solution riche en glucides utilisée pour la croissance de la plante.



**Figure 30 : Réplication des virus à ADN simple brin selon le mécanisme du cercle roulant.** Après décapsidation (étape A), l'ADN simple brin est converti en ADN double brin par extension d'amorce (étape B), le brin néosynthétisé est représenté en rouge. La transcription et la traduction des gènes viraux permettent l'induction de la réplication de l'ADN viral sous forme répllicative double brin. La protéine Rep (un oligomère formé de plusieurs molécules, en rouge) introduit une coupure au niveau de l'origine de réplication (sur un brin viral, en noir) et recrute des enzymes cellulaires de polymérisation de l'ADN (en gris) (étape C). Les enzymes cellulaires vont allonger le brin coupé donnant le brin néosynthétisé (en gris, étape D). La protéine Rep, accrochée à l'extrémité 5' de l'ADN par une tyrosine (le point blanc), fermera le cercle d'ADN déplacé (étape E).

L'ADN simple brin pourra être encapsidé (icosaèdre simple pour les nanovirus, étape F) ou à nouveau transformé en double brin (étape G). Reproduit de (Timchenko & Bernardi 2007).

Le système multipartite montre un grand désavantage pour la transmission du matériel génétique, car pour maintenir l'intégrité du génome, il faudrait que l'ensemble des brins se retrouvent dans la même cellule. Le taux d'infection doit alors être très élevé pour qu'une cellule contienne au moins un brin de chaque segment. Le coût des virus multipartites est d'autant plus élevé que la probabilité de perte d'une partie de l'information génétique est directement liée au nombre de segments. Pour s'ajouter à cela, de récentes études ont montré que chaque segment n'est pas représenté avec la même fréquence (Sicard et al. 2013). En effet, cette étude démontre l'existence d'une formule génomique reproductible au sein d'une plante infectée, c'est-à-dire une fréquence relative à chaque segment. Cette formule génomique expose une grande différence de fréquences entre les segments, avec un ratio allant jusqu'à 1:15 entre les segments le moins et le plus fréquent. La formule génomique est spécifique à une espèce de plante donnée, suggérant une plasticité du virus pour son environnement. De plus, un réarrangement de cette formule génomique se passe lors du passage de la plante infectée vers le puceron. Cette nouvelle formule génomique reste la même avec différentes espèces de pucerons nourris sur même plante infectée, laissant suggérer une relation plus étroite avec le vecteur que ce qui a été décrit jusqu'à présent (Sicard et al. 2015). Pour finir, des études préliminaires réalisées par l'équipe de Stéphane Blanc (CIRAD, UMR BGPI) ont montré que les segments ne sont pas toujours colocalisés dans les cellules, laissant supposer que le maintien de l'intégrité génomique ne dépend peut-être pas d'une co-infection systématique de tous les segments dans la même cellule. Il pourrait alors exister une séparation spatiale des segments avec un échange à distance des produits viraux.

L'étude des *Nanoviridae* présente un enjeu de taille, le BBTV est le virus causant le plus de dommages sur les cultures de bananiers en Afrique, en Asie du Sud-Est et dans le Pacifique Sud (Blomme et al. 2013). Le genre *Nanovirus* s'attaque, quant à lui, à de nombreuses légumineuses, et a émergé en Afrique (Abraham et al. 2010), en Asie (Kumari et al. 2010), et récemment en Europe (Grigoras et al. 2014). Devant le manque de données au niveau moléculaire de leur mode de fonctionnement, mon travail sur ce sujet a été d'essayer de caractériser la structure des protéines du FBNSV, et en particulier de la protéine capsidique S et de la protéine N. Dans cette optique, une approche intégrée alliant cryo-ME et cristallographie aux rayons X a été envisagée pour obtenir des structures à hautes résolutions. La cryo-ME étant destinée à l'étude de la capsidique du virus exprimée *in planta* ou *in vitro*, et la cristallographie pour l'étude *in vitro* des segments d'ADN qui codent pour les protéines de capsidique et de transmission.

## 2. Matériel et méthodes

### 2.1. Expression de protéines recombinantes dans les bactéries *Escherichia coli* (*E. coli*)

#### 2.1.1. Les différents plasmides

Les séquences d'ADN codantes de la protéine de capsid S et de la protéine de transmission N du FBNSV ont été clonées dans le plasmide pETM-41 aux sites BamHI-NcoI, en aval d'une protéine porteuse MBP (*Maltose Binding Protein*), utilisée pour améliorer la solubilité des protéines, d'une étiquette Histidine (His) pour une purification par chromatographie et sous le contrôle du promoteur T7, reconnu par la polymérase du bactériophage T7. De cette manière des protéines de fusion His-MBP-N/S seront synthétisées, et un site de reconnaissance à la protéase du virus de la gravure du tabac (*Tobacco Etch Virus* ou TEV) présent entre les segments His-MBP et N/S permettra de séparer la protéine de fusion et l'étiquette Histidine (**Figure 31**).

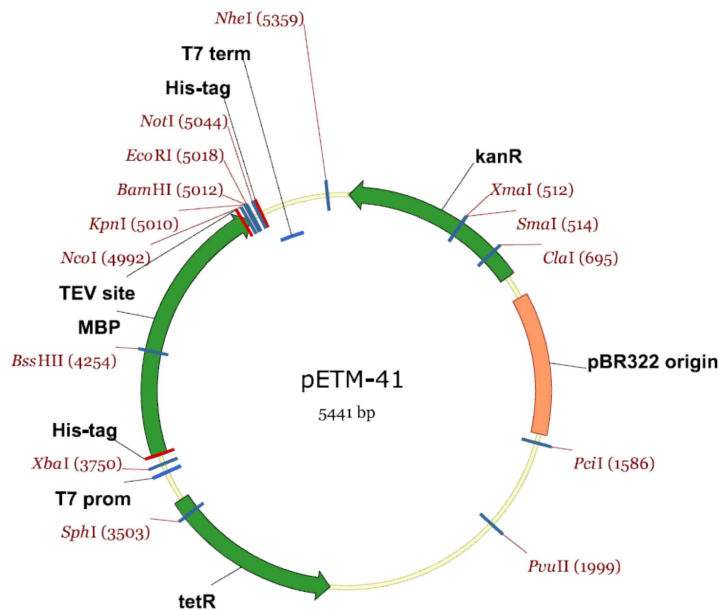


Figure 31 : Carte du plasmide pETM-41 utilisé pour la production des protéines recombinantes en *E. coli*.

Par la suite, les séquences virales ont été également clonées dans trois autres plasmides fournis par le Laboratoire de Biologie Moléculaire Européen (*European Molecular Biology Laboratory* ou EMBL).



Les trois nouveaux plasmides possèdent la même cassette d'expression que pETM-41, mais avec une protéine porteuse différente. Ainsi, les mêmes enzymes de restrictions BamH1-Nco1 ont été utilisées pour insérer l'ADN viral dans les plasmides pour produire quatre protéines de fusion différentes :

- His-MBP-segment viral
- His-GST-segment viral (*Glutathione S-transferase*)
- His-NusA-segment viral
- His-Trx-segment viral (*Thioredoxin*)

### 2.1.2. Les différentes souches *E. coli*

Chaque plasmide a ensuite été transformé dans différentes souches de bactérie *E.coli* rendues chimio-compétentes au laboratoire, qui contiennent toute le gène codant pour la polymérase du phage T7 (souche DE3) sous le contrôle de l'opéron lactose, ainsi qu'un plasmide possédant le gène codant pour le lysozyme du phage T7 pour réduire l'expression basale de la polymérase T7 (plasmide pLysS). Les souches possèdent des modifications permettant un taux d'expression élevée de protéines recombinantes et sont moins susceptibles de dégrader les protéines pendant la purification :

- BL21 (DE3) pLysS : souche déficiente des protéases *Lon* et *OmpT* (qui peuvent dégrader les protéines recombinantes) ;
- Star (DE3) pLysS (Dérivée de la souche BL21) : possède une mutation supplémentaire du gène *RNaseE* impliqué dans la dégradation d'ARN messagers ;
- Rosetta (DE3) pLysS (Dérivée de la souche BL21) : possède également les gènes codants pour les ARN de transfert reconnaissant les codons rares chez *E.coli* ;
- C41 (DE3) pLysS (Dérivée de la souche BL21) : possède des mutations supplémentaires qui confèrent une meilleure tolérance aux protéines toxiques ;
- Origami (DE3) pLysS : souche non dérivée de la souche BL21, mais de la souche K-12 qui possède des mutations au niveau des gènes codant pour les réductase glutathione et thioredoxin, ce qui permettent d'améliorer la formation des ponts disulfures des protéines dans le cytoplasme des bactéries.

### 2.1.3. Les conditions de culture

Un milieu standard et un milieu auto-inductible ont été utilisés pour la croissance des bactéries :

- Le milieu *lysogeny broth* (LB) : composé d'extrait de levure, de chlorure de sodium et de tryptone. C'est un milieu standard utilisé en microbiologie moléculaire pour la préparation de plasmides d'ADN et de protéines recombinantes. En utilisant ce milieu, l'induction de la synthèse des protéines recombinantes est déclenchée par ajout d'isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), un analogue de l'allolactose qui va activer la transcription de l'opéron lactose. Au cours de ces expériences, l'ajout d'IPTG s'est fait quand une densité optique comprise entre 0,6 et 1 a été atteinte.
- Le milieu ZYM-5052 : milieu auto-inductible mis au point par (Studier 2005). L'auto-induction repose sur un équilibre glucose/lactose qui permet d'induire l'expression de la protéine recombinante une fois que les cellules arrivent à saturation.

Les bactéries ont été cultivées à différentes températures pour cribler les meilleures conditions de solubilité :

- Croissance à 37°C jusqu'à une densité optique comprise entre 0,6 et 1 puis 3 h supplémentaires à 37°C ;
- Croissance à 37°C jusqu'à une densité optique comprise entre 0,6 et 1 puis toute la nuit à 17°C ;
- Croissance à 37°C jusqu'à une densité optique comprise entre 0,6 et 1 puis toute la nuit à 25°C.

### 2.1.4. Purification de la protéine

Les bactéries ont été lysées par congélation et sonication avec ajout de lysozyme et de DNase (0,1 mg/ml) dans un tampon Tris 50 mM, NaCl 150 mM et  $\beta$ -mercaptoéthanol 2mM à pH 7,5 (le point isoélectrique de la protéine S étant 10 et celui de la protéine N étant 5,8). Après lyse, un aliquot a été prélevé pour mesurer le taux d'expression total des protéines, puis les bactéries ont été centrifugées, et le surnageant séparé du culot afin de vérifier si les protéines étaient solubles ou pas. Chaque fraction a ensuite été observée par électrophorèse SDS-PAGE (électrophorèse en gel de

polyacrylamide contenant du dodécylsulfate de sodium) qui consiste à faire migrer des protéines dans un gel, sous l'influence d'un champ électrique pour les séparer en fonction de leur poids moléculaire, afin de vérifier l'expression de la protéine recombinante dans chaque fraction.

## **2.2. Expression de protéines recombinantes dans les cellules d'insectes *Spodoptera frugiperda* 9 (Sf9)**

Des baculovirus recombinants (*AcSLP10*) exprimant les gènes S et N avec une étiquette Histidine ont été commandés chez la société Agate bioservices®. Les cellules Sf9 ont été cultivées à 29°C et récoltées 72 h après infection. Les cellules ont été culottées et resuspendues dans différents tampons avant d'être lysées sur glace par agitation mécanique. Après lyse, les cellules ont été centrifugées et le surnageant séparé du culot. Chaque fraction a ensuite été déposée sur SDS-PAGE pour séparer les protéines en fonction de leur poids moléculaire et vérifier l'expression de la protéine recombinante.

## **2.3. Caractérisation de la protéine de capsid par cryo-ME**

### **2.3.1. Purification du virus *in planta***

Le virus FBNSV a été purifié à partir de 100 g de feuilles de fèves infectées. Les feuilles ont été broyées dans un tampon citrate 20 mM, 20 mM Na<sub>2</sub>SO<sub>3</sub> à pH 6. Puis 0,1 % de Macerozyme a été ajouté et la solution a été agitée toute une nuit à 4°C. La solution a ensuite été filtrée et incubée sous agitation avec 25 % d'un mélange de chloroforme et d'isobutanol (ratio 1:1) pendant 30 min à 4°C. La solution a ensuite été centrifugée à 15 200 g pendant 20 min à 4°C. Les virus contenus dans la phase aqueuse ont été centrifugés à 302 000 g pendant 3 h à 4°C. Le précipité a été resuspendu dans un tampon citrate 20 mM à pH 6 et déposés sur un gradient discontinu de chlorure de césium (10%-15%-20%-30%-40%) qui a été centrifugé à 201 000 g pendant 4 h à 4°C pour sédimenter la bande virale. La bande virale a été extraite et centrifugée une dernière fois à 302 000 g pendant 3h à 4°C avant d'être resuspendue dans 20 µl de tampon 20 mM phosphate à pH 7,4 avec une concentration finale de 2 mg/ml.

### 2.3.2. Congélation de l'échantillon

Des grilles commerciales Quantifoil R2/2 (*Quantifoil Instruments GmbH*) ont été utilisées pour la congélation du virus. Ce sont des grilles de cuivre recouvertes d'un film de carbone, et parsemées de trous d'un diamètre de deux microns, espacés chacun de deux microns. Les grilles sont tout d'abord placées dans une chambre connectée à un générateur électrique (*Easy Glow, PELCO*) qui va ioniser l'air dans la chambre, permettant le dépôt d'ions négatifs sur le carbone de la grille, rendant sa surface hydrophile. Une grille est ensuite placée dans la chambre du système de congélation CP3 (Gatan) (**Figure 32**), maintenue à un taux d'humidité 99%. 3  $\mu$ l de l'échantillon ont été déposés sur la grille. L'excès de liquide a été retiré par absorption en pressant deux papiers de chaque côté de la grille pendant 2 secondes. La grille a ensuite été plongée presque instantanément dans de l'éthane liquide refroidi à température de l'azote liquide.



**Figure 32 : Système de congélation Gatan CP3.**

### 2.3.3. Acquisition des images

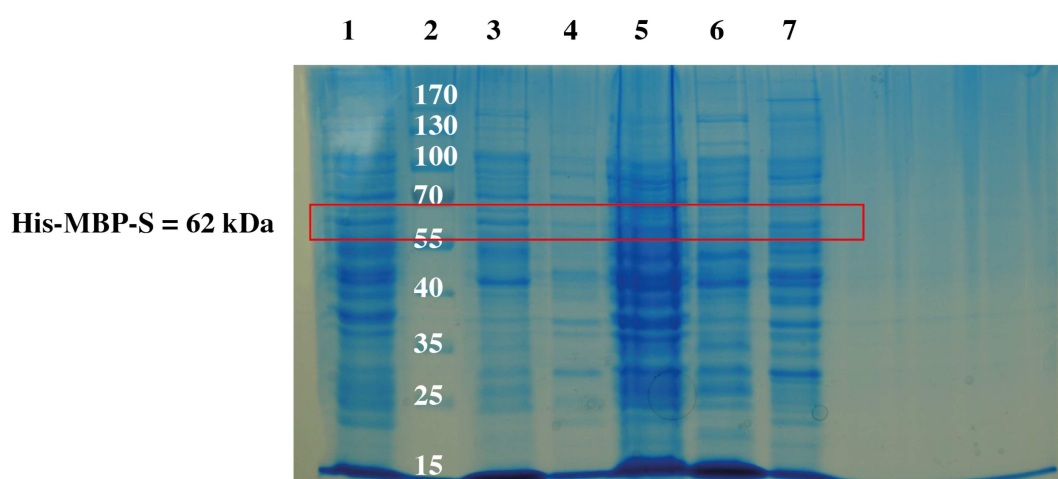
Les premières images de microscopie ont été acquises sur un microscope JEOL 2200FS au CBS (Centre de Biochimie Structurale) à Montpellier, avec une tension accélératrice de 200 kV et équipé d'un canon à émission champs, d'un filtre Oméga et d'une caméra CCD 4k x 4k (Gatan). Une seconde acquisition a été faite sur un Titan Krios (IGBMC, Strasbourg). Le microscope était équipé d'un correcteur d'aberration sphérique, d'une caméra à détection directe Falcon II et du logiciel d'acquisitions automatiques EPU (logiciel propriétaire du constructeur FEI). Le microscope permet d'enregistrer une série de vues qui vont former un film. Un total de 2869 films a été enregistré à un

grossissement x 120000, chacun correspondant à une irradiation totale de  $40 \text{ e}^-/\text{\AA}^2$ . La machine permet de sauvegarder un nombre maximum de 7 vues, nous avons éliminé les deux premières vues qui sont non exploitables à cause de l'instabilité de la caméra. Sur un total de 35 vues enregistrées pour une même acquisition, seuls les vues 3 à 9 ont été sauvegardées pour le traitement d'image. À chaque vue correspond à une dose d'irradiation de  $3,5 \text{ e}^-/\text{\AA}^2$ .

## 3. Résultats

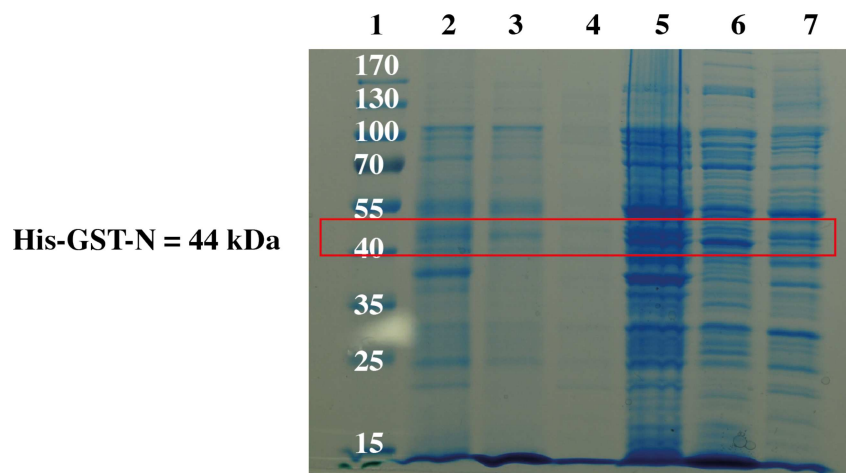
### 3.1. Expression de protéines recombinantes dans la bactérie *E. coli*

Les différents segments ont été insérés dans les plasmides et transformés dans les bactéries. Malgré les nombreuses conditions testées (plasmides, bactéries, milieux et températures de culture), les tests d'expression des protéines recombinantes chez *E.coli* à différentes températures n'ont pas permis de visualiser une surexpression des protéines N et S. Deux exemples d'électrophorèses SDS-PAGE révélées par bleu de Coomassie sont donnés en **Figure 33** et **34**. L'absence de bandes visibles pour les protéines recombinantes, même dans la fraction totale, a empêché l'identification de la meilleure combinaison plasmide/bactérie/milieux de culture vers laquelle se concentrer pour continuer la purification des protéines. L'expression de la protéine étant probablement trop faible pour être détectée par SDS-PAGE.



**Figure 33 : Expression de la protéine de fusion His-MBP-S produite chez la souche BL21 en milieu d'auto-induction. (2) Marqueur de poids moléculaire (kDa). (1, 3 et 4) Culture des bactéries à 37°C pendant 7 h, le dépôt des fractions correspondant respectivement au total, culot, et surnageant. (5, 6 et 7) Culture des**

bactéries à 37°C pendant 3 h, puis 17°C toute la nuit, le dépôt des fractions correspondant respectivement au total, culot, et surnageant.



**Figure 34 : Expression de la protéine de fusion His-GST-N produite chez la souche C41 en milieu d'auto-induction. (1) Marqueur de poids moléculaire (kDa). (2, 3 et 4) Culture des bactéries à 37°C pendant 7 h, le dépôt des fractions correspondant respectivement au total, culot, et surnageant. (5, 6 et 7) Culture des bactéries à 37°C pendant 3 h, puis 17°C toute la nuit, le dépôt des fractions correspondant respectivement au total, culot, et surnageant.**

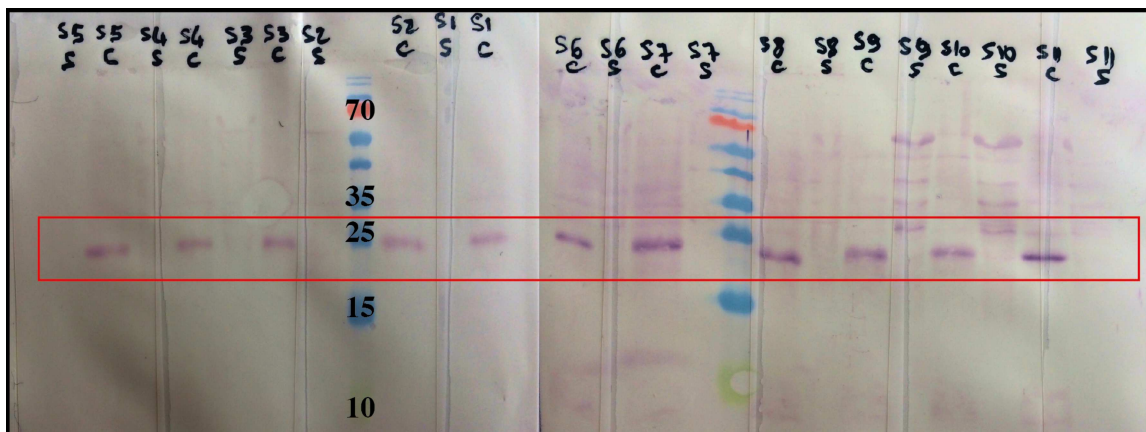
## 3.2. Expression de protéines recombinantes dans les cellules d'insecte *Sf9*

Les tests d'expression des protéines recombinantes en cellules d'insecte n'ont également pas montré de surexpression des protéines. Cependant, ces tests d'expression se sont faits plus tard que l'expression des segments viraux chez *E.coli*. La mise à disposition d'un générateur électrique pour western blot nous a permis de vérifier l'expression des protéines en utilisant un anticorps contre l'étiquette histidine. Par rapport à l'expression en bactérie, l'expression en cellules *Sf9* limite les possibilités de stratégies d'expression, seul un type de plasmide est utilisé et transfecté dans un seul type de cellule. La température de culture des cellules est également fixe et seul le temps de récolte des cellules après infection par les baculovirus varie. Nos collaborateurs du laboratoire BGPI (CIRAD) avaient montré qu'un temps de récolte à 72h après infection donnait le meilleur taux d'expression totale des protéines (données non montrées). Les protéines étant exprimées, les paramètres restant pour obtenir des protéines solubles étaient le tampon utilisé pour la lyse des cellules. C'est pourquoi un criblage de la solubilité des protéines s'est fait avec 12 tampons différents qui ont des différences au niveau du pH, de la force ionique et de la présence ou non de détergent (**Figure 35**). Malgré une

expression des protéines recombinantes, aucun tampon n'a permis d'obtenir les protéines dans la fraction soluble, qui restent concentrées dans le culot de cellules (**Figures 36 et 37**).

1	Acétate 50 mM pH 4	NaCl 5 mM	
2		NaCl 500 mM	
3		NaCl 5 mM	Triton X100 1%
4		NaCl 500 mM	Triton X100 1%
5	MES 50 mM pH 6	NaCl 5 mM	
6		NaCl 500 mM	
7		NaCl 5 mM	Triton X100 1%
8		NaCl 500 mM	Triton X100 1%
9	Tris 50 mM pH 8	NaCl 5 mM	
10		NaCl 500 mM	
11		NaCl 5 mM	Triton X100 1%
12		NaCl 500 mM	Triton X100 1%

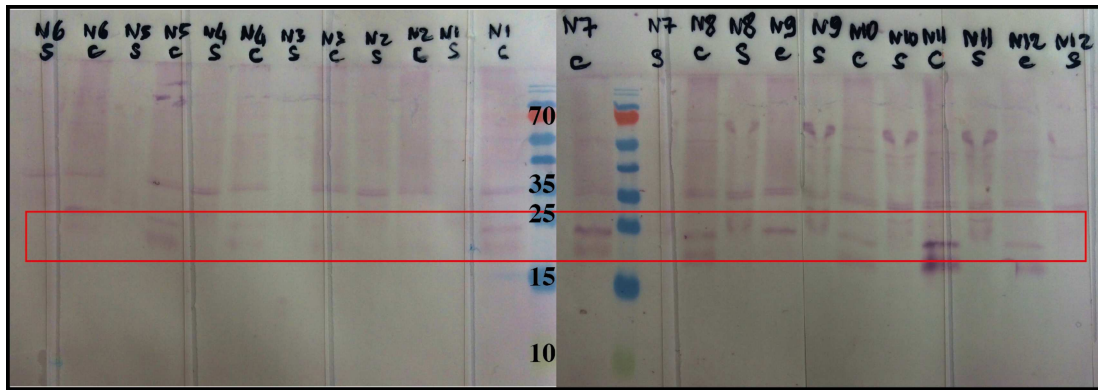
**Figure 35** : Composition des 12 tampons utilisés pour l'expression des segments viraux du FBNSV.



**His-S = 20 kDa**

**Figure 36** : Expression de la protéine de fusion His-S produite chez les cellules *Sf9*. Les nombres correspondent au tampon utilisé (se référer à la Figure 34). C correspond au culot cellulaire et S au surnageant cellulaire après lyse et centrifugation.



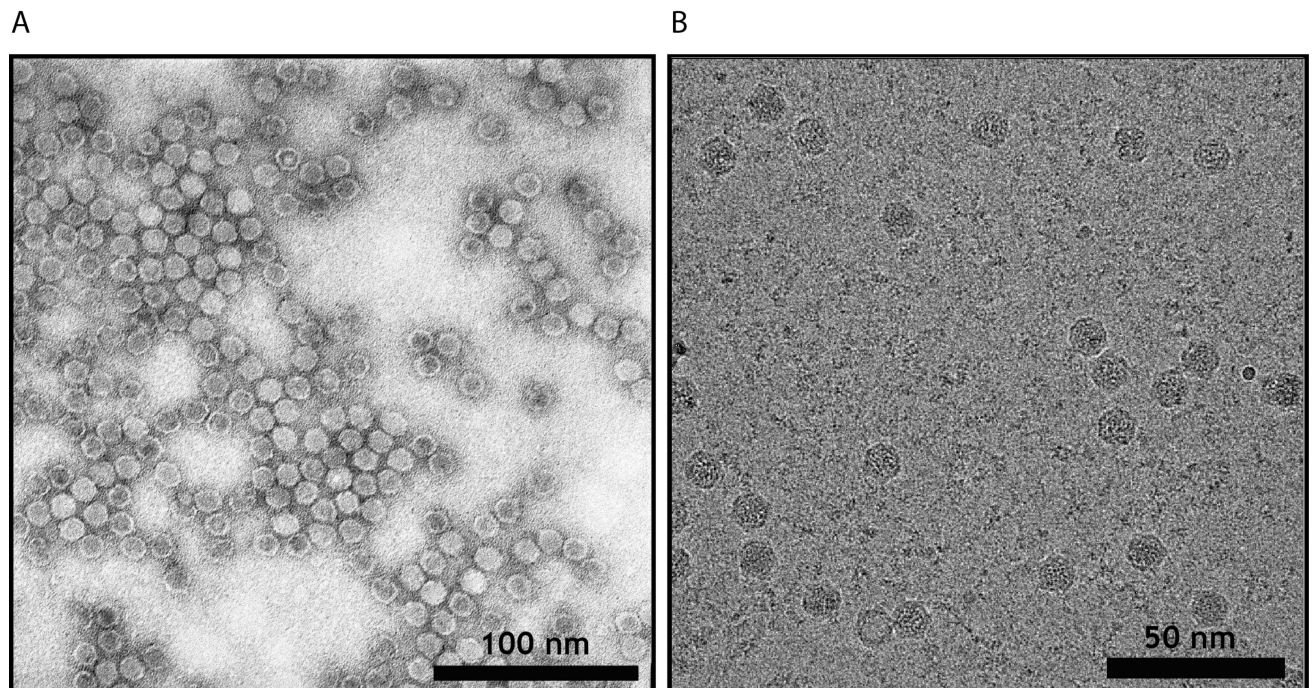


His-N = 18,5 kDa

**Figure 37** : Expression de la protéine de fusion His-N produite chez les cellules *Sf9*. Les nombres correspondent au tampon utilisé (se référer à la Figure 34). C correspond au culot cellulaire et S au surnageant cellulaire après lyse et centrifugation.

### 3.3. Caractérisation de la protéine de capside par cryo-ME

L'étude de la protéine de capside par cryo-ME s'est faite à partir de la purification *in planta* du virus dont la première image de microscopie a été réalisée par nos collaborateurs (CIRAD, UMR BGPI) en coloration négative (**Figure 38A**). Ce même échantillon a permis d'autres séries d'acquisitions en condition cryogénique (**Figure 38B**).



**Figure 38** : Images de microscopie électronique à transmission du virus FBNSV. (A) Image en coloration négative. (B) Image en condition cryogénique.



## 4. Conclusion

Le sujet reste difficile puisque l'expression des protéines de capsid et de transmission du FBNSV n'a pas abouti sur la purification de protéines solubles. La production des protéines de capsid exprimées dans *E.coli* n'a pas montré de surexpression et aucun western blot n'avait été fait pour nous aider à trouver les premières pistes. La production des protéines en *E.coli* avait été laissée de côté pour se concentrer sur la caractérisation structurale du virus purifié directement chez la plante par cryo-ME. La résolution de la structure par cryo-ME avait alors demandé un temps d'apprentissage conséquent du traitement d'image avant de pouvoir commencer à traiter les données. Néanmoins, ce virus s'est avéré être un virus contaminant présent dans le lot de graines de fève utilisées, la caractérisation structurale de ce nouveau virus est décrite Chapitre 4. Des western blot ont pu être réalisés pour l'étude de l'expression des protéines virales exprimées dans les cellules *Sf9*, mais n'ont pas montré la présence de protéines solubles. Ce projet demande encore un temps d'optimisation des protocoles de purification avant de pouvoir exprimer des protéines solubles. Cependant, des western blot seront fait pour repérer l'expression des protéines chez *E.coli*, même si aucune surexpression n'a été visible en bleu de Coomassie, la sensibilité de détection qu'apporte les western blot devrait permettre d'identifier quelle combinaison plasmide/bactérie/température donne des signes d'expression. L'identification d'une fraction soluble par western blot, même en petite quantité, nous aiguillera sur la meilleure voie à suivre. De cette manière, il sera possible de continuer la purification jusqu'à la séparation de la protéine virale avec sa protéine porteuse. S'il est possible d'obtenir une fraction soluble de la protéine seule, il suffira d'ajuster la quantité de bactéries cultivées en amont de la purification pour pouvoir purifier la protéine virale en quantité suffisante et commencer une première étude par cristallographie aux rayons X.

# CHAPITRE 3

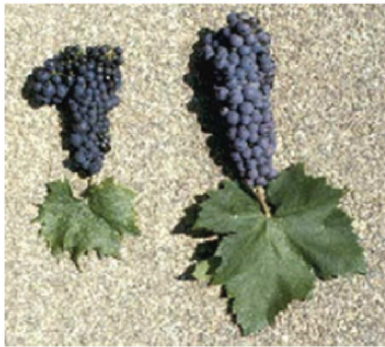
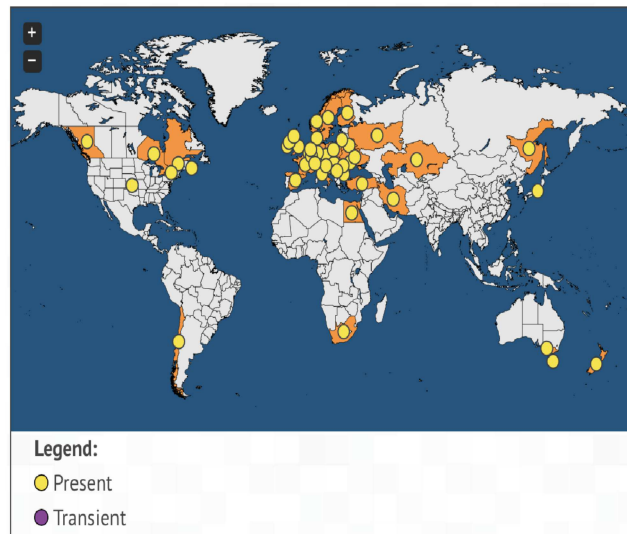
# Chapitre 3 : Étude structurale du virus de la mosaïque de l'arabette

---

## 1. Introduction

Le virus de la mosaïque de l'arabette (ArMV) est un virus de la sous-famille des *Comovirinae* appartenant au genre *Nepovirus*, responsable de la maladie du court-noué qui infecte les vignes. Il a été découvert pour la première fois en 1944 dans une serre à Cambridge par Smith et Markham. L'ArMV est un virus à symétrie icosaédrique contenant un génome ARN bipartite, sa capsid virale possède une triangulation  $T=1$ , elle est composée de 60 copies d'une même protéine structurale formant une particule de 30 nm de diamètre. La capsid contient un génome ARN composé de deux segments simples brins à polarité positive : ARN1 et ARN2. L'ARN1 jouant un rôle dans la réplication du virus et l'ARN2 dans le mouvement du virus et encode sa protéine de capsid.

L'ArMV est transmis par les nématodes qui acquièrent le virus en se nourrissant au niveau des racines en croissance d'une vigne. Le nématode va ingérer les virions d'une plante infectée qui seront retenus spécifiquement au niveau de son appareil alimentaire, les virions seront relâchés plus tard lorsque le nématode se nourrira de nouveau sur une autre plante (Demangeat 2007). La transmission se fait seulement de proche en proche puisque les nématodes se déplacent seulement de quelques dizaines de centimètres par an. Cependant, la longévité des nématodes peut dépasser 5 ans grâce à un passage en forme quiescente, leur permettant de survivre en l'absence de racines et de conserver leur pouvoir infectieux (Demangeat et al. 2005). L'ArMV infecte les grappes de raisins provoquant une dégénération de la plante et des fruits (**Figure 39A**), il est responsable d'importantes pertes économiques pour les vignerons partout dans le monde (**Figure 39B**).

**A****B**

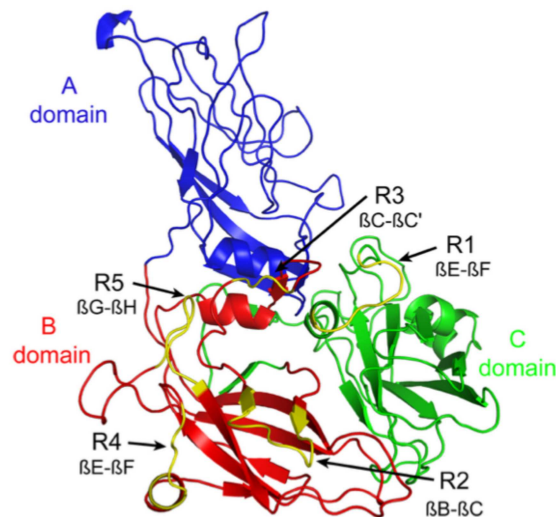
**Figure 39 : L'impact du virus de l'ArMV. (A)** Grappes de raisins issues d'une plante infectée (gauche) et saine (droite) (Martelli 2014). **(B)** Distribution mondiale des pays touchés par le virus de l'ArMV (epppo.int)

Les moyens de luttés actuels restent inefficaces puisqu'il ne suffit pas d'arracher tous les pieds infectés d'une culture pour se débarrasser du virus. En effet, il est possible de trouver un million de nématodes dans un mètre cube de terre, s'ils sont déjà porteurs du virus, la nouvelle culture saine sera infectée de nouveau.

Les virus de plantes possèdent une spécificité plus ou moins grande pour leurs vecteurs. Cette spécificité va jouer un rôle important dans la distribution du virus puisqu'un même virus peut très bien être transmis par différentes espèces ou par une seule et unique espèce. L'ArMV entre dans cette dernière catégorie, il est exclusivement transmis par le nématode *Xiphinema diversicaudatum*. Cette transmission ne fait intervenir que la capsid virale qui se lie directement au récepteur hypothétique du vecteur, seule stratégie connue pour les Népovirus.

Des études ont été préalablement réalisées pour comprendre la spécificité de transmission entre le virus et le nématode. Afin d'identifier la zone de la capsid virale impliquée dans l'interaction du virus avec le récepteur présent chez le nématode, des études comparatives entre les déterminants structuraux de l'ArMV et du *Grapevine fanleaf virus* (GFLV), un autre *Nepovirus* infectant les vignes, sont réalisées. Ces deux virus partagent une identité de séquence de leur ARN2 entre 60 et 78% (Vigne et al. 2008). Cependant, l'ArMV se transmet uniquement par le nématode *X. diversicaudatum*, alors que le GFLV est uniquement transmis par le nématode *Xiphinema index*.

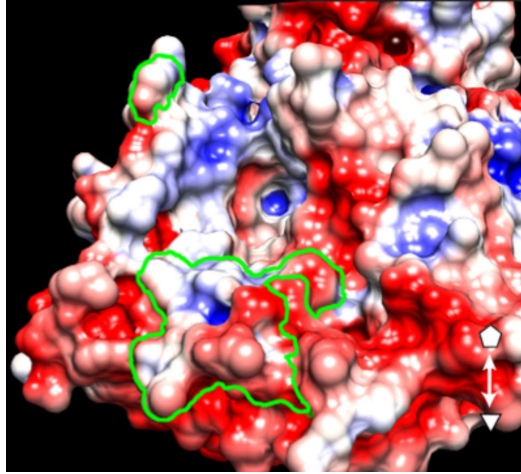
En 2010, un premier modèle 3D du GFLV a été proposé par homologie de structure avec *Tobacco ringspot virus* résolue par cristallographie aux rayons X pour prédire les séquences d'acides aminés impliquées dans la spécificité de transmission. Des tests de transmission d'ARN chimères basés sur ce modèle a permis d'identifier une région, appelé R2, impliquée dans la transmission du GFLV à son vecteur *X. index* (**Figure 40**).



**Figure 40 : Structure de la protéine de capsid du GFLV obtenue par homologie de structure.** Les couleurs bleu, rouge et vert désignent les domaines structuraux. Les couleurs jaunes désignent les régions mutées utilisées les tests de transmission. Reproduit de (Schellenberger et al. 2010).

Lorsque la région R2 du GFLV était remplacée par celle de l'ArMV, le virus perdait sa capacité de transmission à son vecteur *X. index*, mais ne gagnait pas non plus de capacité de transmission au vecteur de l'ArMV. Ce manque de réciprocité suggère qu'une différence de conformation spatiale de cette région R2 existe chez les deux virus.

Un an plus tard, la structure de la capsid du GFLV a été résolue par cristallographie aux rayons X à une résolution de 3,0 Å pour le GFLV sauvage et de 2,7 Å pour un mutant du GFLV avec une substitution de la glycine 297 par un acide aspartique. Ce mutant est apparu spontanément pendant les tests de transmission et entraîne une diminution du taux de transmission à son vecteur de 80% à 8%. Les structures du GFLV ont révélé la présence d'une cavité à la surface de la capsid, proche de l'acide aminé 297 et de la région R2, suggérant que cette région pourrait être le site de liaison avec le récepteur du nématode, de plus, la diminution du taux de transmission pourrait s'expliquer par un changement du potentiel électrostatique (**Figure 41**) (Schellenberger et al. 2011).



**Figure 41 : Représentation du potentiel électrostatique de la capsid du GFLV.** Les lignes vertes délimitent la position de la Glycine 297 en haut et la région R2 en bas. Ces deux régions entourent une poche positive. Reproduit de (Schellenberger et al. 2011)

Dans le but de comparer cette région entre le GFLV et l'ArMV, une première étude structurale de l'ArMV par cryo-ME a permis d'obtenir une structure de la capsid à une résolution estimée de 6,5 Å (Lai-Kee-Him et al. 2013). La comparaison entre cette structure et celle du GFLV précédemment obtenue par cristallographie aux rayons X a montré une distribution différente des charges au niveau du domaine B de la protéine de capsid, suggérant d'autant plus que cette région est responsable de la spécificité de transmission des deux *Néporvirus*. Cependant, la différence de résolution ne permet pas de comparer efficacement ce domaine B, et empêche de proposer une hypothèse viable sur les déterminants structuraux impliqués dans la spécificité de transmission entre les deux virus.

Mon travail sur cette étude a donc été d'utiliser les derniers outils mis à disposition avec les dernières avancées de la cryo-ME pour résoudre la structure de la capsid de l'ArMV à haute résolution. Le cœur de ce projet a été le criblage des différents paramètres liés au traitement d'image dans le but d'obtenir la meilleure résolution possible, au travers de ce qu'il a été possible de mettre en place au niveau local.

## 2. Matériel et méthodes

### 2.1. Congélation de l'échantillon

L'échantillon a été congelé sur des grilles Quantifoil R2/2 (*Quantifoil Instruments GmbH*), avec le même protocole que décrit dans le sous-chapitre 2.3.2 du chapitre 2.

### 2.2. Acquisitions des films

Les images ont été acquises sur le microscope Titan Krios installé au NeCEN (Netherlands Centre for Electron Nanoscopy) à Leiden aux Pays-Bas. Le microscope était équipé d'un correcteur d'aberration sphérique, d'une caméra à détection directe Falcon II et du logiciel d'acquisitions automatiques EPU (logiciel propriétaire du constructeur FEI). Le microscope permet d'enregistrer une série de vues qui vont former un film. Un total de 2680 films a été enregistré à un grossissement  $\times 120000$ , chacun correspondant à une irradiation totale de  $40 \text{ e}^-/\text{\AA}^2$ . Par rapport aux vues enregistrées depuis le Titan Krios de Strasbourg (décrit dans le sous-chapitre 2.3.3), ces vues ont été enregistrées par paire. C'est-à-dire que les vues 3 à 16 ont été enregistrées et combinés de la manière suivante : les vues 3 et 4 ont été sommés ensemble, puis les vues 5 et 6 ont été également sommés ensembles, et ainsi de suite. De ce fait, 7 paires de vues ont été enregistrées, avec une dose de  $7 \text{ e}^-/\text{\AA}^2$  pour une paire. Dans la suite du chapitre, une vue désignera maintenant la somme de deux vues successives.

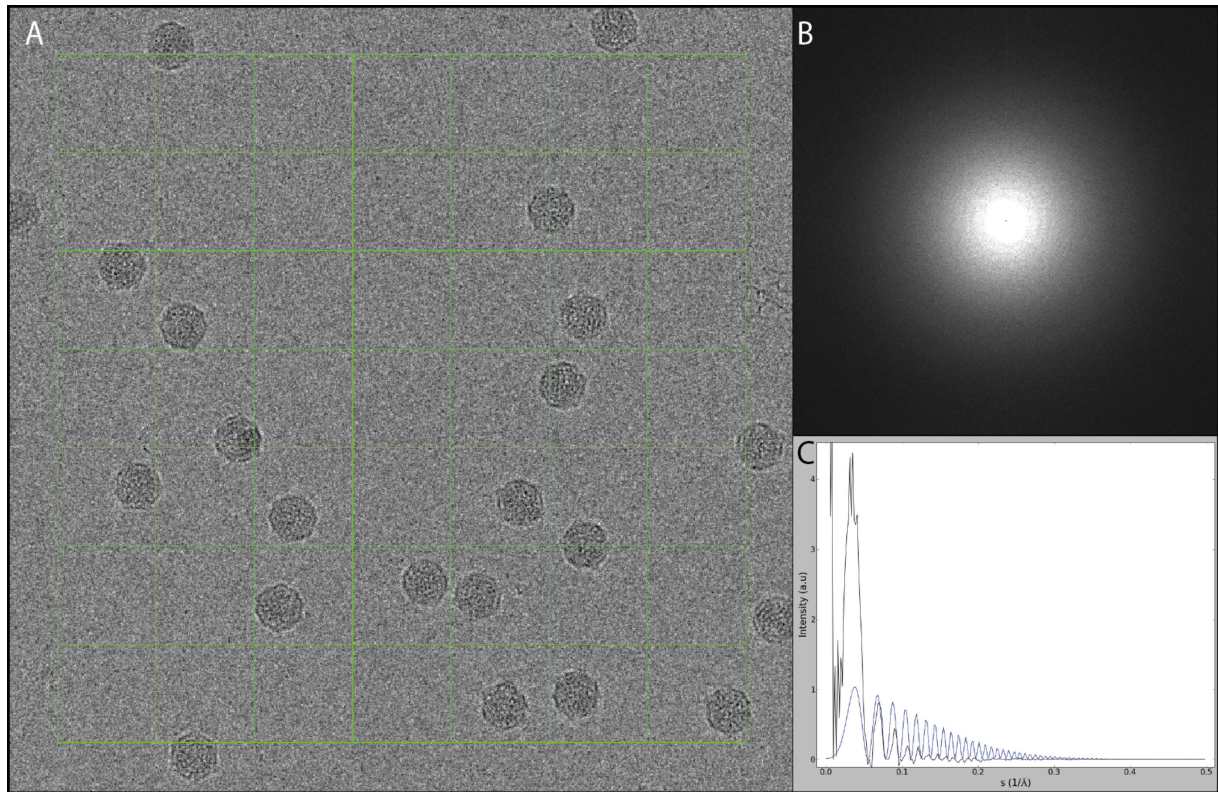
### 2.3. Le traitement d'image et la reconstruction 3D des cartes de densité

Ce chapitre regroupe l'ensemble des méthodes utilisées pour obtenir la structure en 3-dimensions du virus de l'ArMV. Le traitement d'image débute depuis l'analyse des films enregistrés sur le Titan Krios jusqu'à l'obtention de la structure finale.

#### 2.3.1. Le tri des films

Les films ont tout d'abord été triés visuellement, en utilisant le programme `e2evalimage.py` appartenant à la suite de logiciels EMAN2 (Tang et al. 2007) qui permet à la fois de visualiser l'image, sa transformée de Fourier, et l'oscillation de la fonction de transfert de contraste (**Figure 42**).





**Figure 42 : Interface du programme e2evalimage.py permettant de visualiser les films. (A)** L'image du film entier. **(B)** La transformée de Fourier du film. **(C)** Le profil de sa fonction de transfert du contraste qui correspond à la somme de la fonction de transfert du contraste de chaque zone délimitée par des carrés verts.

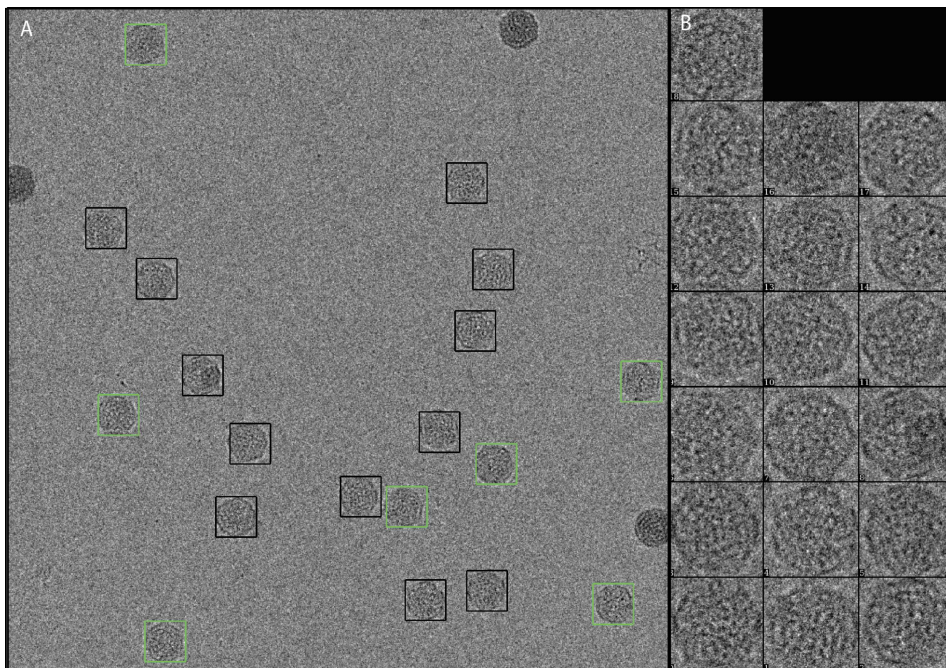
Les paramètres pris en considération pour choisir entre garder ou retirer un film pour la suite du traitement d'image sont :

- L'analyse visuelle de l'image : Les images montrant une mauvaise qualité de l'image sont écartées (glace trop épaisse, absence de glace, présence de contamination, mouvement de l'échantillon).
- L'analyse de la transformée de Fourier : Elle peut mettre plus facilement en évidence un mouvement de l'échantillon.
- Le profil de la fonction de transfert du contraste : Le programme permet de recalculer une courbe théorique de FTC par rapport aux données expérimentales, donnant une idée de la résolution maximale de l'image. Si la courbe tombe à 0 dans les basses fréquences spatiales, notamment à cause des problèmes cités plus haut, l'image est écartée.



### 2.3.2. La sélection des particules des films

La prochaine étape après le tri des films est la sélection des particules. Cette étape consiste à délimiter les régions (qu'on appellera « boîtes ») du film contenant chacune une particule virale. Les particules ont été sélectionnées avec le programme e2boxer.py (EMAN2) qui permet une sélection semi-automatique des particules. Les premières particules sont d'abord sélectionnées manuellement, puis le programme permet une reconnaissance automatique des particules semblables à celles sélectionnées manuellement (**Figure 43**). Les coordonnées de chaque particule sont ensuite écrites dans un fichier texte contenant le centre de chaque boîte ainsi que ses dimensions.



**Figure 43** : Interface du programme e2boxer.py. (A) Les carrés noirs correspondent aux particules sélectionnées manuellement, les carrés verts correspondent aux particules reconnues automatiquement par le programme. Les particules noircies correspondent aux particules manuellement retirées de la sélection. (B) Ensemble des particules sélectionnées dans le film.

### 2.3.3. Calcul de la FTC et des valeurs de défocalisation

Les valeurs de défocalisation des images ont été affinées avec le logiciel Ctfind4 (Rohou & Grigorieff 2015) afin de corriger au mieux la fonction de transfert du contraste. À partir de la valeur de défocalisation initiale du film, une courbe de FTC théorique est générée et comparée à la courbe de

FTC expérimentale. Cette valeur de défocalisation va ensuite être affinée pour que la courbe théorique corresponde au mieux à la courbe expérimentale.

### **2.3.4. La reconstruction 3D du virus**

La résolution de structures par cryo-ME nécessite une quantité de données et un temps de calcul très important pour pouvoir retrouver l'orientation des particules à partir d'un jeu de données très bruité. L'avantage de travailler avec un virus avec une capsidie icosaédrique est qu'elle présente un haut degré de symétrie : l'image d'un même virion contient jusqu'à soixante vues différentes de l'unité asymétrique de la capsidie, l'orientation relative de ces vues étant précisément connue (une fois que l'orientation de la capsidie est établie). Il en résulte une redondance élevée des données par particule imagée et une diminution des erreurs d'alignement. De plus, la grande taille des virions donne un excellent contraste qui permet une estimation plus précise de l'orientation des particules. La reconstruction 3D a été réalisée avec le logiciel Relion (REgularized Likelihood Optimisation) avec une approche statistique de vraisemblance maximale (Scheres 2012). Dans cette approche il n'y a pas d'attribution discrète de l'orientation des particules, elles seront assignées à toutes les orientations, mais pondérées par une probabilité. Relion présente l'avantage d'être simple à utiliser de par le nombre de paramètres à choisir qui est limité. Cependant, l'approche statistique qu'il utilise demande une grande puissance de calcul, et nécessite de travailler sur des infrastructures informatiques dédiées au traitement d'image.

#### **2.3.4.1. La classification 2D**

Une classe est un ensemble de particules, la classification 2D correspond à un alignement 2D des particules entre elles et permet d'estimer statistiquement la qualité des données en séparant les données en différentes classes. Au départ les particules séparées aléatoirement dans différentes classes, puis les particules dans chaque classe sont moyennées entre elles pour créer une classe référence. Chaque particule est alors soumise à une série d'opérations de translation et rotation dans le plan dans le but d'obtenir la meilleure corrélation possible avec une des classes références. Les particules sont enfin réattribuées aux classes et de nouvelles classes références sont générées. La classification 2D se déroule ainsi de manière itérative jusqu'à l'obtention de classes satisfaisantes. Si aucun détail n'est visible dans les classes 2D, il est peu probable d'observer des détails dans la reconstruction 3D. Relion ne demande que le nombre de classes voulu comme paramètre d'entrée

pendant la classification 2D. Il utilise ensuite la méthode de vraisemblance maximale pour créer les classes, dans cette approche une même particule appartiendra à toutes les classes en même temps, mais se verra attribué un poids correspondant à la probabilité d'appartenant à cette classe. L'avantage de la méthode est de ne pas faire de choix strict de classe lorsque l'on travaille avec images très bruitées, ce qui empêche de rester bloqué dans un minimum local. La classification 2D permet de sélectionner seulement les particules qui nous intéressent pour la suite de la reconstruction.

### **2.3.4.2. La classification 3D**

Relion propose également une classification 3D où différentes cartes de densité sont reconstruites à partir du jeu de données. Au contraire de la classification 2D, une référence 3D est demandée, celle-ci a été préalablement générée avec le logiciel Auto3dem (Yan et al. 2007). Avec la même approche que la classification 2D, les particules sont réparties aléatoirement dans des classes 3D références, les premières classes 3D référence correspondent à la référence 3D qui a été filtrée avec un filtre passe-bas. Ce filtre permet d'enlever les hautes fréquences spatiales, pour éviter de créer un biais dans la reconstruction. La méthode de vraisemblance maximale va une nouvelle fois attribuer chaque particule à toutes les orientations possibles dans chacune des classes 3D, avec un poids dépendant de la probabilité d'appartenance à cette orientation et à cette classe. La classification 3D se fait également de manière itérative, où chaque cycle servira à générer de nouvelles références 3D utilisées pour le cycle suivant. En partant du principe que tous les jeux de données sont hétérogènes, la classification 3D permet de choisir à quel point on souhaite tolérer cette hétérogénéité. Elle permet de déceler les différences conformationnelles de l'échantillon et de sélectionner seulement un sous-jeu de particules pour la suite du traitement d'image. Le choix des classes 3D dépend aussi bien des différences entre les classes que du nombre de particules par classes, et il est important de conserver un équilibre entre l'homogénéité des données et le nombre de particules nécessaires pour obtenir une reconstruction 3D à haute résolution.

### **2.3.4.3. L'affinement 3D automatique**

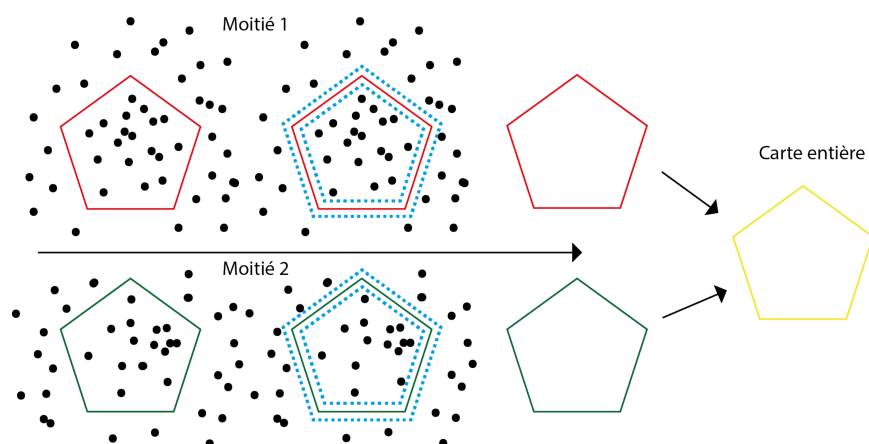
Après classification, la reconstruction 3D est affinée à partir des particules sélectionnées. L'orientation des particules est affinée contre une même référence que celle utilisée pour la classification 3D (sur laquelle un filtre passe-bas a été appliqué). L'affinement se fait selon la « Gold

standard FSC » (Scheres & Chen 2012), les données sont séparées en deux lots indépendants qui ne se croisent jamais, ainsi deux reconstructions indépendantes sont réalisées en parallèle, qui seront appelées demi-cartes. À chaque itération, un nouveau filtre passe-bas, estimé automatiquement par le programme, est appliqué aux demi-cartes pour éviter l'accumulation de bruit dans les reconstructions. L'affinement se déroule de manière automatique jusqu'à ce que la résolution ne s'améliore plus, les demi-cartes sont alors fusionnées pour obtenir la reconstruction 3D finale.

#### 2.3.4.4. Le post-traitement d'image

Le post-traitement d'image est la dernière étape de la reconstruction de la carte de densités ayant principalement pour but de valider la résolution finale estimée pendant l'affinement 3D et d'appliquer un filtre sur la reconstruction visant à améliorer son rendu. Relion permet de réaliser les trois opérations suivantes sur la carte de densité :

- Un masque gaussien couvrant la région d'espace occupée par les molécules est appliqué à chacune des demi-cartes avant qu'elles ne soient fusionnées. L'application de masques permet d'éliminer du calcul de la FSC, le bruit restant autour des molécules (**Figure 44**) ;
- Les phases de la carte sont remplacées par des valeurs aléatoires au-delà d'une résolution choisie automatiquement par Relion pour simuler du bruit. Une comparaison est faite avec les phases originales pour vérifier la présence de bruit au-delà de la résolution choisie (voir 1.3.4.1) ;
- Un aiguisement gaussien est appliqué pour augmenter le signal des hautes fréquences spatiales, et faire ressortir les détails du virus correspondant au signal à haute résolution. La valeur de l'aiguisement gaussien est déterminée et appliquée automatiquement selon la procédure décrite dans (Rosenthal & Henderson 2003).



**Figure 44 : Représentation schématique de l'application du masque gaussien.** Un masque est appliqué (pointillés bleus) pour faire abstraction du bruit (points noirs) autour molécules avant la fusion des deux demi-cartes.

### 2.3.5. Le traitement des vues

L'incorporation des vues permet d'accéder à la haute résolution. Dans le cadre de l'étude de la protéine de capsid de l'ArMV, deux approches ont été abordées pour utiliser les vues de la meilleure manière possible :

- Un traitement des vues en amont de la reconstruction 3D a été fait avec les logiciels Unblur et Summovie qui ont permis d'aligner l'ensemble des vues avant de les sommer.
- Un traitement des vues après la reconstruction 3D a été fait avec Relion. À la différence de Unblur, il y a un alignement local des particules présentes dans chaque vue.

## 2.4. La construction et le premier affinement du modèle atomique

Un modèle atomique 3D a été ensuite construit en plaçant les atomes dans l'espace en accord avec la carte de densité issue de la reconstruction 3D. Les positions des atomes sont alors affinées pour que la carte de densité théorique du modèle soit le plus proche possible de la carte de densité expérimentale. La fonction d'optimisation de l'affinement du modèle se base sur une approche statistique de vraisemblance maximale (Murshudov et al. 2011). Les reconstructions 3D issues de la cryo-ME n'étant pas encore à de très hautes résolutions, il existe une ambiguïté dans le recalage des

atomes à cause du manque de détails. Pendant l'affinement, les atomes peuvent être piégés dans un minimum local qui n'est pas forcément la meilleure conformation possible. Pour cette raison, il est nécessaire d'introduire dans la fonction d'optimisation des restrictions géométriques supplémentaires aux données expérimentales. Cela permet d'ajouter des informations supplémentaires et stabiliser l'affinement du modèle atomique lorsque les données expérimentales ont une résolution inférieure à  $\sim 2,8$  Å. Pour l'affinement du modèle de l'ArMV dans la reconstruction 3D, les logiciels suivants ont été utilisés :

- La suite de logiciel CCP4 (Winn et al. 2011) : Coot pour l'affinement manuel (Emsley & Cowtan 2004), Refmac5 pour l'affinement automatique (Brown et al. 2015), ProSMART pour l'ajout de restrictions externe (Nicholls et al. 2012) ;
- Le logiciel Mapman pour la manipulation des cartes de densité (Kleywegt & Jones 1996) ;
- Le logiciel Super pour la superposition de modèles et de cartes de densité (Trapani & Navaza 2007) ;
- Le logiciel Molprobity pour vérifier la normalité de la géométrie moléculaire (Chen et al. 2010).

## 3. Résultats et discussion

### 3.1. Reconstruction 3D de la carte de densité

Sur un total de 2680 films enregistrés sur le Titan Krios du Ncec, 1225 ont été sélectionnés avec e2boxer.py. À partir de ces 1225 films, 16245 coordonnées de particules ont été écrites avec une taille de boîte de 360 x 360 pixels. Les étapes de la reconstruction présentées ci-après ont été réalisées avec l'utilisation des 7 vues alignés par le logiciel Unblur.

#### 3.1.1. Classification 2D

Une classification 2D a été faite avec les 16245 particules pendant 25 cycles itératifs avec 100 classes demandées, et un pas d'échantillonnage de  $5^\circ$  pour la rotation dans le plan, suivi de 5 cycles supplémentaires avec un pas d'échantillonnage de  $2^\circ$ . L'évolution des classes 2D est illustrée avec 4 classes en **Figure 45**, 15889 particules ont été sélectionnées après cette classification 2D (**Figure 46**).

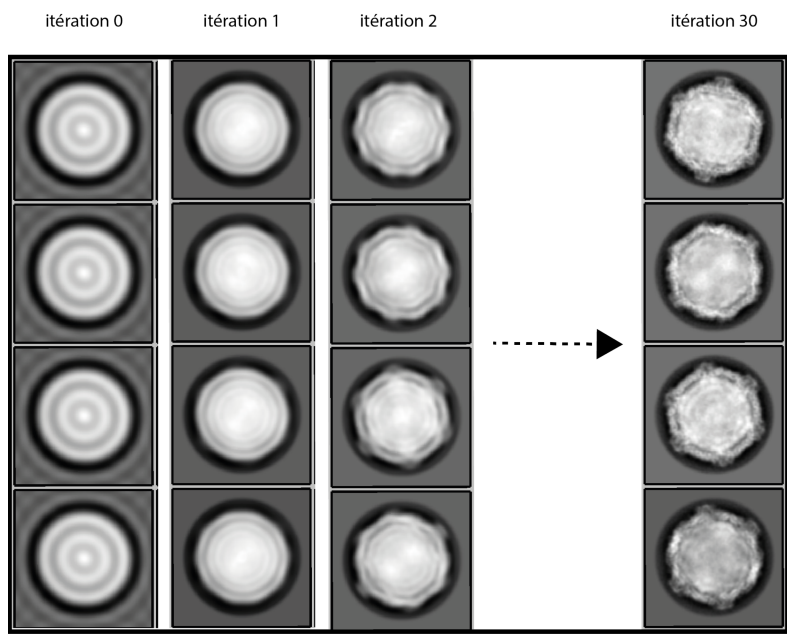


Figure 45 : Évolution des classes 2D du virus de l'ArMV au fil des 30 itérations.



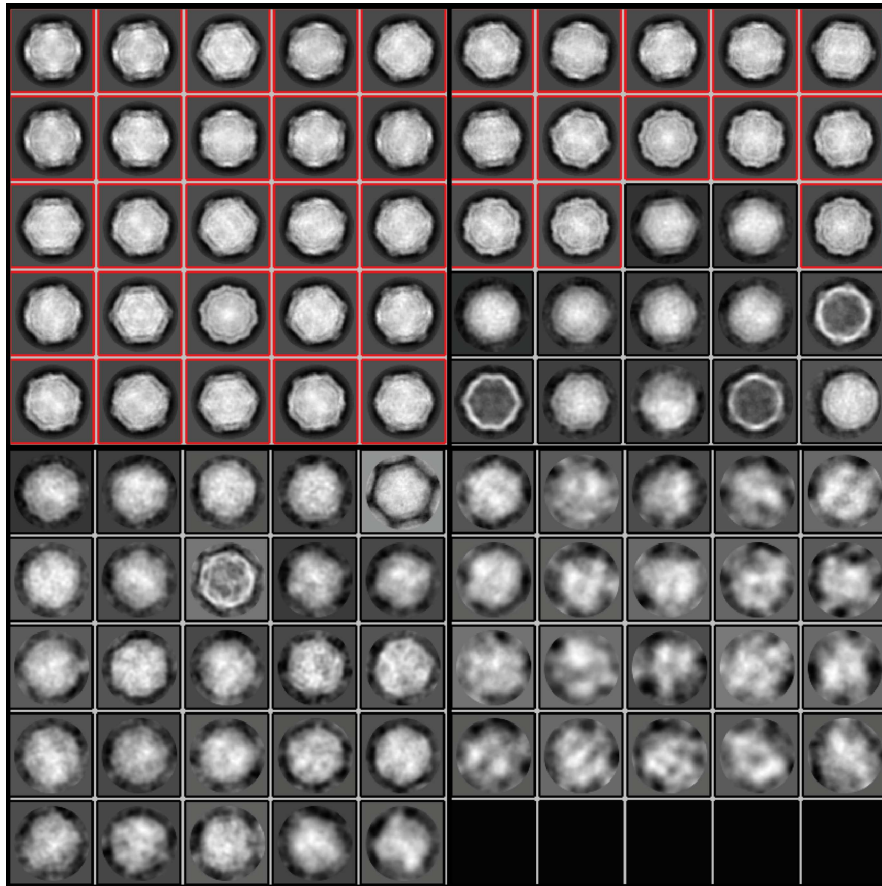
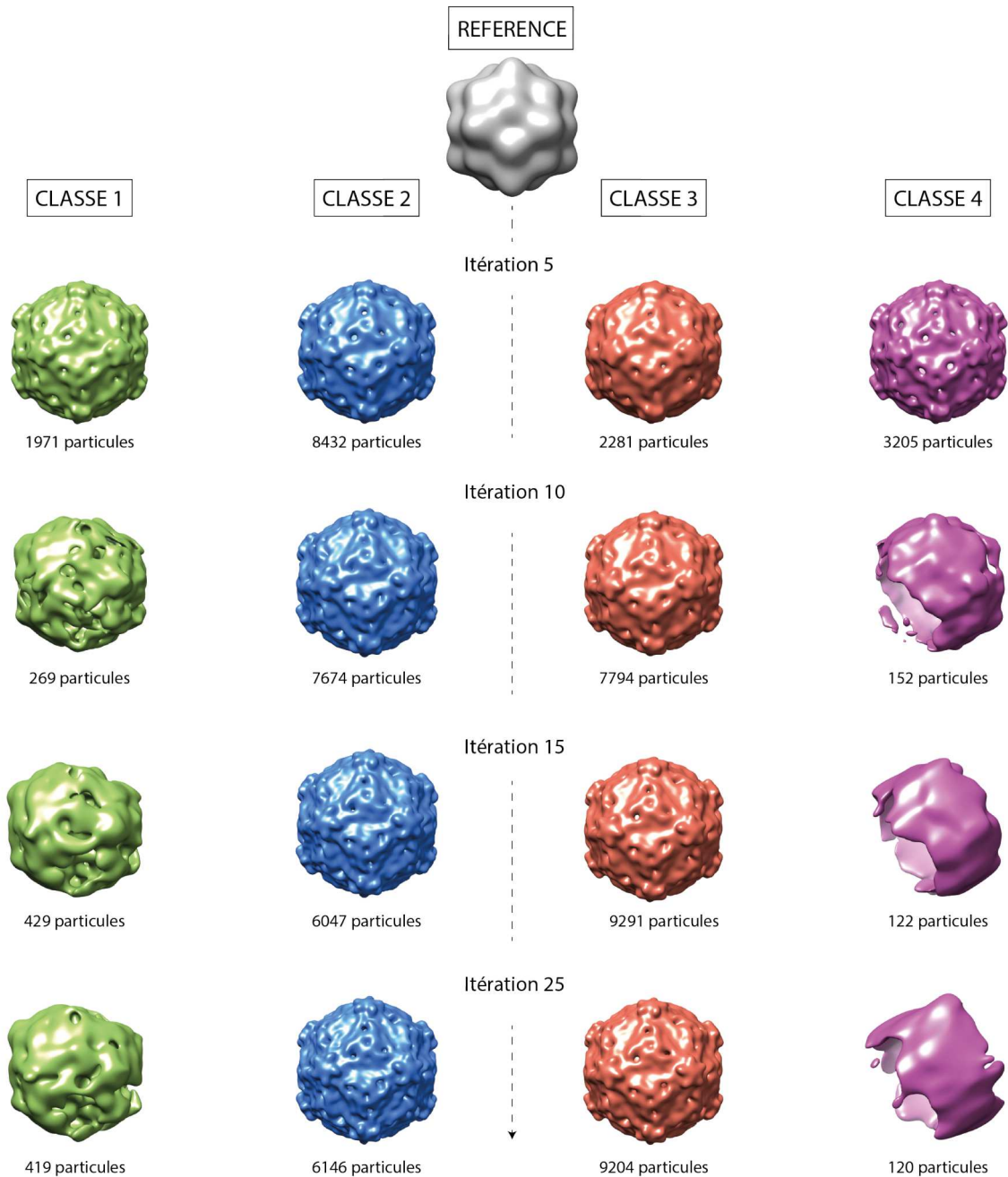


Figure 46 : Visuel des classes 2D après 30 itérations. Les classes sélectionnées pour la suite de la reconstruction sont en rouge).

### 3.1.2. Classification 3D

Les 15889 particules issues de la classification 2D ont été soumises à une classification 3D pendant 25 cycles itératifs avec 4 classes demandées. Un filtre passe-bas à 60 Å a été appliqué sur la référence et aucune symétrie n'a été imposée (Figure 47).



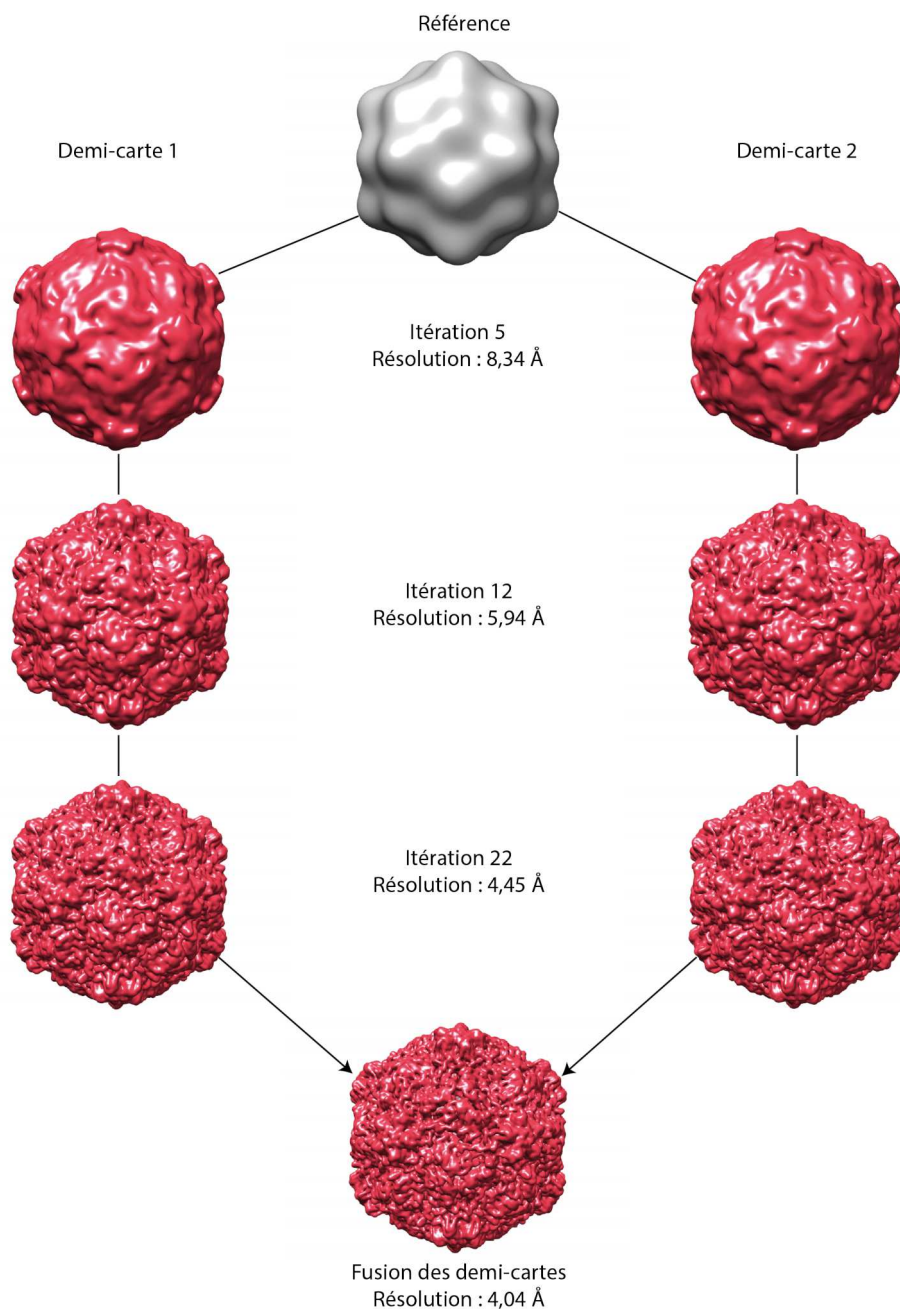


**Figure 47 : Classification 3D du virus de l'ArMV à partir des 15889 particules issues de la classification 2D.**

La classification 3D a permis d'obtenir 2 classes avec des détails structuraux visibles et 2 classes incohérentes. Les particules appartenant aux classes 3D 1 (419 particules) et 4 (120 particules) ont été écartées pour la suite de la reconstruction, celles appartenant aux classes 3D 2 (6146) et 3 (9204) ont été soumises à un affinement automatique.

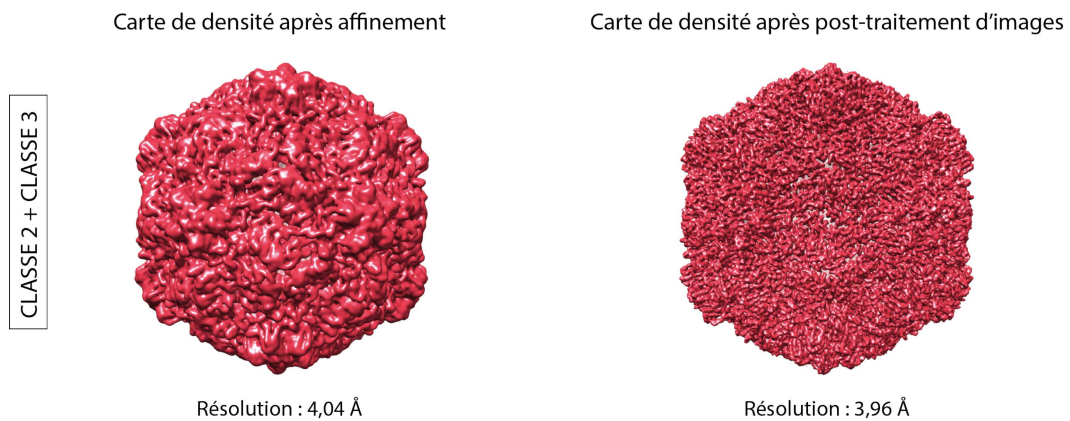
### 3.1.3. Affinement automatique et post-traitement d'image

La somme des particules appartenant aux classes 3D 2 et 3 a été utilisée pour un premier affinement automatique comprenant un total de 15350 particules. La référence utilisée a été la même que la classification 3D avec application d'un filtre passe-bas à 60 Å également, une symétrie icosaédrique I2 a également été imposée pendant l'affinement, qui a convergé au bout de 22 cycles et a généré une carte de densité à 4,04 Å (**Figure 48**).



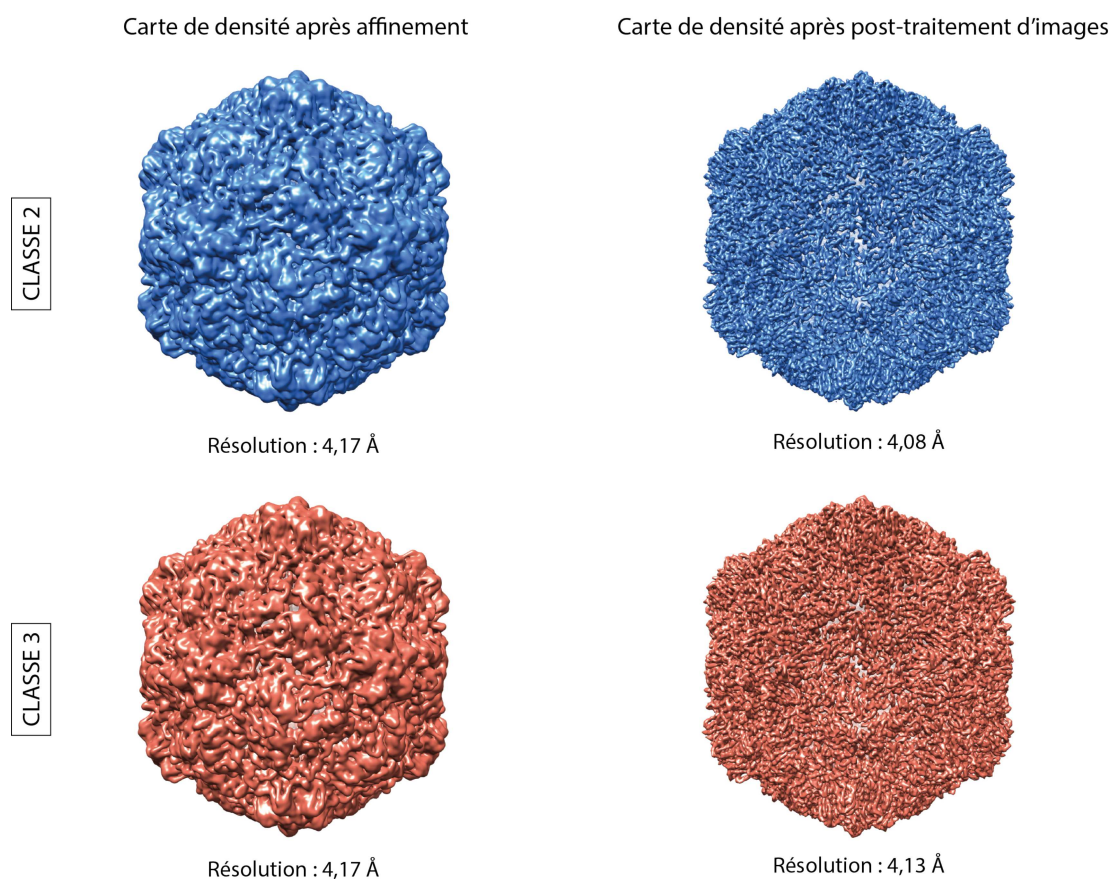
**Figure 48** : Affinement automatique des 15350 particules issues de classes 3D 2 et 3.

La carte de densité a ensuite été soumise au post-traitement d'image de Relion, un masque gaussien a été appliqué avec un seuil de binarisation de 0,02, et un aiguisement gaussien d'une valeur de -170 Å<sup>2</sup> a été appliqué automatiquement. La résolution de la carte s'est légèrement améliorée pour une valeur estimée à 3,96 Å (critère de résolution à 0,143 FSC) (**Figure 49**).



**Figure 49 : Carte de densité des particules contenues dans les classes 2 et 3, avant et après post-traitement d'image.**

De la même manière, les particules contenues dans les classes 3D 2 et 3 ont été affinées séparément pour vérifier si des lots de particules plus homogènes permettaient d'améliorer la résolution de la carte de densité. Cependant, les résolutions obtenues étaient légèrement inférieures à la première reconstruction (**Figure 50**).



**Figure 50 : Carte de densité des classes 2 et 3 avant et après post-traitement d'image.**

Plusieurs autres classifications 2D et 3D ont été réalisées pour essayer d'obtenir des lots de particules plus homogènes. Néanmoins, la carte de densité avec la plus haute résolution a été obtenue avec les 15350 particules contenues dans les classes 2 et 3.

### 3.1.4. L'impact des vues sur la reconstruction 3D

Un travail de recherche sur l'impact des vues a été fait à partir des 15350 particules retenues. La première reconstruction ayant été faite avec les 7 vues alignées par Unblur, le premier test a été d'aligner les vues de la même manière, mais en ajoutant le filtre de dose proposé par Unblur. Cette approche n'a pas permis d'améliorer la résolution de la carte (**Figure 51**).

Alignement global des photogrammes avec Unblur	
Nombre de photogrammes alignés et sommés	Résolution avant post-traitement d'images (Å)
7 photogrammes	4,04
7 photogrammes + filtre de dose	5,45

**Figure 51 : Impact du filtre de dose proposée par Unblur sur la résolution.**

L'impact du nombre de vues utilisées (et donc de la totale d'électron utilisée) avec un alignement par Unblur a été étudié. Il semble qu'avec ce jeu de donnée un nombre de vues compris entre 2 et 4 permet d'améliorer la résolution finale. La variation de résolution comprise dans cette plage reste néanmoins négligeable (Figure 52).

Alignement global des photogrammes avec Unblur	
Nombre de photogrammes alignés et sommés	Résolution après -traitement d'images (Å)
1 photogramme	4,45
2 photogrammes	3,80
3 photogrammes	3,80
4 photogrammes	3,80
5 photogrammes	3,84

**Figure 52 : Impact du nombre de vues alignés par Unblur sur la résolution.**

Le traitement des vues par Relion a ensuite été réalisé avec un nombre de 7 et 3 vues pour une comparaison avec les résultats obtenus avec Unblur (Figure 53).

Alignement local des photogrammes avec Relion	
Nombre de photogrammes alignés et sommés	Résolution après -traitement d'images (Å)
Image entière	4,45
7 fractions d'images + polissage	3,96
3 fractions d'images + polissage	3,84

**Figure 53 : Impact du traitement des vues par Relion sur la résolution.**

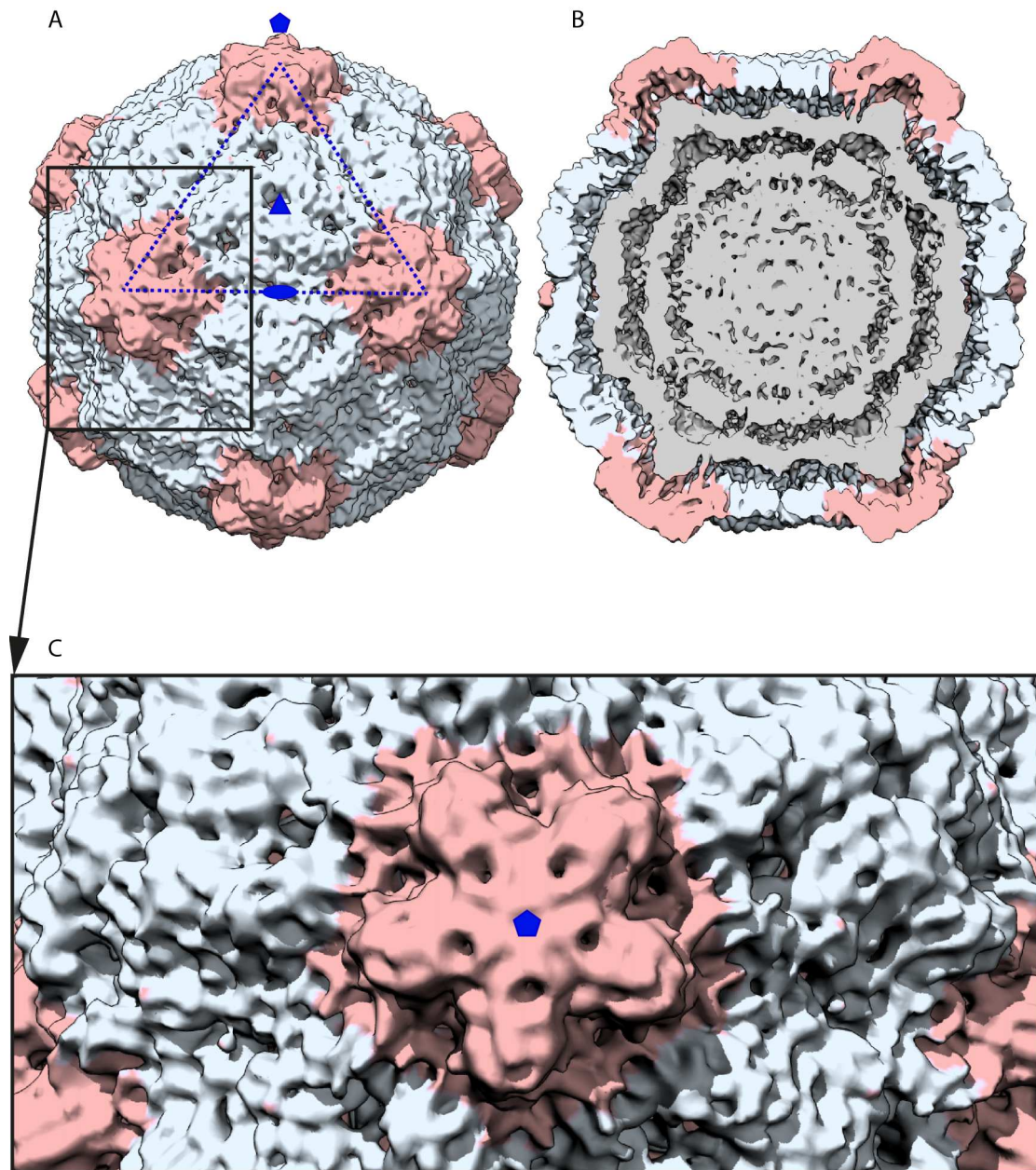
Ainsi, les deux approches montrent des résultats très comparables, une diminution de la dose d'électron totale utilisée permet d'améliorer la résolution de la reconstruction. Des tests supplémentaires ont été réalisés, comme l'alignement individuel des particules par Unblur, ou une

nouvelle extraction des particules dont le centre de la boîte a été recentré à partir de la reconstruction 3D. Cependant, ces approches n'ont pas amélioré la résolution par rapport aux résultats obtenus précédemment.

### 3.1.5. Structure de la capsid de l'ArMV

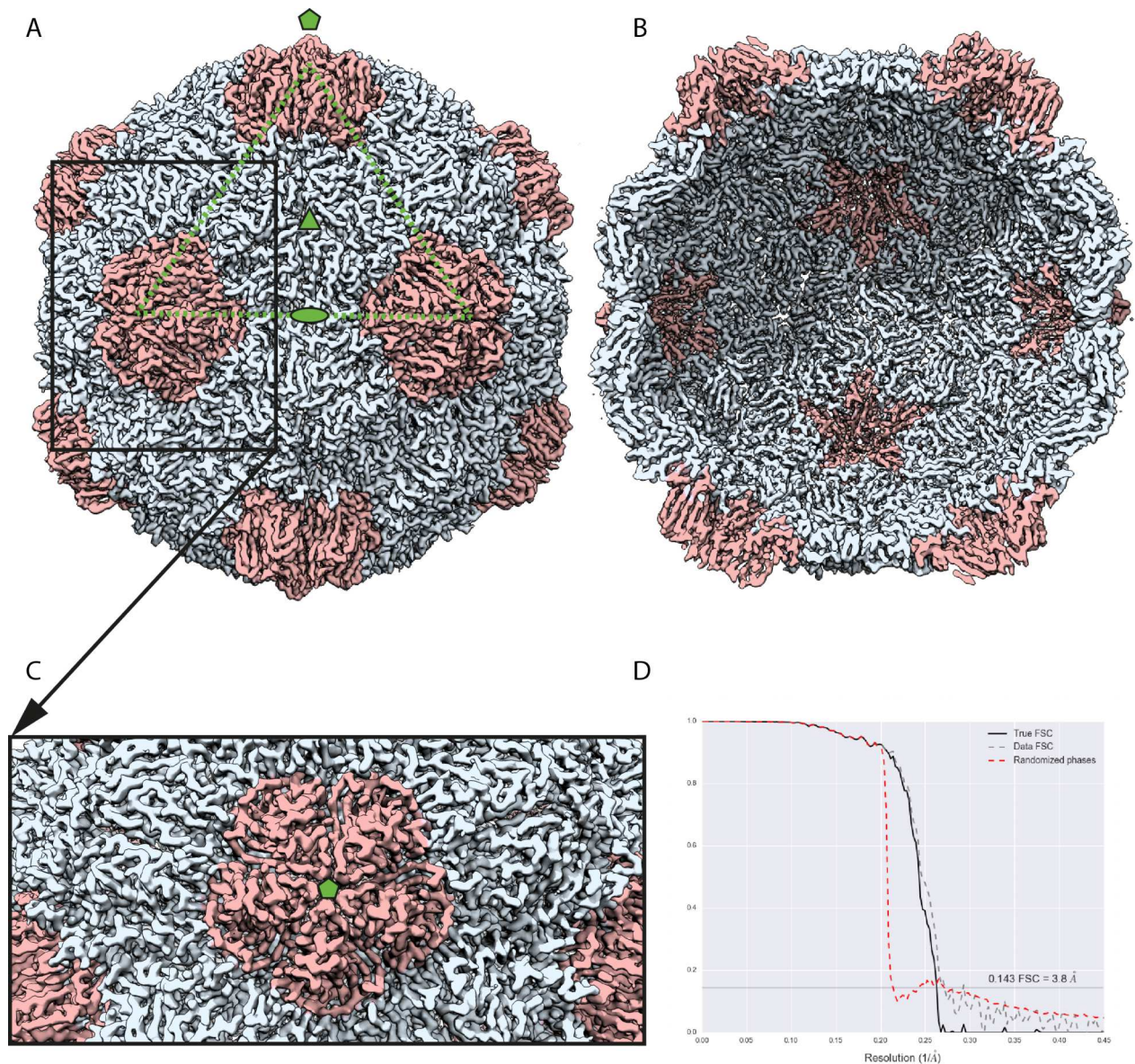
La carte de densité retenue pour l'affinement du modèle atomique a été celle obtenue par un alignement global de 3 vues par Unblur, avec une résolution estimée à 3,80 Å. La différence de résolution entre l'ancienne carte résolue 6,5 Å et la nouvelle peut être observée en **Figure 54** et **Figure 55** respectivement. Le logiciel Resmap a été utilisé pour visualiser la variation de résolution dans chaque région de l'icosaèdre. La résolution semble globalement homogène, excepté au centre de l'axe de symétrie d'ordre 5 qui montre une plus basse résolution que le reste de l'icosaèdre (**Figure 56**).





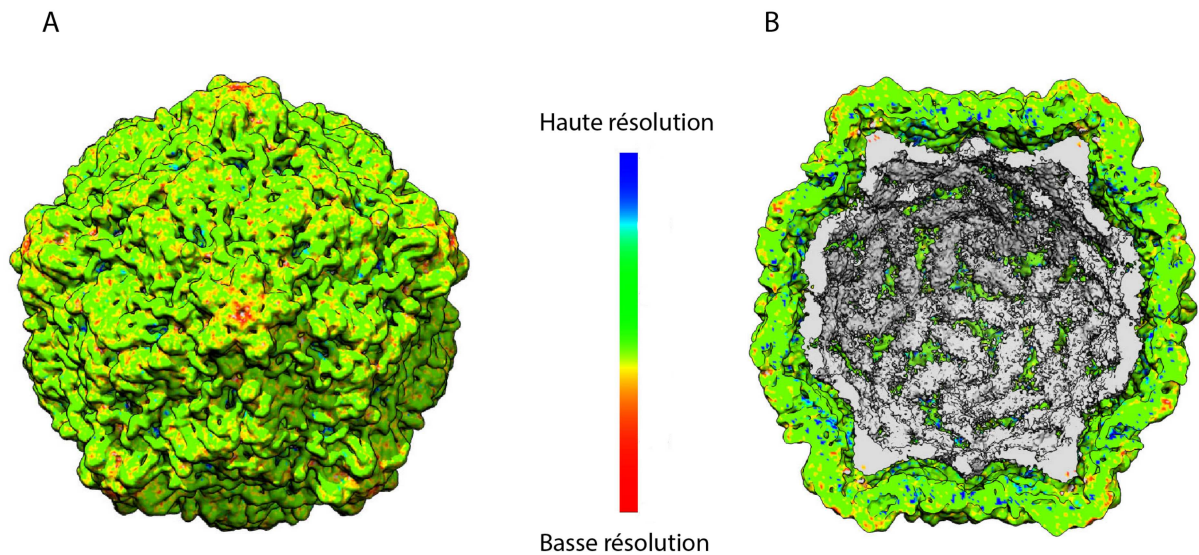
**Figure 54 : Reconstruction du virus de l'ArMV à 6,5 Å. (A) Vue externe du virus avec les axes de symétrie d'ordre 5, 3 et 2, représentés respectivement par un pentagone, un triangle, et une ellipse. (B) Vue interne du virus avec le génome coloré en gris. (C) Zoom sur un pentamère du virus.**





**Figure 55 : Reconstruction du virus de l'ArMV à 3,8 Å. (A)** Vue externe du virus avec les axes de symétrie d'ordre 5, 3 et 2, représentés respectivement par un pentagone, un triangle, et une ellipse. **(B)** Vue interne du virus, le génome n'est pas visible car le post-traitement agit comme un filtre passe-haut, enlevant le signal basse résolution. **(C)** Zoom sur un pentamère du virus. **(D)** *Fourier Shell Correlation* de la reconstruction 3D. La courbe noire correspond à la valeur correcte de FSC, estimée depuis la valeur de FSC de la reconstruction finale (pointillés gris) dont la contribution du bruit a été enlevée après randomisation des phases à haute résolution (pointillés rouges).





**Figure 56 :** Figure : Estimation de la résolution locale par Resmap. Cette carte correspond à une carte de différence entre les deux demi-cartes. (A) Vue externe du virus. (B) Vue interne, le génome est coloré en gris.

### 3.2. Affinement du modèle atomique

L'affinement du modèle atomique s'est fait en utilisant le modèle homologue du GFLV résolu par cristallographie comme référence (pdb : 2y26, résolution : 2,7 Å). Ce modèle a d'abord été recalé dans la carte de densité, et remplacé par la séquence de l'ArMV. Ensuite, des cycles d'affinement automatique et manuel ont été réalisés, avec vérification de la géométrie après chaque cycle, jusqu'à convergence. L'affinement du modèle est résumé dans la **Figure 57** et décrit en détail par la suite.

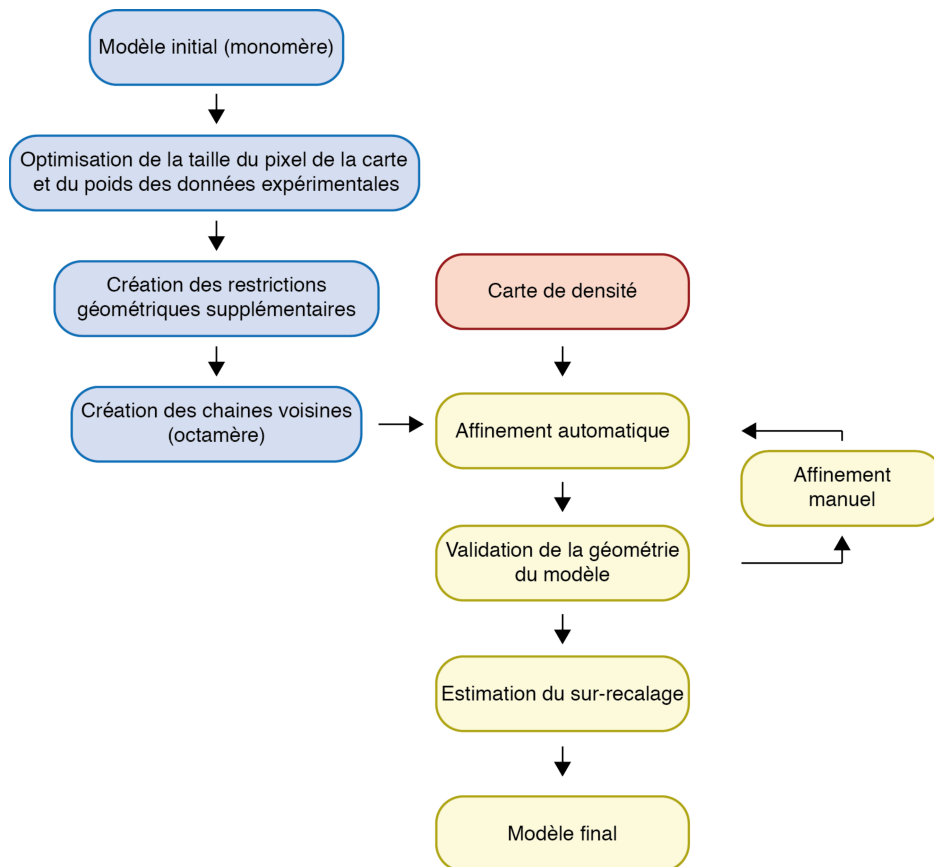


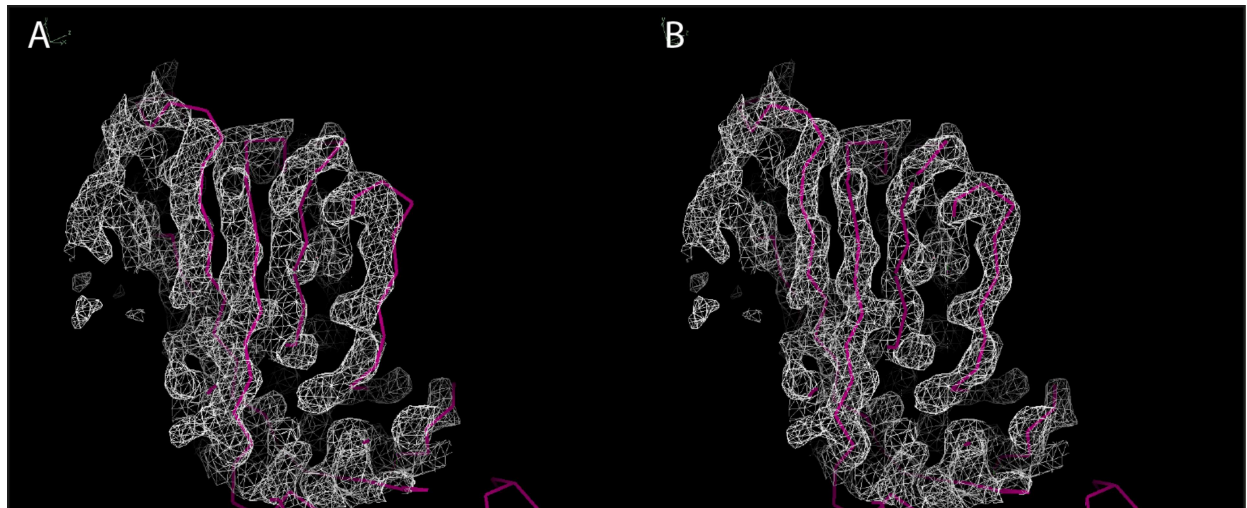
Figure 57 : Diagramme récapitulatif du protocole utilisé pour l'affinement du modèle atomique.

### 3.2.1. Positionnement de la carte de densités et du modèle

Les centres de masse de la carte de l'ArMV et de l'assemblage biologique du modèle de la capsid du GFLV (icosaèdre entier) ont été transférés à l'origine du système cartésien ( $x=0$ ,  $y=0$ ,  $z=0$ ) pour faciliter la manipulation du modèle dans la carte. En utilisant les axes de symétrie de la carte de densité de l'ArMV comme repère, le modèle du GFLV a ensuite été réorienté de manière à retrouver la même orientation que la carte. Le logiciel Mapman a permis de calculer et déplacer le centre de masse, le logiciel Super a permis de superposer le modèle du GFLV sur lui-même afin de retrouver les opérations de rotation pour réorienter le modèle.

### 3.2.2. Ajustement du monomère du GFLV dans la carte de densité

Un seul monomère du modèle du GFLV est conservé, il a alors subi un affinement de corps rigides autour de sa position initiale pour ajuster le monomère dans la carte de densité. L'affinement de corps rigide a été réalisé avec la commande « Jiggle Fit » du logiciel Coot (**Figure 58**).



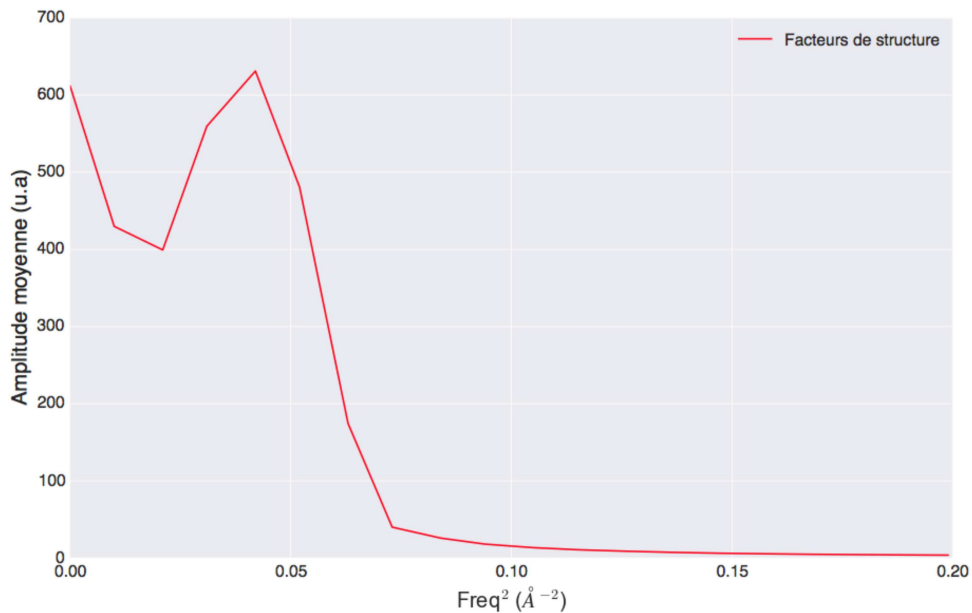
**Figure 58 : Affinement de corps rigides par « Jiggle Fit ».** Les bâtons roses correspondent au tracé de carbone alpha du monomère du GFLV. Les densités blanches correspondent à la carte de densité ArMV.

### 3.2.3. Construction du modèle

La séquence d'acides aminés du monomère du GFLV a été mutée manuellement avec le logiciel Coot pour correspondre à la séquence de l'ArMV, en se basant sur l'alignement de leurs séquences. L'alignement a montré la présence d'une délétion en position 306 chez le GFLV, entraînant un changement de numérotation dans la séquence de l'ArMV. Après observation du modèle dans la carte, cette délétion s'est avérée être finalement en position 304. Nous avons également remarqué la présence de trois domaines structuraux dans le monomère, et avons décidé de modéliser initialement le désordre de la molécule en supposant qu'une grande partie du désordre des atomes puisse être vu comme un désordre des trois domaines en tant que corps rigides (**Figure 59**).



doucement jusqu'à 0, aucun filtre passe-bas n'a été appliqué (**Figure 60**). Cette étape a été réalisée avec Refmac.



**Figure 60** : Graphique du nombre de facteurs de structure par rapport à la résolution.

### 3.2.4.2. Ajustement de la taille de pixels de la carte

En microscopie électronique à transmission, la taille du pixel des films dépend de la taille physique du pixel du détecteur, du grandissement utilisé pendant l'acquisition et des imperfections de la calibration du microscope. Il y a normalement une imprécision de plus ou moins 5 % de la taille du pixel donnée pour l'image. Nous avons donc regardé la variation de l'accord entre la carte et le modèle lorsque la taille du pixel était changée. Plusieurs affinements automatiques (dont les détails seront décrits par la suite) du monomère de l'ArMV dans la carte ont été réalisés en parallèle, et la taille du pixel donnant le meilleur accord a été choisie. La taille initiale du pixel (1,09 Å) a été celle donnant la meilleure corrélation entre le modèle et la carte (**Figure 61**). Cette étape a été réalisée avec Refmac pour l'affinement du modèle, et avec le logiciel Imod (Kremer et al. 1996) pour la modification de la taille du pixel.

Taille du pixel (Å)	Corrélation (FSC moyenne)
1,068	0,8032
1,079	0,8223
1,090	0,8367
1,100	0,8322
1,111	0,8268

Figure 61 : Variation de la taille du pixel et l'accord entre le modèle et la carte.

### 3.2.4.3. Ajustement des paramètres liés à l'affinement du modèle atomique

#### 3.2.4.3.1. Création des restrictions géométriques supplémentaires

Le logiciel Refmac, utilisé pour l'affinement automatique du modèle, possède une librairie de restrictions géométriques permettant au modèle de conserver une géométrie normale des structures cristallographiques résolues à haute résolution :

- Les distances et angles de liaisons chimiques ;
- Les angles de torsions ;
- La chiralité des centres chiraux ;
- La planéité des groupes aromatiques et des liaisons peptidiques ;
- Les répulsions de Van der Waals ;
- Les valeurs du désordre (Facteur B) d'atomes liés.

À côté de ces restrictions générales, il est possible d'ajouter des restrictions supplémentaires s'il existe un modèle homologue résolu à haute résolution. Le modèle du GFLV a donc servi à générer ces restrictions supplémentaires afin que le modèle de l'ArMV garde un repliement similaire. Les restrictions supplémentaires ont été créées avec Prosmart.

#### 3.2.4.3.2. Optimisation du poids des données expérimentales

Pendant l'affinement du modèle atomique, un poids est attribué aux données expérimentales par rapport aux restrictions géométriques. Ce poids régule le degré de liberté autorisé pendant l'affinement du modèle afin que celui-ci conserve une géométrie normale. Plusieurs affinements automatiques du modèle atomique avec différents poids imposés ont été réalisés en parallèle. Le

choix de ce poids s'est fait en observant l'écart-type des distances atomique du modèle par rapport aux distances idéales (**Figure 62**).

Poids des données expérimentales	Ecart-types des distances de liaison idéales	Corrélation entre le modèle et la carte
0.1	0.009	0.8303
0.2	0.009	0.8320
0.3	0.010	0.8357
0.4	0.010	0.8338
0.5	0.010	0.8343
0.6	0.011	0.8331
0.7	0.011	0.8305
0.8	0.012	0.8329

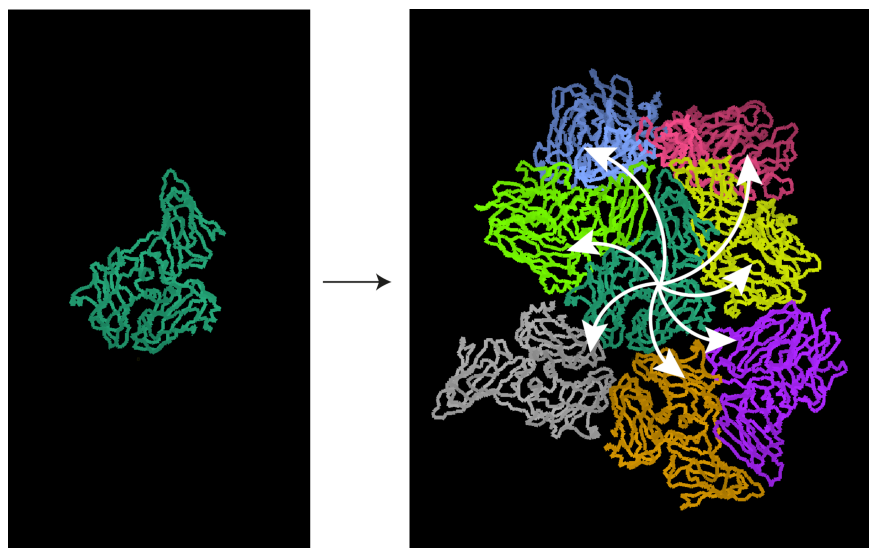
**Figure 62 : Variation du poids des données expérimentales avec la géométrie du modèle.**

Le choix du poids se fait tout d'abord par rapport à l'écart-type des distances de liaison idéales, le but étant que celui-ci ne dépasse pas 0,010. Les valeurs de poids entre 0,1 et 0,5 ayant toutes données un écart-type égal ou inférieur à 0,010. Dans cet intervalle, le choix du poids s'est alors fait par rapport à plus haute valeur de corrélation entre le modèle et la carte a été choisi : 0,3. Cette étape a été réalisée avec Refmac.

### 3.2.4.3.3. Création d'un octamère

Nous n'affinons pas le modèle de l'icosaèdre entier pour des raisons de temps de calcul. Cependant, un octamère est généré autour du monomère afin d'avoir toutes les zones de contact du monomère avec ses voisins, les chaînes voisines sont des copies conformes du monomère central. L'affinement de l'octamère permet ainsi d'éviter les débordements du modèle atomique dans les densités des chaînes voisines et de prendre en compte les chocs entre atomes de chaînes voisines (**Figure 63**). La création de l'octamère a été réalisée avec Coot.





**Figure 63 : Création de l'octamère.** Les voisins de l'unité asymétrique ont été générés par les opérations de symétrie icosaédrique.

#### 3.2.4.3.4. Création des restrictions NCS

Les chaînes de l'octamère seront affinées indépendamment les unes des autres, cependant elles doivent être identiques entre elles. C'est pourquoi des restrictions NCS (« Non-Crystallographic Symmetry », terme désignant des copies multiples d'une molécule à l'intérieur de l'unité asymétrique du cristal) sont ajoutées pendant l'affinement du modèle. Ces restrictions vont empêcher que la géométrie des huit chaînes ne diffère trop entre elles. Le poids de ces restrictions est arbitraire, le plus fort niveau de restriction a été choisi pour l'affinement de l'octamère.

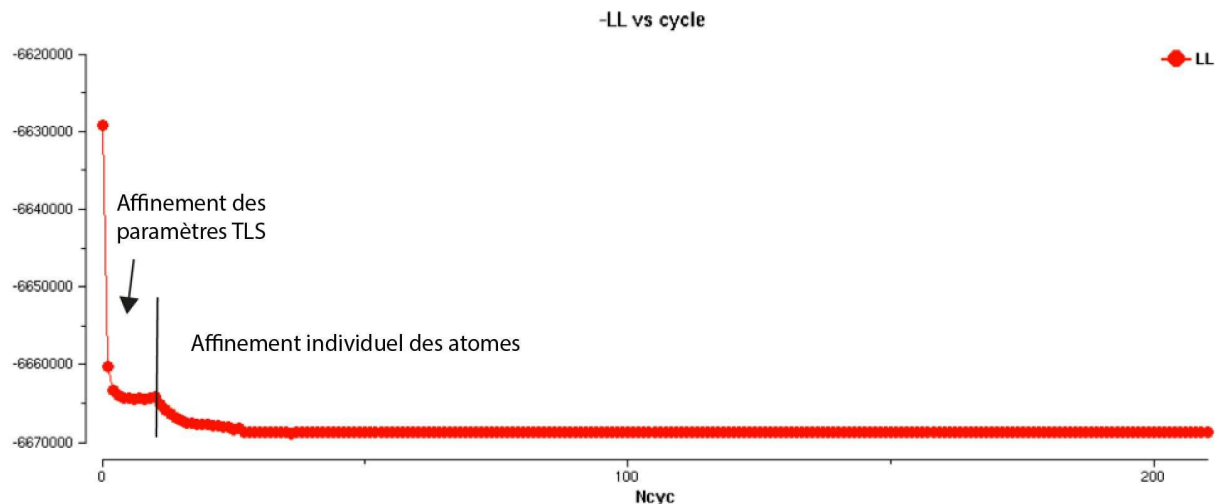
### 3.2.5. Affinement du modèle atomique

Après ajustement des différents paramètres décrits plus haut, l'affinement automatique du modèle atomique s'est déroulé de la manière suivante :

- 10 cycles d'affinement des paramètres TLS sont réalisés : les paramètres TLS (« Translation, Libration and Screw ») sont les paramètres décrivant un désordre gaussien anisotropique d'un corps rigide dans l'espace. Pour une position donnée, il y a une certaine fréquence d'appartenance à cette position qui est décrite par 20 paramètres TLS. Dans notre cas, ces corps rigides correspondent aux trois domaines décrits pendant la construction du modèle.



- Des cycles d'affinement des positions atomiques individuelles jusqu'à convergence : après affinement des paramètres TLS, chaque atome est maintenant affiné individuellement. La convergence de l'affinement est vérifiée par une fonction d'optimisation, qui peut être visualisée à travers un graphique (**Figure 64**).



**Figure 64** : Graphique de la fonction d'optimisation du modèle atomique dans la carte de l'ArMV par rapport au cycle d'affinement.

### 3.2.6. Vérification de la géométrie du modèle

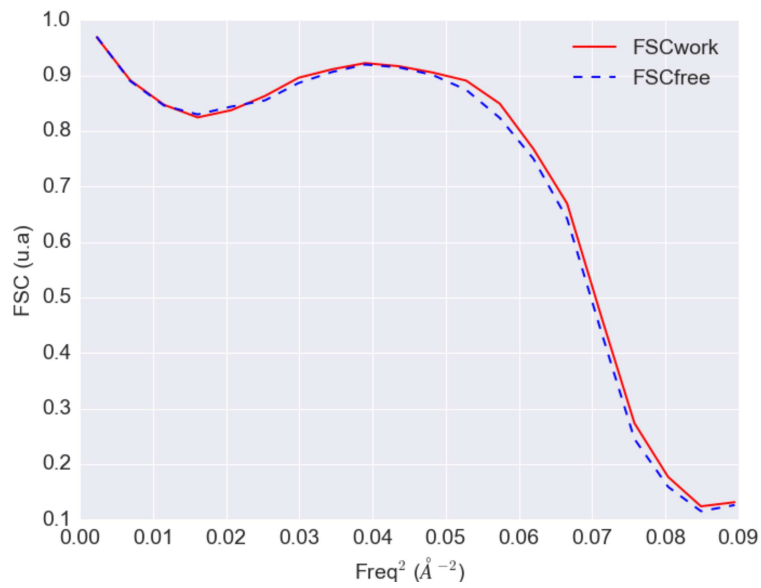
L'affinement automatique ajuste la position des atomes en essayant de maintenir un équilibre géométrie/données expérimentales. Néanmoins, les atomes peuvent se retrouver piégés dans un minimum local, qui n'est peut-être pas le minimum global. Si le déplacement nécessaire au positionnement correct des atomes est très important, la position des atomes ne pourra pas être corrigée automatiquement. C'est pourquoi le logiciel Molprobita a été utilisé pour vérifier la bonne géométrie du modèle, afin de corriger manuellement les problèmes. Après analyse de la géométrie, Molprobita propose une charte détaillant les acides aminés présentant un ou plusieurs problèmes. Cet outil permet d'estimer à quel point le modèle s'éloigne de la géométrie des protéines connues et de corriger manuellement le modèle dans Coot à l'aide de la charte, avant de recommencer l'affinement automatique par Refmac. Les statistiques de la géométrie du modèle atomique de l'ArMV sont présentés en **Figure 65**.

Validation de la géométrie par Molprobit	
<i>Score de chocs entre atomes</i>	12.6
<i>Rotamères non favorisés</i>	1.20 %
<i>Rotamères favorisés</i>	89.90 %
<i>Ramachandran aberrants</i>	0.15 %
<i>Ramachandran favorisés</i>	93.76 %
<i>Score Molprobit</i>	2.12
<i>Mauvaise distances de liaison</i>	0.00 %
<i>Mauvais angles de liaison</i>	0.00 %
<i>Géométrie aberrante du tracé de carbones alpha</i>	0.40 %

Figure 65 : Tableau statistique de la validation de la géométrie du modèle de l'ArMV par Molprobit.

### 3.2.7. Estimation du surrecalage (« overfitting »)

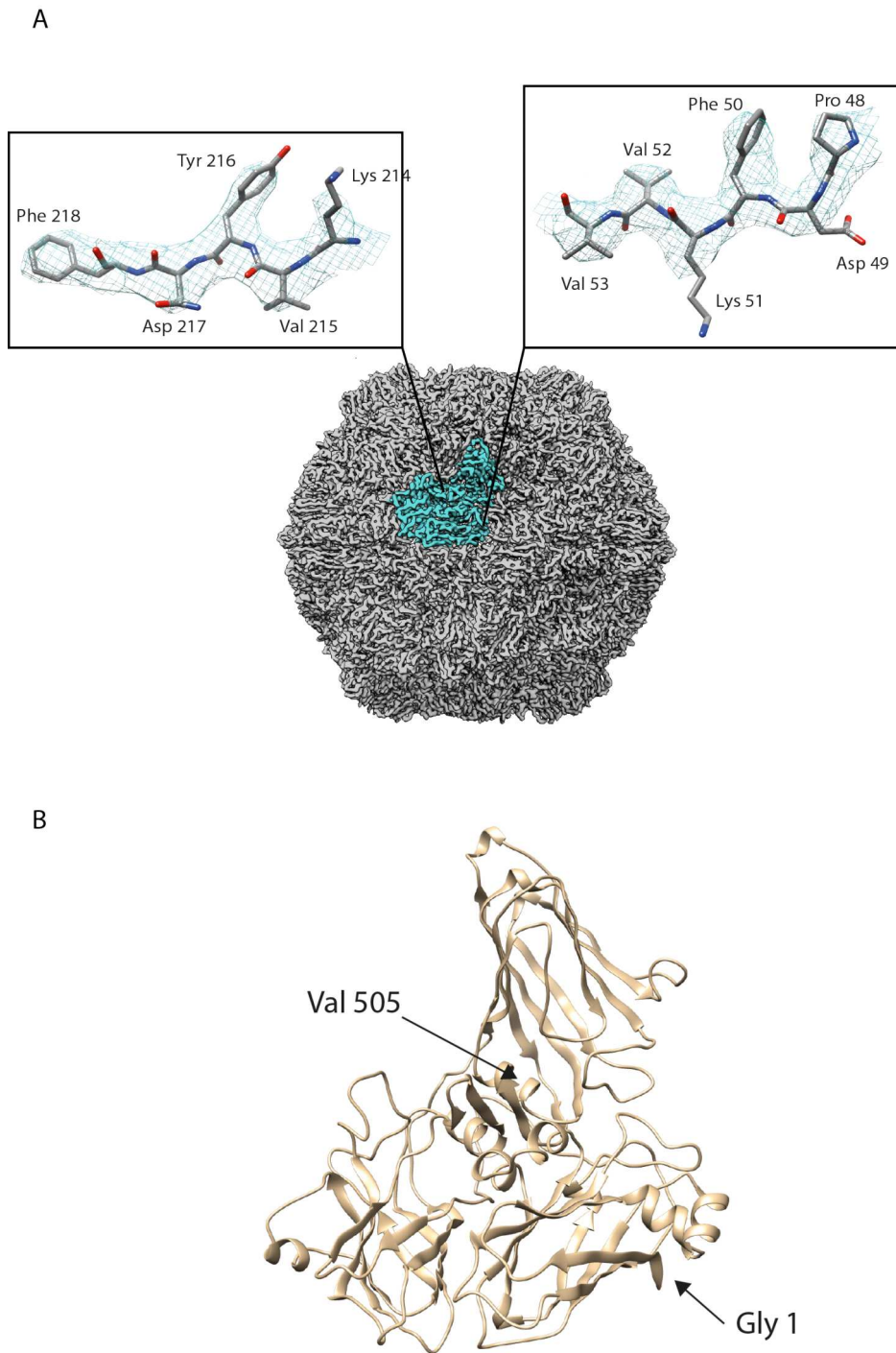
Il n'existe pas de procédure bien établie de validation de l'affinement du modèle atomique à partir des données issues de la cryo-ME, car les données sont très corrélées entre elles. Afin de vérifier si l'affinement ne s'est pas fait contre du bruit, une approche proposée est l'affinement du modèle final contre les demi-cartes ayant servi à générer la reconstruction 3D (Brown et al. 2015). Les atomes du modèle atomiques sont d'abord déplacés aléatoirement pour éviter tout biais lié au modèle, ce nouveau modèle est ensuite affiné contre une des demi-cartes et une corrélation est calculée entre le modèle nouvellement affiné et la demi-carte. Une nouvelle corrélation est ensuite calculée entre ce dernier modèle et l'autre demi-carte. Les valeurs de ces deux corrélations sont ensuite comparées graphiquement, le modèle est alors validé si sa corrélation est la même avec les deux demi-cartes (Figure 66).



**Figure 66 : Corrélation du modèle atomique avec les demi-cartes.** FSCwork correspond à la corrélation entre le modèle nouvellement affiné et la première demi-carte (contre laquelle s'est fait le nouvel affinement). FSCfree correspond à la corrélation entre le modèle nouvellement affiné et la seconde demi-carte.

### 3.2.8. Modèle atomique de l'ArMV

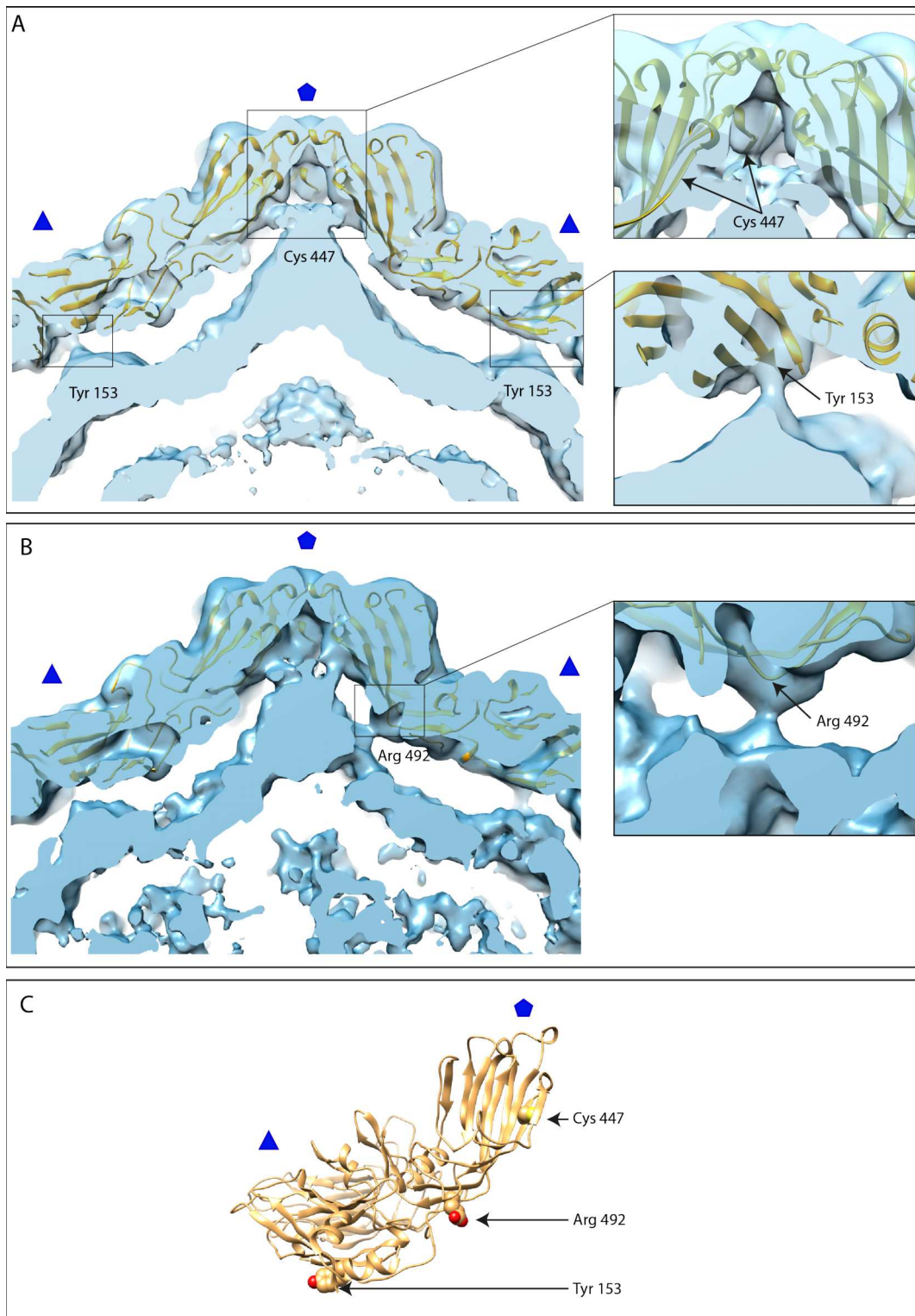
Le modèle atomique de l'ArMV a ainsi pu être construit à l'intérieur de la carte de densité obtenue par cryo-ME (**Figure 67**).



**Figure 67 : Modèle atomique de la protéine de capsid de l'ArMV.** (A) Représentation des acides aminés du modèle atomique à l'intérieur de la carte de densité d'un monomère de la protéine de capsid. (B) Représentation en ruban du modèle atomique du monomère de la protéine de capsid.

### 3.2.9. Interaction ARN/protéine de capsid

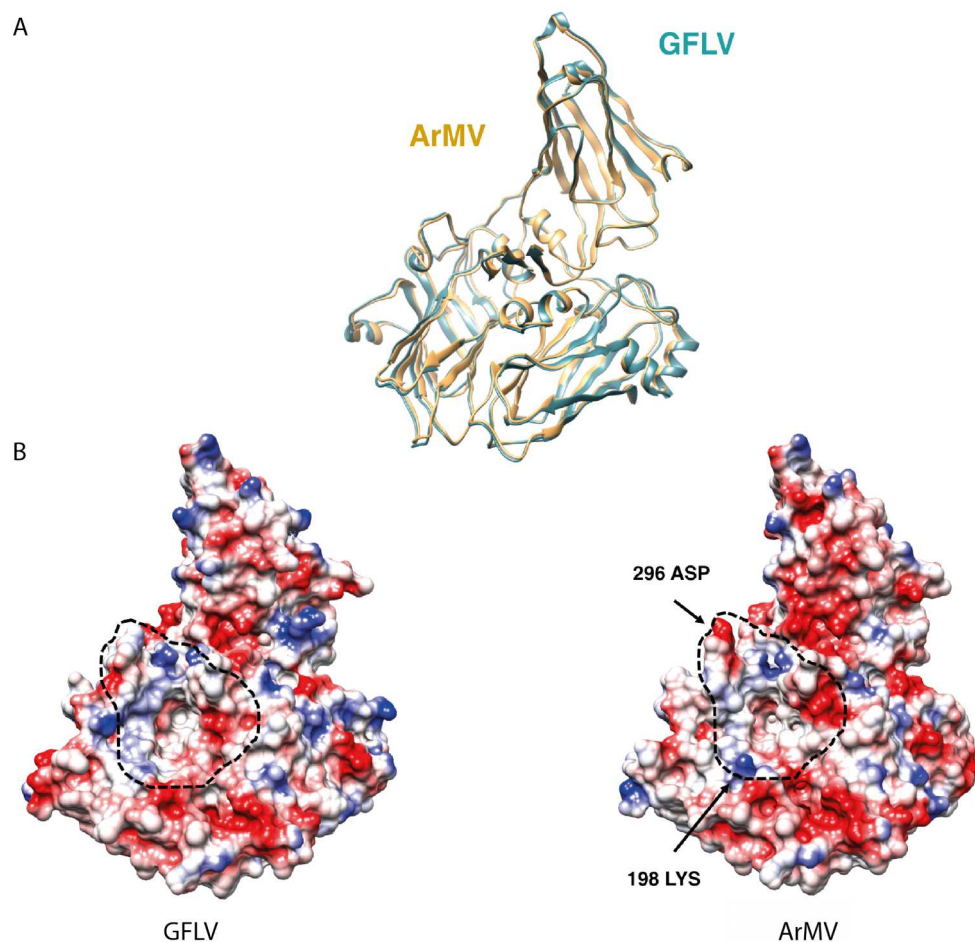
L'affinement du modèle atomique a également permis d'identifier les résidus impliqués dans l'interaction de la protéine avec l'ARN viral. Deux sites d'interactions ont pu être identifiés avec la carte de densité ayant servi à l'affinement du modèle. De plus, un site d'interaction supplémentaire a été identifié dans une carte de densité qui a été reconstruite sans symétrie icosaédrique imposée. Cette deuxième carte avait été reconstruite avec exactement les mêmes particules, que la carte de densité finale, à une résolution de 5 Å. Les cartes de densité présentées dans la **Figure 68** correspondent aux cartes avant post-traitement d'image par Relion, afin de visualiser l'ARN viral.



**Figure 68 : Interaction ARN/protéine de capsid de l'ArMV. (A)** Les deux sites d'interactions visibles dans la carte de densité reconstruite avec une symétrie icosaédrique imposée. **(B)** Le site d'interaction visible dans la carte de densité reconstruite sans symétrie imposée. **(C)** Les trois résidus qui interagissent avec l'ARN viral sont représentés en sphère sur le modèle atomique.

### 3.2.10. Première comparaison GFLV/ArMV

Une première comparaison entre les deux modèles atomiques montre un repliement semblable des deux virus avec une superposition parfaite des éléments de structures secondaires. Par ailleurs, une première observation de la répartition des charges au niveau de la poche responsable de la transmission du virus met en évidence la présence de charges additionnelles en position 296 et 198, ainsi qu'une cavité plus étroite chez l'ArMV, qui pourraient correspondre aux déterminants viraux impliqués dans la spécificité de transmission (**Figure 69**).



**Figure 69 : Première comparaison des protéines de capsid de l'ArMV et du GFLV. (A)** Superposition des modèles atomiques du monomère de la capsid du GFLV et de l'ArMV en représentation ruban. **(B)** Comparaison du potentiel électrostatique autour de la cavité de la protéine de capsid du GFLV et ArMV.



## 4. Conclusion

Au travers de l'étude structurale de l'ArMV, ce travail de méthodologie m'a permis d'acquérir les compétences nécessaires à la reconstruction 3D de structure à haute résolution et à la construction et affinement d'un modèle atomique. Le criblage des méthodes de traitement des vues a notamment montré que, non seulement l'alignement des vues joue un rôle important dans la résolution finale, mais que la dose totale d'électrons vient également influencer la résolution finale. Une carte de densité de la capsid virale a ainsi été reconstruite à 3,8 Å, ainsi que son modèle atomique, qui nous a permis d'identifier les résidus impliqués dans l'interaction protéine de capsid/génome viral. Ce travail entre dans l'étude comparative de la capsid du GFLV (précédemment obtenue par cristallographie aux rayons X) avec celle de l'ArMV, cette dernière structure étant à plus haute résolution que précédemment (de 6,5 Å à 3,8 Å), la comparaison entre les régions de la capsid impliquées dans la transmission virale devient maintenant plus pertinente. Bien qu'ils soient transmis par deux nématodes différents, les structures du GFLV et de l'ArMV restent très similaires, mais une première observation de la région impliquée dans la transmission montre la présence de charges additionnelles chez l'ArMV qui ne sont pas présentes chez le GFLV, et d'une cavité à l'intérieur de cette région qui semble différente entre les deux virus. Le but de ce projet était d'utiliser ce sujet d'étude comme outil d'apprentissage des méthodes de traitement d'image en utilisant les derniers logiciels mis à disposition par la communauté de microscopie électronique, ainsi que l'apprentissage de la construction et affinement d'un modèle atomique en utilisant les logiciels initialement dédiés à la cristallographie aux rayons X. J'ai ainsi pu transposer les compétences acquises pendant ce projet sur les autres sujets d'études présentés dans les chapitres suivants. La suite logique de ce travail sera une analyse plus fine de la stéréochimie de cette région entre les deux virus, pour proposer une hypothèse viable concernant les déterminants structuraux responsables de la spécificité de transmission des deux virus.



# CHAPITRE 4

# Chapitre 4 : Etude structurale du virus de la tâche de la fève

---

## 1. Introduction

Le virus de la tâche de la fève (*Broad bean stain virus* ou BBSV) est un virus de la sous-famille des *Comoviridae* à particules icosaédriques d'environ 30 nm de diamètre, composées de 60 copies d'une protéine de capsid, elle-même composée d'une grande et d'une petite sous-unité. Chaque particule contient un génome d'ARN à polarité positive qui peut être soit le segment 1 qui encode pour la polymérase virale, ou le segment 2 encodant pour une protéine de mouvement et la protéine de capsid. Le BBSV infecte la famille des *Fabaceae*, qui correspond aux plantes légumineuses comme les lentilles, les pois ou encore les fèves. Le virus est très présent en Asie, au Moyen-Orient, mais également en Europe. Il peut être transmis de plantes à plantes en contaminant les graines de la plante, conduisant à une infection qui est présente dès les premiers stades de développement de la plante, ou bien en utilisant les coléoptères comme vecteur de transmission, qui peut alors retenir le virus jusqu'à plusieurs semaines avant d'infecter une nouvelle plante saine. Le virus provoque un retard de croissance qui peut aller jusqu'à la mort de la plante dans les cas les plus sévères, entraînant une perte de rendement pouvant s'élever jusqu'à 77 %. Le taux maximum de transmission par les graines ne dépassant pas les 50 %, il est naturel de penser que sans transmission par les coléoptères, la transmission du virus diminuerait chaque année. C'est pourquoi il est important de mieux comprendre la transmission par les coléoptères, et particulièrement les interactions virus/vecteur. La structure 3D de la capsid du BBSV a été résolue par cryo-microscopie électronique à une résolution de 3,8 Å. Ce travail est présenté sous forme d'un article scientifique qui sera soumis prochainement pour publication, il présente les premières données moléculaires de la capsid du virus, et un premier pas vers la compréhension des mécanismes de transmission du BBSV.

## 2. Article scientifique

### **The atomic structure of Broad Bean Stain Virus revealed by cryo-electron microscopy**

Lecorre<sup>1</sup> F., Lai-Kee-Him<sup>1</sup> J., Blanc<sup>2</sup> S., Zeddami<sup>2\*</sup> J.-L., Trapani<sup>1\*</sup> S., and Bron<sup>1\*</sup> P.

<sup>1</sup> Centre de Biochimie Structurale, INSERM U1054 / CNRS UMR 5048 and Université Montpellier, 29 rue de Navacelles, 34090 Montpellier, France.

<sup>2</sup> INRA, Virus Insect Plant Laboratory, Joint Research Unit UMR 385 BGPI, Campus International de Baillarguet, Montpellier, France.

Address correspondence to Patrick Bron: Fax; +33 491 266 720; e-mail: [patrick.bron@cbs.cnrs.fr](mailto:patrick.bron@cbs.cnrs.fr); or to Stefano Trapani: e-mail: [Stefano.trapani@cbs.cnrs.fr](mailto:Stefano.trapani@cbs.cnrs.fr); or to Jean-Louis Zeddami: [jean-louis.zeddami@cirad.fr](mailto:jean-louis.zeddami@cirad.fr).

## **Abstract**

The *Broad bean stain virus* (BBSV) is a member of the genus Comovirus that infects *Fabaceae*, causing severe and widespread disease in Europe, North Africa, Asia, Middle East and Australia. The virus can be transmitted through seed or by plant weevils. BBSV has a bipartite, positive-sense and single-stranded RNA genome. The two segments are separately encapsidated in icosahedral particles that have a maximum diameter of about 30 nm consisting in 60 copies of large and small coat proteins. In the present work, we investigated the structural organization of the BBSV capsid proteins by cryo-EM and single particle analysis. We obtained a cryo-EM density map of the BBSV at 3.8 Å resolution allowing us to compute the atomic model of its capsid proteins. The N terminus parts of small subunits form a pentameric annulus structure around the fivefold axis, consequently placing the BBSV into the subgroup of CPMV-like viruses. This study allowed us to delineate connections between BBSV capsid shell and its RNA, and the structural organization of the C-terminal domain of the small coat protein, suggesting a common key role of this region in capsid assembly of comoviruses.

## ***Introduction***

Broad bean stain virus (BBSV) belongs to *Comoviridae* subfamily of the plant-infecting *Secoviridae* in the picornavirus superfamily. It induces severe and widespread disease in Europe, North Africa, Asia, Middle East and Australia. The natural host range of BBSV is restricted to *Fabaceae* such as *Lens culinaris*, *Pisum sativum*, *Vicia faba* including faba bean, broad bean and tick bean, and *Vicia palaestina* (Makkouk *et al.*, 1986). BBSV is transmitted from plant to plant according to two distinguished strategies. It can be spread through its weevil vectors (*Sitona* spp., *Apion* spp.) to other host plants that retain the ability to transmit for days or weeks, or it can be spread through sowing infected seeds leading to an initial infection of crops from the very first stage of plant development.

BBSV has a number of different strains, which vary in the severity of symptoms ranging from mild degradation to plant death (Kumari *et al.*, 1996). It has also been reported that symptoms varied cyclically during the growth of the plant (Gibbs *et al.* 1968). The symptoms are similar to those caused by a wide range of seed-borne viruses, with the consequence that BBSV is usually not detected immediately. The standard detection methods for BBSV are serological tests through ELISA and tissue blot immunoassay procedures (Kumari, Makkouk, 1993). The seed transmission rates of BBSV have been reported to be very high: up to 20% in faba beans (Mali *et al.*, 2003), 50% in field peas (Fiedorow, Szlachetka-Wawrzyniak, 2002) and 27% in lentils (Al-Khalaf *et al.*, 2002). Yield losses can also be high. For instance, pre-flowering infection in some lentils can lead up to 77% seed yield loss (Mabrouk, Mansour, 1998). The maximum rate of seed transmission that has been reported is 50%, suggesting that without BBSV field spread the transmission rate will decrease each year. Hence, there is now a necessity to focus on BBSV field spread and in a more general way on virus/insect vector interaction. As little is known about BBSV, the determinants of the BBSV transmission by these insect vectors have not been found yet. Therefore, the structural characterization of BBSV capsid at high resolution would allow us a better understanding of the general organization of the virus, its encapsidation and the interaction with its putative receptor, opening the door for further investigation.

BBSV is composed by two positive-sense, single-stranded RNA genomes, termed RNA-1 and RNA-2, which are separately encapsidated and both required for virus infection (Fig. 1). The icosahedral virion of BBSV has a diameter of ~30 nm and is composed of 60 copies each of a Large (L) and Small (S) coat protein arranged into a pseudo T3 capsid. L and S are

processed from a single precursor polyprotein, the RNA-2-encoded VP60, by the action of the RNA-1-encoded 24K proteinase. At the present time, crystallographic structures are available for three comoviruses (Lin and Johnson, 2003): cowpea mosaic virus (CPMV) (Lin et al., 1999; Ochoa et al., 2006), bean pod mottle virus (Chen et al., 1989) and red clover mottle virus (Lin et al., 2000).

The mechanism by which RNA is selected and packaged are poorly understood. The only portion of the CPMV capsid proteins currently implicated in RNA packaging is a segment of 24 amino acids at the C terminus of the S subunit (Taylor et al., 1999). This sequence is proteolytically cleaved during maturation without affecting particle stability or infectivity (Lomonossoff and Johnson, 1991), and is therefore missing from X-ray structures of CPMV (Lin et al., 1999). The suggestion that this region may promote interaction with RNA was supported by the fact that the S subunit loses its suppressor of host RNA silencing activity when the C terminus is removed (Cañizares et al., 2004). It is clear that this extension is also essential for capsid assembly itself. Indeed, deletion of the entire 24-amino acid sequence almost completely abolished empty particle formation. This effect could be partially reversed by substituting the 24 amino acids with 6 histidines, suggesting that basic amino acids in the C-terminal extension may play a role in capsid assembly (Sainsbury et al., 2011). A recent study performed by cryo-electron microscopy (cryo-EM) on the CPMV, revealed the essential role of the C-terminal region of the S coat protein subunit for virus assembly (Hesketh et al., 2015). Indeed, it was shown from the cryo-EM structure of empty virus-like particle (eVLP) of CPMV that the C-terminal region interacts with the neighbouring S subunit therefore stabilizing the formation of the penton and subsequently the formation of the particle. More recently, the crystal structure of eVLP reinforced the critical role of the C terminus of the S subunit in the efficient assembly of CPMV virions and eVLPs, but also revealed that the C termini of S subunits undergo proteolytic cleavages at multiple sites instead of a single cleavage sites as previously reported (Huynh N et al., 2016).

Here, we present the three-dimensional density map of BBSV by cryo-electron microscopy (cryo-EM) allowing us to compute the atomic model of the capsid protein. Thus, we have accessed to residues involved in RNA/capsid interactions, as well as the organization and interactions of the C-terminal domain of the small coat protein, which is significantly different from CPMV.

## ***Material & Methods***

### ***Isolation and purification of the BBSV virions***

One hundred grammes of fresh *Vicia faba* FBNSV-infected leaves (with petioles) exhibiting symptoms were collected and homogenized in 20 mM citrate buffer containing 20 mM Na<sub>2</sub>SO<sub>3</sub> (pH 6). Macerozyme at 0.1% was added and left overnight at 4°C under agitation. The extract was filtered in cheesecloth (Miracloth), 25% of cold chloroform/iso-butanol (v/v) was progressively added before being incubated under agitation at 4°C for 30 min. Then, the mixture was centrifuged (15,200 x g for 20 min at 4°C), resulting in the butanol phase containing virus particles in the upper layer and the chloroform phase in the lower layer. The virus particles were precipitated by centrifugation (302,000 x g for 3 h at 4°C). The resulting precipitate was resuspended in 20 mM citrate buffer (pH6). The solution was deposited on a discontinuous cesium chloride gradient (10-15-20-30-40 %) and centrifuged (201,000 x g for 4 h at 4°C). Virus bands were collected and centrifuged again (302,000 x g for 3h at 4°C). The pellet was finally resuspended in 20 mM phosphate buffer (pH 7.4).

### ***Cryo-EM sample preparation and imaging***

Three microliters of purified BBSV viral particles at ~2 mg/ml were applied to glow-discharged Quantifoil R 2/2 grids (Quantifoil Micro tools GmbH, Germany), blotted for 1s and then flash frozen in liquid ethane using the semi-automated plunge freezing device CP3 (Gatan inc.) at 95% relative humidity. First images were recorded on a JEOL 2200 FS electron microscope operating at 200 kV in the zero-energy-loss mode with a slit width of 20 eV, after which high resolution data was collected on the in-house spherical aberration (Cs) corrected Titan Krios S-FEG instrument (FEI, Eindhoven, Netherlands) (IGBMC-CERBM, Strasbourg, France) transmission electron microscope operating at 300 kV at a nominal magnification of X 59, 000 under low-dose conditions and with defocus ranging from -0.8 to -2.5 μm. The final pixel size was of 1.13 Å at the specimen level. A total of 2859 movies were recorded using the EPU automated acquisition software on a FEI Falcon II camera with dose fractionation. Seven individual frames were collected with an electron dose of 3.5 electrons/Å<sup>2</sup> per frame, starting from the second one. Total exposure time was 2s with a dose of 45 e<sup>-</sup>/Å<sup>-2</sup> (or 3.5 e<sup>-</sup>/Å<sup>-2</sup> per frame).



### ***Image processing and preliminary reconstructions***

The defocus and astigmatism values of each image recorded with the JEOL 2200 FS electron microscope were determined using CTFFIND3 (Mindell and Grigorieff, 2003) giving defocus values ranging from 1.2 to 3.2  $\mu\text{m}$ . Seven thousand and nine particles were semi-automatically selected and extracted using the e2boxer.py program (Tang et al., 2007). Flip of phases was then carried out on each individual particle image followed by band-pass filtering between 350 and 5  $\text{\AA}$ , normalizing and masking using the IMAGIC-5 software package (Van Heel et al., 1996; Van Heel et al., 2011). Particles were initially centered by iteratively alignments against their rotationally averaged sum and subjected to a multi-statistical analysis giving first class averages. Two class averages corresponding to RNA filled and empty capsids were selected manually and used as references for refining of particle orientations in Multi-Reference Alignment cycle. From this alignment, two set of particles corresponding to views of RNA filled and empty capsids were generated and further analyzed separately. Determination of image parameters including image shift and Euler angles assignment were performed with the AUTO3DEM software package (Yan et al., 2007). Refinement of image parameters was iteratively carried out using the PO2R program.

### ***Movie processing and final reconstruction***

Particles were selected using e2boxer.py program in EMAN2 (Tang et al. 2007), the seven frames of each movie were drift-corrected and averaged using UNBLUR (Grant & Grigorieff 2015b), and the contrast transfer function of each was determined using CTFFIND4 (Rohou & Grigorieff 2015) in the RELION workflow (Scheres 2012). First, particles were classified using reference-free 2D classification to remove bad particles. 3D classification was then performed to search for a homogenous subset of particles using a *de novo* 3D reference created with AUTO3DEM software (Yan et al. 2007) filtered to  $\sim 60$   $\text{\AA}$  resolution and no symmetry was imposed during the classification. The same 3D reference was used for 3D refinement, and post-processing was applied to the final maps to appropriately mask the model, B-factor sharpening, and resolution validation as described in (Scheres 2012). The final resolution was determined using the so-called “gold standard FSC” at 0.143 FSC criterion (Scheres & Chen 2012) as 3.9  $\text{\AA}$ . The same workflow was performed using only three frames drift-corrected and averaged using UNBLUR and SUMMOVIE (Grant & Grigorieff 2015b) that improves the final resolution to 3.8  $\text{\AA}$ . Local resolution estimation was

done with ResMap (Kucukelbir A et al.). Figures were prepared using the software Chimera (Pettersen EF et al.).

### ***Model Building and refinement***

The homologue structure CPMV (pdb: 5a33) was used as reference to build the asymmetric unit of ArMV model in Coot (Emsley and Cowtan 2004). Prior to refinement, the pixel size and experimental weight values were optimized in REFMAC5 (Murshudov et al. 2011). Reference and secondary structure restraints were applied using ProSMART (Brown et al. 2015). The neighbours of the asymmetric were then created in Coot to assess for inter-chain clashes and to prevent refinement in nearby densities. The asymmetric unit and its neighbours were automatically refined in REFMAC5. After each refinement, MolProbity (Chen et al. 2010) was used to assess for the model quality, and Coot (Emsley and Cowtan 2004) for manual corrections before other REFMAC5 rounds. The fully refined model for overfitting validation using the two half maps as described in (Brown et al. 2015).

## Results

Viral particles of BBSV were purified from infected leaves. Images of Frozen-hydrated BBSV virion were first recorded using a JEOL 2200FS electron microscope operating at 200 KV equipped with a 4K X 4K slow-scan CCD camera (Gatan inc.). Visual inspection of images reveals low concentration levels of viral particles, mainly composed of RNA-filled capsids and, occasionally, empty particles. Although the number of BBSV particles was not high, a preliminary cryo-EM study of both types of particle was performed (Supplementary Fig. 1). The first initial reconstructions were obtained using the Imagic (van Heel & Keegstra 1981) and Auto3dem packages (Baker and Cheng, 1996; Fuller et al., 1996; Ji et al., 2006; Marinescu and Ji, 2003), revealing that the outermost layer corresponds to the capsid protein shell formed by self-assembled CP, whereas RNA contributes to the densities of the innermost layers.

Although preliminary, the three-dimensional map of RNA-filled BBSV particle was iteratively refined up to 8.5 Å, indicating that the viral particles are homogeneous and suitable for further structural characterization at high resolution. We therefore collected a cryo-EM data set on a FEI Titan Krios microscope using a Falcon II direct electron detector. On a total of 2899 micrographs recorded, 1779 were used for further processing. Thus, 37, 326 particles were semi-automatically selected. Micrographs were recorded with an exposure movie consisting of 7 frames and computationally corrected for microscope drift and beam-induced movement (Grant & Grigorieff 2015b). Iterative rounds of two-dimensional (2D) and three-dimensional (3D) classifications were then used to select a homogeneous subset of 36, 261 particles for 3D structure refinement.

The final resolution was estimated at 3.8 Å by the so-called gold-standard FSC using the 0.143 criterion (Scheres & Chen 2012). The data set was split into two independent halves during image processing to prevent over fitting. Besides, the high resolution noise substitution method has been showed to be a robust way to assess the amount of noise in high spatial frequencies and has been used as a phase randomization beyond 4.5 Å in order to validate the resolution (Chen et al. 2013). The FSC of randomized phase data set was then subtracted to the FSC of the original data set to give the true FSC curve (Supplementary figure 2). The final density map was sharpened using an empirically derived *B*-factor of  $-202 \text{ \AA}^2$  to 3.8 Å resolution (Fig 2A and B, EMD-xxxx).

As shown in [Fig. 2](#), the resolution of the wild-type BBSV map is high enough to clearly resolve amino acid side chains in the density (Fig. 2C and D). The atomic model of the cowpea mosaic virus capsid protein (PDB: 1ny7), which shares 55% of sequence identity with BBSV was used as a template for the model building. The process used to perform this step is presented as a flow chart in Figure 3. The template was first fitted into the asymmetric unit of the BBSV and manually mutated with BBSV amino acid sequence. Modeling started from bulky side chains to each extremity of protein chains. Structure restraints from the cowpea mosaic virus template were generated prior to model refinement using ProSMART (Brown et al. 2015). Therefore, the model was refined and then assessed for geometry quality in an iterative fashion until optimum. The optimized model was fitted in one of the two halves maps to be further refined, the new fitted model was then fit in the other half map. Finally, the correlation value between the model and the second half map was done to assess for noise overfitting. The resulting model contained information for the majority of the polypeptide sequence, critically including the terminal residues in the C- and N- terminal regions of S subunit.

As expected, the BBSV CP architecture (Fig. 2E) is similar to that of other Comoviruses. Briefly, the BBSV capsid is formed by 60 copies of the CPs arranged according to a pseudo  $T = 3$  symmetry. The L subunit is composed of two jellyroll  $\beta$ -barrel domains while the S subunit forms a single  $\beta$ -barrel domain. The outer radius at the fivefold, twofold and threefold axes are 153 Å, 128 Å and 136 Å respectively.

The superimposition of BBSV and CPMV coat proteins mainly reveals local structural discrepancies in loops and in the C-terminal part of the S subunit. The EM density map of the unsharpened BBSV map shows that the C-terminus part of S subunit forms a bridge with neighboring S subunits that composes the turret at the icosahedral fivefold vertex (Fig. 4A, green dot line). In the sharpened BBSV map, these connecting densities between S subunits disappeared, but well-resolved densities with clear density for side chains could be observed at the basis of the penton, which indicates that the C-terminal segment of S subunit then becomes ordered once more. These extra densities have been colored in green in Figure 4B and would correspond to residues 182–189 in the C terminus. When playing with the threshold of the sharpened BBSV map, Arg182 and Val183 are also visible. We have been able to build a convincing model into this region of the map corresponding to residues 185 to 189 of S subunit (Fig. 4C).

At this stage of our study, we can now reconsider BBSV CP–RNA interplays. Wild-type BBSV particles have packaged either the full-length single-stranded RNA-1 or RNA-2, so the density present inside the density map is an averaged picture of both RNA molecules. The Fig. 5A shows a central view of the unsharpened BBSV map. As previously described, we see a significant extra density inside the capsid that we ascribe to packaged RNA genome. The RNA appears as a concentric shell of density just beneath the capsid extending out through the fivefold symmetry axis. The shell of density has a thickness of  $\sim 20$  Å. Three different interaction areas are observed which are spread into the S and L subunits of the capsid. Strongest RNA interactions with the capsid shell exist near CP fivefold axis through Arg382 while other interactions take place in the L subunits through residues Asn175 and Gln190.

## ***Discussion***

The preliminary study allowed us to compute the 3D cryo-EM maps of RNA-filled and empty BBSV particles revealing the overall BBSV organization and allowing the unambiguous distinction between CP and RNA regions. Although preliminary, refinement of the RNA-filled BBSV virion allowed us to compute a density map at 8.5 Å resolution, indicating that BBSV particles were homogeneous in size and shape, and consequently suitable for a structural characterization at higher resolution. We then collected cryo-EM dataset on a Titan Krios microscope equipped with a Falcon II direct detector camera. As with this technical configuration, only seven frames could be recorded, we decided to apply an electron dose of 3.5 electrons/Å<sup>2</sup> per frame, without recording the two first ones. Image processing allowed us to compute a final map of RNA-filled BBSV virion at an average resolution of 3.8 Å. Local resolution estimation with ResMap shows that many regions reach lower resolution in the central part of the capsid shell (Fig. 6A and B) while main flexible parts of the map are related to pentons.

Contrary to other cryo-EM studies of viral particles, it was not possible to define a smaller set of particles producing a better resolution for BBSV reconstruction. Iterative cycles of 2D and 3D classification allowed us to exclude bad particles but not to distinguish between different BBSV virion conformations. We suppose that it is due to the fact that during the purification process, both type of BBSV particles having incorporated RNA-1 and RNA-2 are present.

The direct sequence homology are 54%, 54% and 75% between BBSV and CPMV, BPMV and RCMV coat proteins respectively. So, the overall structure of BBSV CPs revealed, as expected, is very similar to that of other comoviruses. Each icosahedral asymmetric unit is occupied by one S and one L subunit. However, some significant structural differences have been identified, mainly related to the position of loops and the N- and C- terminus of S subunits.

Although there is no available information about the cleavage of the BBSV C-terminus of S subunit, our cryo-EM map confirmed that domain is proteolytically cleaved at position 189 during maturation.

The existing structural information for the CPMV capsid whatever by X-ray crystallography or cryo-EM show the C terminus of S subunit after cleavage (ending at residue 189) in an extended conformation running across the exterior surface of the capsid. In case of BBSV,

residues 181 to 189 of S subunit adopts a different conformation but still acting as a bridge between S subunits forming pentons. These ordered C-terminal segments described here form an intimate network of interactions with the neighbouring S subunit around the pentameric ring that forms the fivefold vertex of the particle. It is worth noting that local resolution estimation indicates that external parts of pentamers are the most variable regions of BBSV virions. In the map, the location of residues 181 to 189 of S subunit corresponds to an average position, sufficiently redundant to be present in the cryo-EM map of BBSV. In this context, Arg565 and Asp418 from the neighboring S subunit are involved in the interaction with the C-terminus domain of the S subunit. When inspecting the surface representation of the local resolution map, we can see that the most variable part of the map corresponds approximately to the place of C-terminus of S subunit. A high variable region is located in the upper side of the S subunit forming the penton (black arrow in Fig. 6C), corresponding in case of CPMV to the interaction area between the S subunit and the small helix formed by residues 190 to 202 that are subsequently released during maturation.

This high variable region suggests that some BBSV particles should contain the entire C-terminus region of S subunit that could interact with the adjacent S subunit forming the penton, then suggesting that the cleavage of the C-terminal extremity of S subunit is not total and that the role of the C-terminal extremity of S subunit in viral assembly seems a general property in Comoviruses.

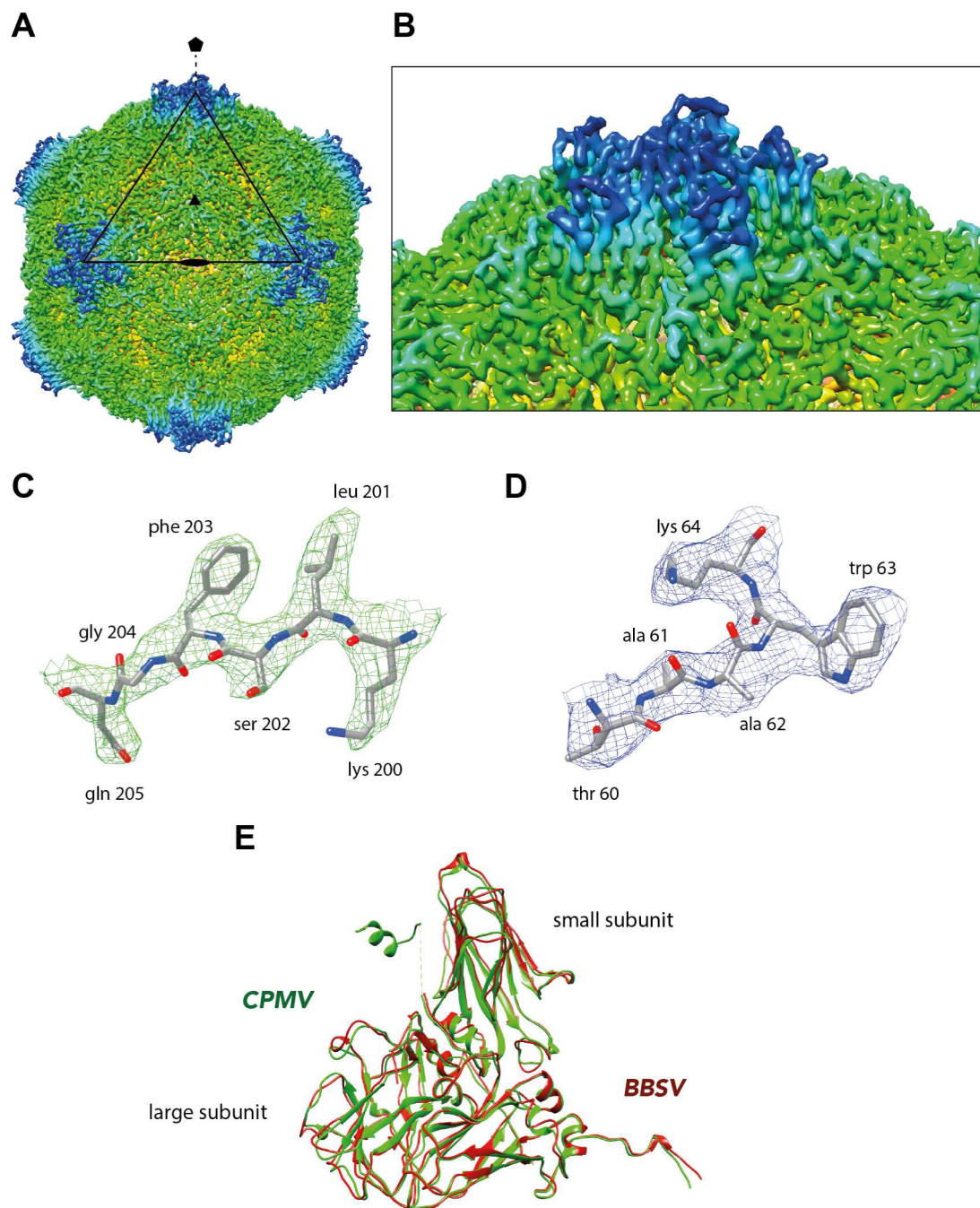
The BBSV cryo-EM map shows internal details of the capsid and reveals connections between the capsid shell and related RNA densities. The interaction between the BBSV capsid shell and the RNA density involve both L and S subunits, with a strong connection involving Arg382 observed around the fivefold symmetry axis. This is different from the CPMV since the interaction between the CPMV capsid shell and the RNA density is associated with Arg17 and Trp190, both from the L subunit, with Trp190 being by far the strongest density feature connecting the capsid shell to the RNA. However, the position 190 present in both BBSV and CPMV seems to be a key point in capsid RNA interaction. Although the general form of the BBSV packaged RNA is also different, since only one layer is observed in our cryo-EM map whereas CPMV displays at least two layers of RNA, we also observed that this layer of RNA is dodecahedral rather than icosahedral with the strongest RNA feature in the map forming a truncated dodecahedral cage beneath the capsid shell (Supplementary figure 3). This is strongly reminiscent of RNA packaging in insect viruses of the *Nodaviridae*, where ordered genomic RNA is packaged as a dodecahedral cage in both X-ray and EM structures (Tang et al., 2001; Johnson et al., 2004). Nearly 20 % of the

packaged RNA in BPMV binds to the capsid interior in a symmetric fashion forming a trefoil shape centered about the threefold particle axes. Most of interactions between the RNA and protein are nonbonded electrostatic and van des Waals interactions although a few specific contacts may exist. Whatever in CPMV or BBSV, the two comoviruses whom structures were resolved by cryo-EM, it was not possible to observe any repeated organization in RNA layer under the threefold symmetry axes.

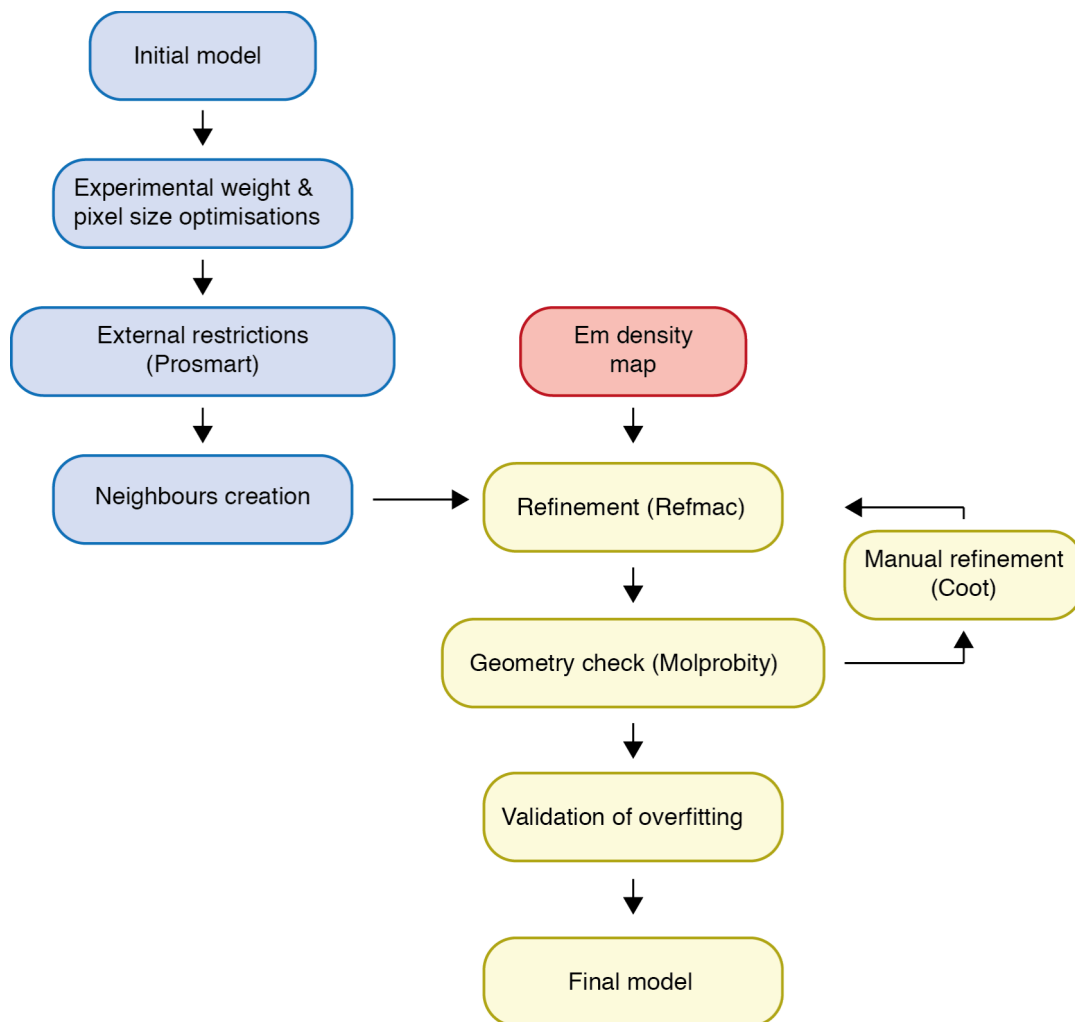
In conclusion, this cryo-EM structural study allowed us to determine the density map of BBSV at 3.8 Å resolution and to compute an atomic model of coat proteins forming the capsid of viral particles. As expected, this study confirms that the overall organization of capsid proteins is conserved among comoviruses but also reveals structural discrepancies specific to BBSV viral particles, as residues involved in RNA/capsid interactions, as well as the organization and interactions of the C-terminal domain of the small coat protein, a key region involved in viral assembly.



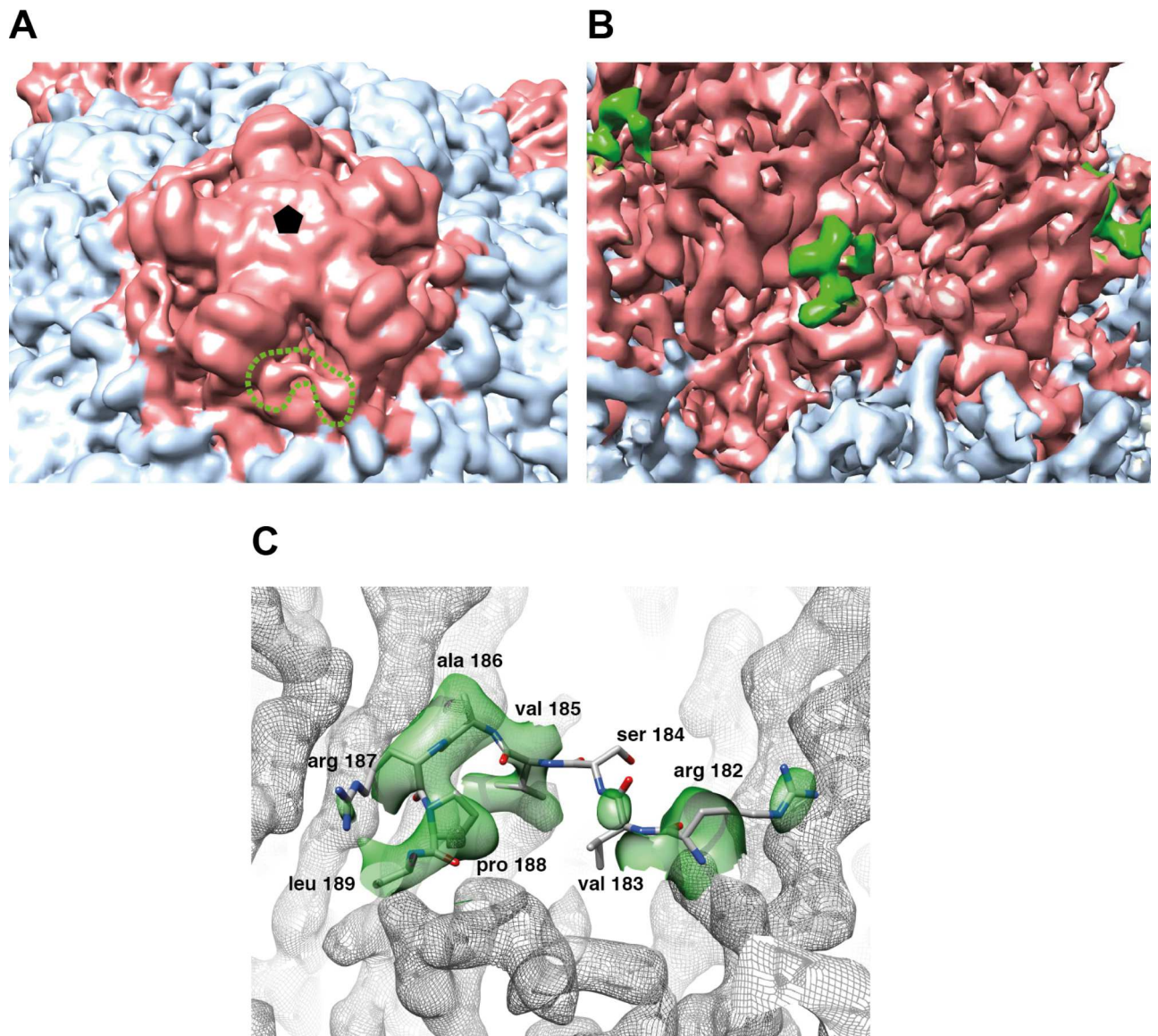




**Figure 2. The cryo-EM structure of BBSV.** (A) Icosahedral organization of BBSV determined at 3.8 Å resolution. (B) Closed-view of the top five-fold axis. (B and C) Side-chain details of the density map superimposed with atomic models of the large and small subunits respectively. (E) Structure match of BBSV (red) with CPMV (green) showing similar spatial arrangements.

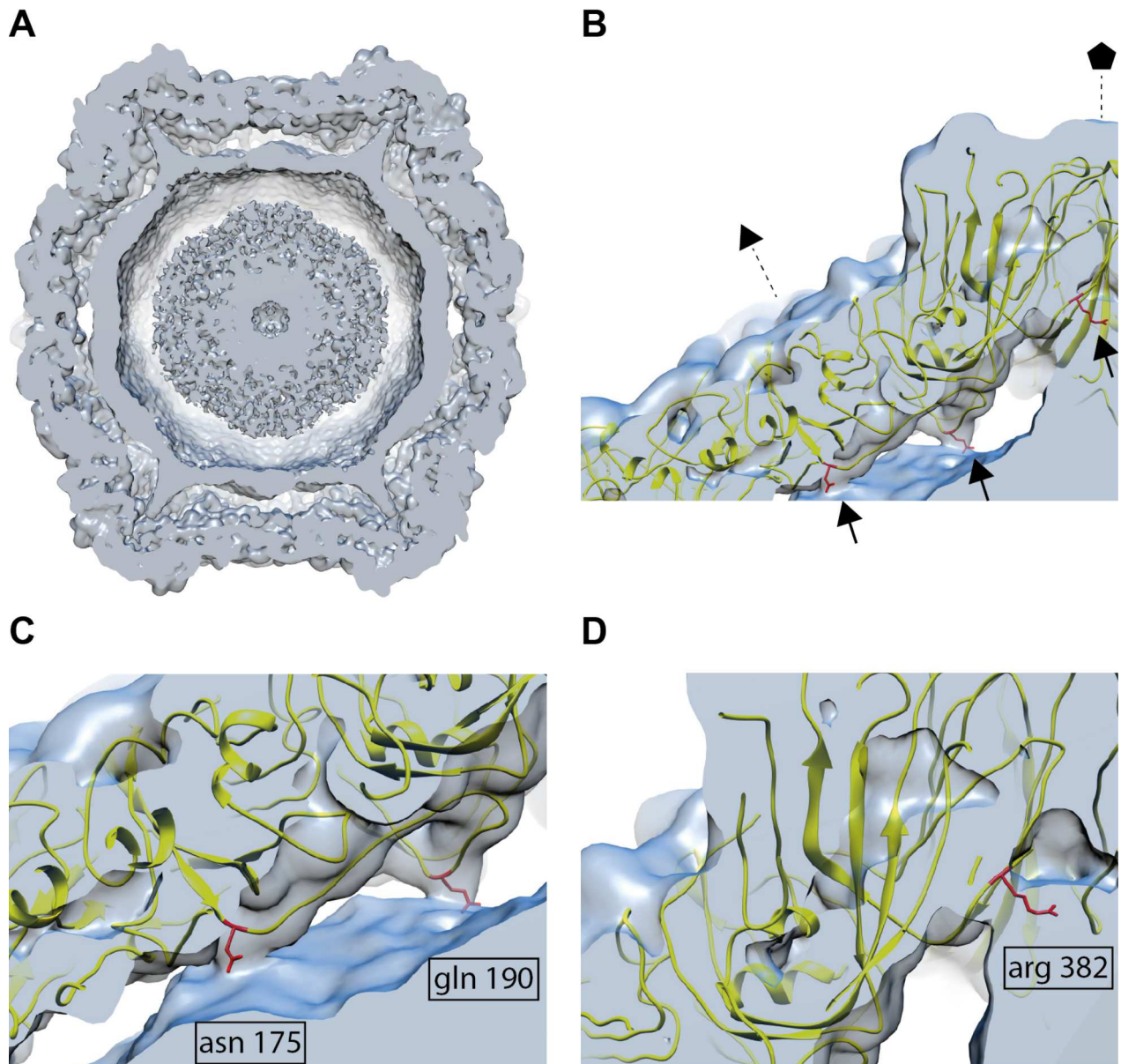


**Figure 3. The flowchart of the atomic model refinement and validation.**

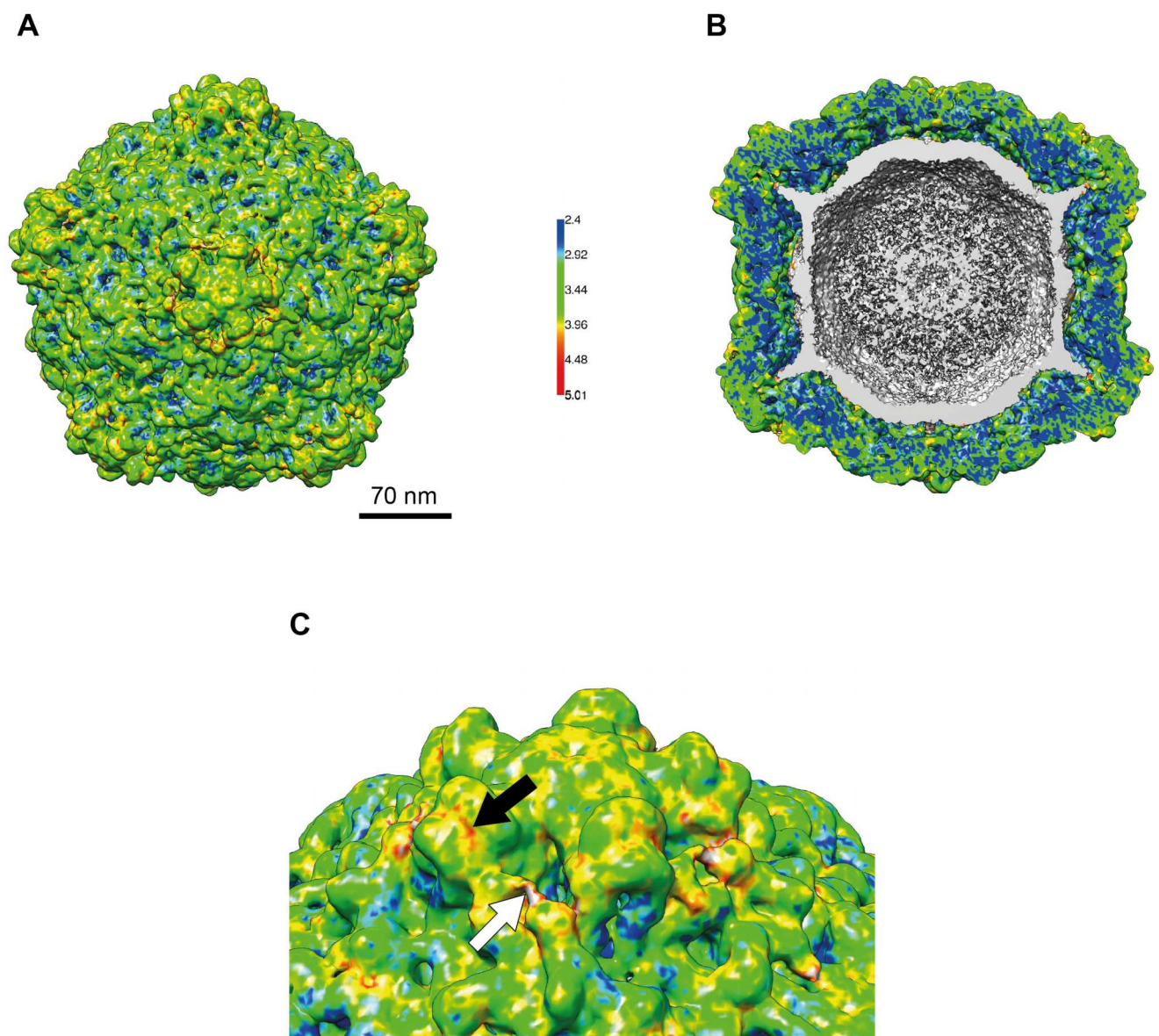


**Figure 4. The structure of the C-terminal extension of the small subunit.** (A) Closed-view of the C-terminal extension observed in the unsharpened BBSV map. The corresponding density is forming a bridge between adjacent arms of the five-fold axis (surrounded by green dotted lines). (B) Closed-view of the C-terminal extension observed in the sharpened BBSV map (coloured in green). (C) Closed-view of the atomic model inside the C-terminal extension from the sharpened BBSV map, where the density of the amino acid 184 is not visible.





**Figure 5. Interactions between the capsid protein and the RNA genome.** (A) Central view of the unsharpened BBSV map, the RNA genome is coloured in yellow, the large capsid subunit in red and the small in blue. (B) Closed-view of the three interaction sites between the capsid and the RNA genome. (C) Closed-view of the two interactions between the RNA genome and the residues Asn175 and Gln190 of the small subunit of the capsid protein. (D) Closed-view of the two interactions between the RNA genome and the residue Arg382 of the large subunit of the capsid protein. (E) Atomic model of BBSV observed from inside of capsid in which residues involved in CP-RNA interactions are drawn with spheres.



**Figure 6. Surface representation of the local resolution map.** (A) Surface view of the unsharpened cryo-EM map colored by local resolution as evaluated by ResMap. (B) Cut-through view of the unsharpened cryo-EM map colored by local resolution as evaluated by ResMap. (C) Top-view of the local resolution of the 5-fold axis with the C-terminus of S subunit (white arrow). (D) High variable region located in the upper side of the S subunit forming the penton (black arrow).

## ***Acknowledgments***

F.L. acknowledges funding from the Gouvernement de Nouvelle-Calédonie.

We would like also to thank the « Unité mixte de services » Biocampus (Montpellier) and the CBS for their constant support. The CBS is a member of the French Infrastructure for Integrated Structural Biology (FRISBI), a national infrastructure supported by the French National Research Agency (ANR-10-INBS-05).

## ***Accession numbers***

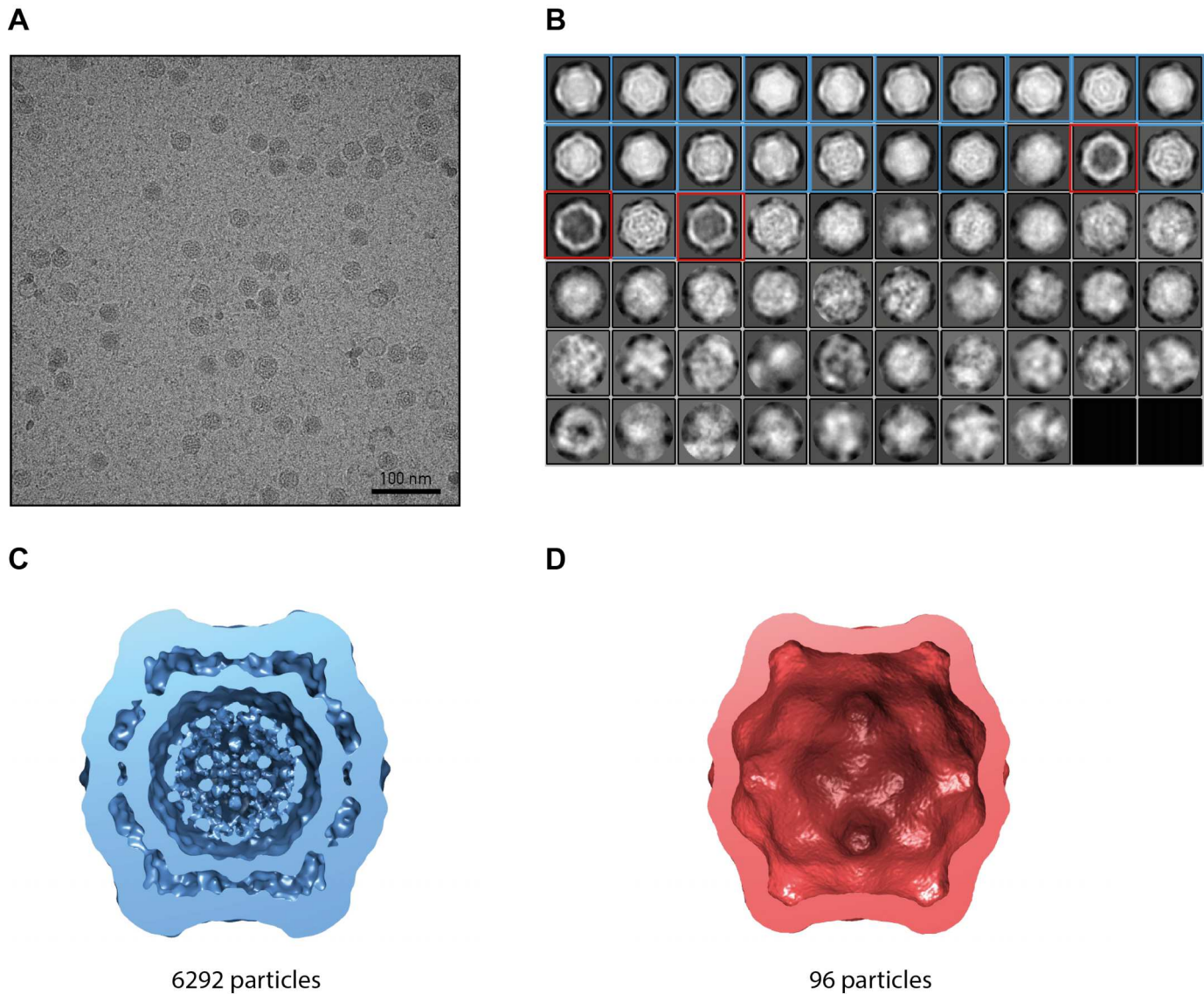
Cryo-EM density map of BBSV has been deposited in the Electron Microscopy Data Bank with accession number EMD-XXXX. The atomic structure of the S and L coat proteins has been deposited in the Protein Data Bank under ID code XXXX.



## Appendix A. Supplementary data

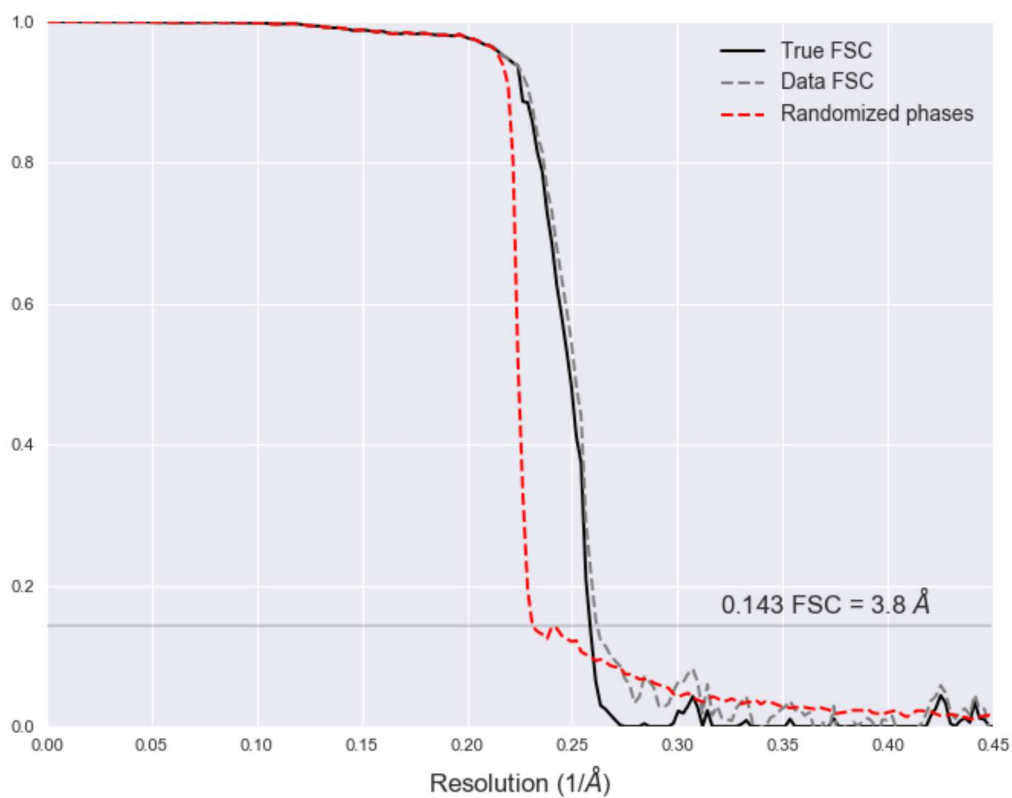
Supplementary data associated with this article can be found, in the online version, at <http://>

.....



**Supplementary Figure 1. The preliminary cryo-EM study of BBSV.** (A) A two-dimensional cryo-EM image of the BBSV particles, showing strong background and heterogeneity of the sample after *in planta* purification. (B) Two-dimensional class averages of BBSV particles permitting to separate particles with RNA (blue) and without RNA (red). (C and D) Three-dimensional reconstructions of a full particle (blue) and an empty particle (red) from the two-dimensional class averages.





**Supplementary Figure 2. Resolution estimation and validation.** Three FSC curves computed between two independent three-dimensional reconstructions generated from two independent data sets using different methods. The “True” FSC is shown with a black line and the gold standard FSC is shown with a grey dotted line, which is nearly identical with the “True FSC”. The two FSC curves validate the resolution of the density map at 3.8 Å; the FSC phase randomization at 4.5 Å is shown with a red dotted line.

## References

- Al-Khalaf M, Makkouk KM, Kasem AH. Seed transmission of broad bean stain virus in lentil with respect to genotype variability and seed size. *Arab Journal of Plant Protection* 20:106-110 (2002).
- Baker TS, Cheng RH. A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J Struct Biol.* 116:120–130 (1996).
- Brown, A., Long, F., Nicholls, R.A., Toots, J., Emsley, P., Murshudov, G. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 71, 136–153 (2015).
- Cañizares, M. C., Taylor, K. M. & Lomonosoff, G. P. Surface-exposed C-terminal amino acids of the small coat protein of Cowpea mosaic virus are required for suppression of silencing. *J. Gen. Virol.* 85, 3431–3435 (2004).
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 66, 12–21 (2010).
- Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H.W., Henderson, R. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* 135, 24–35 (2013).
- Chen, Z. G. et al. Protein-RNA interactions in an icosahedral virus at 3.0 Å resolution. *Science* 245, 154–159 (1989).
- Emsley, P., Cowtan, K. Coot -model building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* D60, 2126–2132 (2004).
- Fiedorow Z, Szlachetka-Wawrzyniak. Transmission of broad bean stain virus (BBSV) by seeds of pea *Pisum sativum* L.). *Plant Breeding and Seed Science* 46:81-88 (2002).
- Fuller SD, Butcher SJ, Cheng RH, Baker TS. Three-dimensional reconstruction of icosahedral particles--the uncommon line. *J Struct Biol.* 116:48–55 (1996).
- Grant, T., Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *Elife* 4, e06980 (2015).

Hesketh, E.L., Meshcheriakova, Y., Dent, K.C., Saxena, P., Thompson, R.F., Cockburn, J.J., Lomonosoff, G.P., Ranson, N.A. Mechanisms of assembly and genome packaging in an RNA virus revealed by high-resolution cryo-EM. *Nat. Commun.* 6, 10113 (2015).

Huynh, N.A.T., Hesketh, E.A.L., Saxena, P., Meshcheriakova, Y., Ku, Y.C., Hoang, L.A.T., Johnson, J.A.E., Ranson, N.A.A., Lomonosoff, G.A.P., Reddy, V.A.S. Crystal Structure and Proteomics Analysis of Empty Virus-like Particles of Cowpea Mosaic Virus. *Structure* 24, 567–575 (2016).

Johnson, K. N., Tang, L., Johnson, J. E. & Ball, L. A. Heterologous RNA encapsidated in Pariacoto virus-like particles forms a dodecahedral cage similar to genomic RNA in wild-type virions. *J. Virol.* 78, 11371–11378 (2004).

Kucukelbir A, Sigworth FJ, Tagare HD. Quantifying the local resolution of cryo-EM density maps. *Nat Methods.* Jan; 11(1):63-5 (2014).

Kumari SG, Makkouk KM. Evaluation of different ELISA procedures for the detection of pea seed-borne mosaic potyvirus and broad bean stain comovirus in lentil leaf extracts. *Arab Journal of Plant Protection* 11:86-91 (1993).

Kumari SG, Makkouk KM. Transmission of broad bean stain comovirus and broad bean mottle by weevils in Syria. *Phytopathologia Mediterranea* 35:124-126 (1996).

Lin, T. & Johnson, J. E. Structures of picorna-like plant viruses: implications and applications. *Adv. Virus Res.* 62, 167–239 (2003).

Lin, T. et al. The refined crystal structure of cowpea mosaic virus at 2.8 Å resolution. *Virology* 265, 20–34 (1999).

Lin, T. et al. Structural fingerprinting: subgrouping of comoviruses by structural studies of red clover mottle virus to 2.4 Å resolution and comparisons with other comoviruses. *J. Virol.* 74, 493–504 (2000).

Lomonosoff, G. P. & Johnson, J. E. The synthesis and structure of comovirus capsids. *Prog. Biophys. Mol. Biol.* 55, 107–137 (1991).

Mabrouk O, Mansour AN. Effect of pea seedborne mosaic and broad bean stain viruses on lentil growth and yield in Jordan. *Scientia Horticulturae* 73:175-178 (1998).

Makkouk KM, Azzam OI. Detection of broad bean stain virus in lentil seed groups. *LENS Newsletter* 13:37-38 (1986).

Marinescu DC, Ji Y. A computational framework for the 3D structure determination of viruses with unknown symmetry. *J Parallel and Distrib Comput.* 63:738–758 (2003).

Mindell, J.A., Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* 142, 334–347 (2003).

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67, 355–367 (2011).

Ochoa, W. F., Chatterji, A., Lin, T. & Johnson, J. E. Generation and structural analysis of reactive empty particles derived from an icosahedral virus. *Chem. Biol.* 13, 771–778 (2006).

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* Oct; 25(13):1605-12 (2014).

Rohou, A., Grigorieff, N. CTFIND4: Fast and accurate defocus estimation from electron micrographs, *Journal of Structural Biology* (2015).

Sainsbury, F., Saunders, K., Aljabali, A. A. A., Evans, D. J. & Lomonosoff, G. P. Peptide-controlled access to the interior surface of empty virus nanoparticles. *Chembiochem.* 12, 2435–2440 (2011).

Scheres, S.H.W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* 180, 519–530 (2012).

Scheres, S.H.W., Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* 9, 853–854 (2012).

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., Ludtke, S.J. EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* 157, 38–46 (2007).

Tang, L. et al. The structure of pariacoto virus reveals a dodecahedral cage of duplex RNA. *Nat. Struct. Biol.* 8, 77–83 (2001).

Taylor, K. M., Spall, V. E., Butler, P. J. & Lomonosoff, G. P. The cleavable carboxyl terminus of the small coat protein of cowpea mosaic virus is involved in RNA encapsidation. *Virology* 255, 129–137 (1999).

van Heel, M., Harauz, G. Resolution criteria for three-dimensional reconstructions. *Optik* 73, 119–122 (1986).

van Heel, M., Portugal, R., Rohou, A., Linnemayr, C., Bebeacua, C., Schmidt, R., Grant, R., and Schatz, M. Four-Dimensional Cryo Electron Microscopy at Quasi Atomic Resolution: "IMAGIC 4D". *International Tables for Crystallography Volume F: Crystallography of biological macromolecules (Second Edition)* (2011).

Yan, X., Sinkovits, R.S., Baker, T.S. AUTO3DEM-an automated and high throughput program for image reconstruction of icosahedral particles. *J. Struct. Biol.* 157, 73–82 (2007).

### 3. Conclusion

La structure de la capside du BBSV a pu être caractérisée par cryo-ME à 3,8 Å de résolution et le modèle atomique, composé d'une petite et d'une grande sous-unité a ainsi pu être construit et affiné. Les acides aminés en interaction avec le génome ARN ont pu être identifiés (arginine 4 de la petite sous-unité, asparagine 175 et glutamine 190 de la grande sous-unité). La structure du BBSV a été comparée à celle du virus de la mosaïque du niébé (*Cowpea Mosaic Virus*, ou CPMV), un autre Comovirus dont la structure a déjà été résolue par cryo-ME et cristallographie aux rayons X. Comme attendu, les deux structures sont très similaires, mais une différence structurale a été identifiée au niveau de la partie C-terminale de la petite sous-unité qui est impliquée dans l'assemblage de la capside. A l'instar du CPMV, la partie C-terminale de la petite sous-unité du BBSV se retrouve clivée après la leucine 189 pendant la maturation du virus. Chez le BBSV, l'extrémité C-terminale forme un pont entre les petites sous-unités qui forme les pentamères du virus, alors que chez le CPMV, elle va s'élever à la surface de la capside. L'estimation de la résolution locale de la capside du BBSV laisse suggérer que certaines particules virales contiennent une extrémité non clivée qui pourrait interagir avec la protéine adjacente du pentamère. La manière dont l'extrémité C-terminale de la petite sous-unité intervient dans l'assemblage du virus semble ainsi être commune parmi les Comovirus.

# CHAPITRE 5



# Chapitre 5 : Etude structurale du virus de la mosaïque du chou-fleur

---

## 1. Introduction

Le Virus de la Mosaïque du chou-fleur (CaMV) est le virus le plus étudié du genre *Caulimovirus* appartenant à la famille des *Caulimoviridae*. Il fut le premier virus de plante découvert comme contenant un génome d'ADN. Le CaMV est un pararétrovirus, dont le mode de répllication fait intervenir une transcriptase inverse virale qui permet la conversion de l'ADN viral en ARN (Hohn 2013).

### 1.1. Génome du virus

Il possède un génome d'ADN circulaire doubles brins d'environ 8000 pb, comprenant sept cadres de lectures ouverts (*open reading frames* ou ORFs), dont six ont été identifiés. Le génome se retrouve encapsidé dans une particule icosaédrique non enveloppée d'environ 500 Å de diamètre et composée de 420 protéines de capsides formant une triangulation  $T = 7$ . Les ORFs vont être transcrits en deux ARNs messagers : l'ARN messager 19S qui code seulement pour la protéine P6 issue de l'ORF VI, et l'ARN 35S qui code pour les protéines P1 à P5, respectivement issues des ORFs I à V (Figure 70).

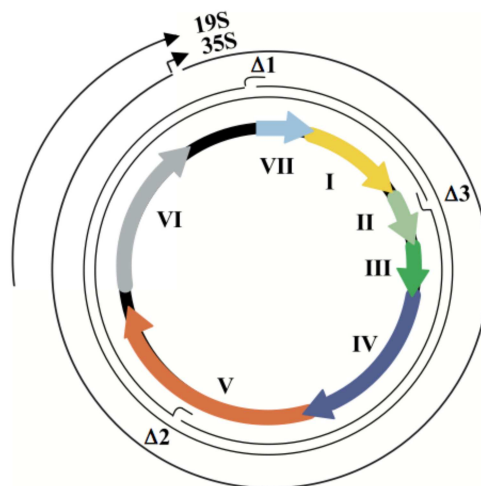


Figure 70 : Diagramme schématisé du génome du CaMV. Reproduit de (Haas et al. 2002)

La protéine P1 est une protéine de mouvement (MP). Elle permet le déplacement du virus à travers les cellules de la plante infectée. La protéine P1 forme une structure tubulaire insérée dans les plasmodesmes permettant de transporter le virus. Les plasmodesmes sont des canaux traversant la paroi cellulaire végétale et faisant communiquer deux cellules adjacentes. Les mouvements du virus entre cellules semble se faire par des interactions entre les domaines superhélice (motif structurel dans lequel deux à sept hélices alpha sont enroulées ensemble) de la protéine P1 formant les tubules, et le complexe transmissible, composée de la protéine P3 associée à la protéine de capsid P4 (Stavolone et al. 2005).

La protéine P2 est un facteur assistant la transmission (FAT). Elle sert de lien entre le récepteur supposé du virus présent à l'intérieur du stylet du puceron (extrémité des pièces buccales du puceron), et le complexe transmissible (Uzest et al. 2007; Leh et al. 1999).

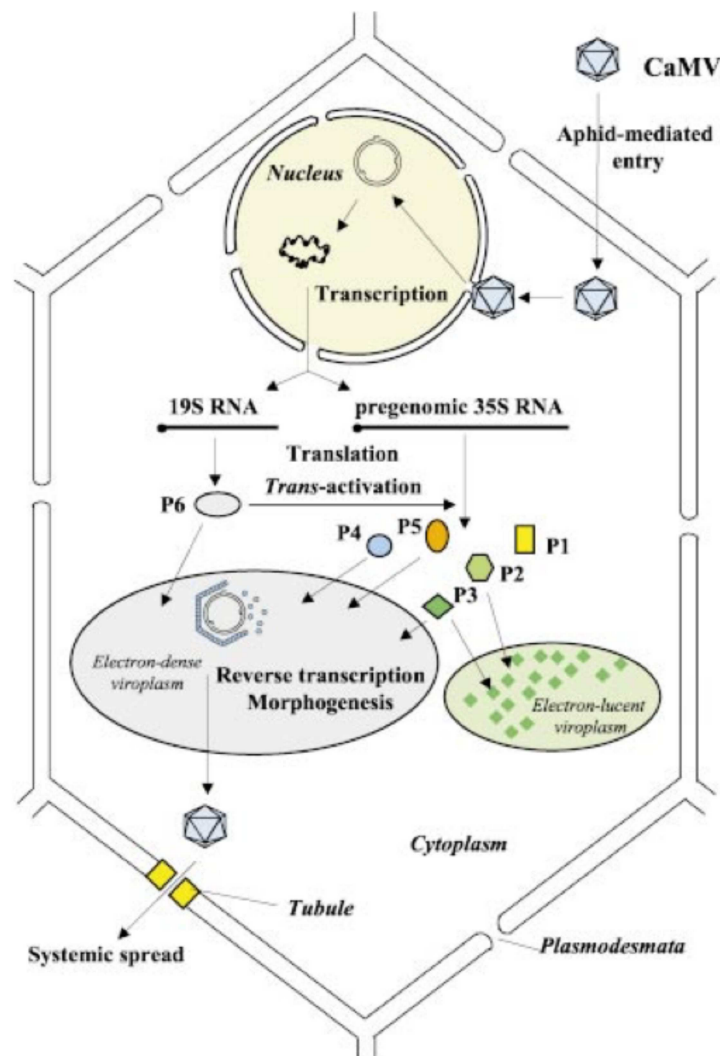
La protéine P3 est une protéine associée sous forme de dimère à la protéine de capsid P4, pour former le complexe transmissible. Elle possède un domaine superhélice en N-terminale, qui semble être responsable de la liaison du complexe transmissible avec les protéines P1 et P2 (Plisson et al. 2005).

La protéine P4 est une protéine de capsid formant le complexe transmissible avec la protéine P3. La protéine P5 est une transcriptase inverse synthétisant le génome viral. La protéine P6 est une protéine multifonctionnelle. Elle transactive la traduction de l'ARN 35S pour permettre la synthèse des autres protéines virales, et elle est responsable de la formation de viroplasmes, supposés être des usines virales où la synthèse de l'ADN viral par la transcriptase inverse P5 semble avoir lieu, et où les particules virales restent ensuite stockées (Haas et al. 2002).

## **1.2. Multiplication virale dans la cellule de plante**

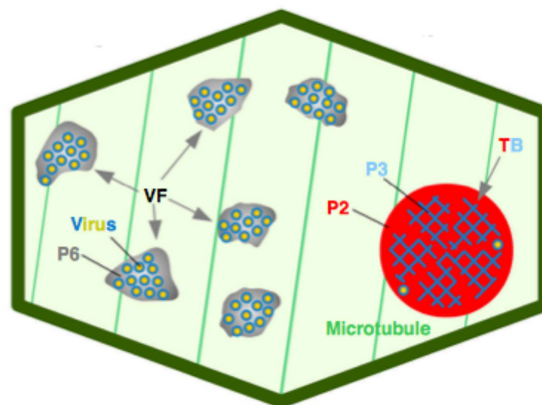
Lors de son entrée dans la cellule de la plante hôte, la particule est transportée jusqu'au pore nucléaire de la cellule et l'ADN viral entre dans le noyau. Les discontinuités présentes dans l'ADN viral sont réparées et l'ADN s'associe à des histones pour former un mini-chromosome. La transcription de l'ADN viral se fait alors par l'ARN polymérase II cellulaire qui synthétise l'ARNm 19S monocistronique et l'ARNm 35S polycistronique. Les ARNm sont transportés dans le cytoplasme, où l'ARNm 19S est traduit en protéine P6 qui va former les viroplasmes et va intervenir dans la traduction de l'ARNm 35S polycistronique. Les protéines P1 à P5 seront ainsi traduites en périphérie

des viroplasmes, et la transcriptase inverse P5 utilisera l'ARN 35S comme matrice pour synthétiser le nouvel ADN double-brin, destiné à être encapsidé dans de nouvelles capsides (**Figure 71**).



**Figure 71** : Cycle de multiplication du CaMV à l'intérieur d'une cellule de plante. Reproduit de (Haas et al. 2002).

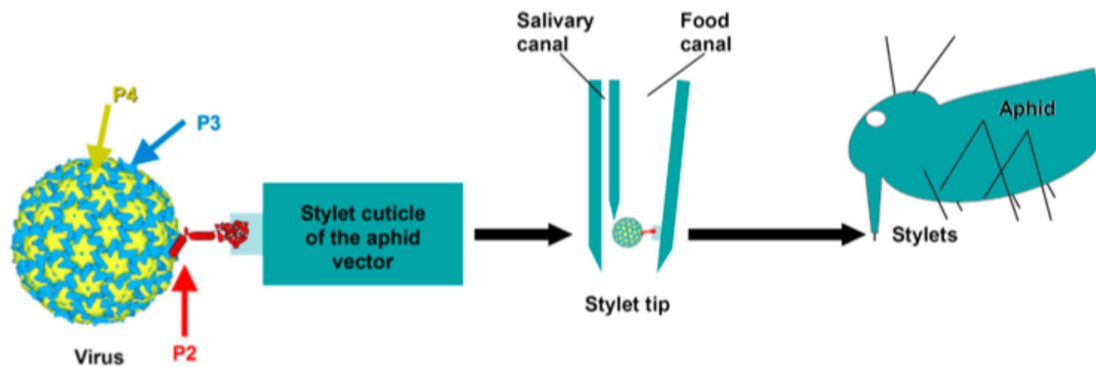
Après synthèse des protéines virales, la quasi-totalité des complexes transmissibles (nouveaux virions + P3) nouvellement synthétisés sera compartimentée dans plusieurs viroplasmes formés par la protéine P6. De plus, la totalité des protéines P2 se verra compartimentée à l'intérieur d'un corps de transmission, qui contiendra également des protéines P3 qui auront coagrégées avec P2, ainsi que quelques complexes transmissibles (Drucker et al. 2002). De cette manière, à l'intérieur d'une cellule infectée, il y aura une séparation spatiale entre les viroplasmes contenant le complexe transmissible virion:P3, et le corps de transmission contenant le complexe P2:P3 (**Figure 72**).



**Figure 72 : Séquestration des particules virales.** Séparation spatiale entre les viroplasmes formées par la protéine P6 (gris), contenant le complexe transmissible P4 (jaune) et P3 (bleu), et le corps de transmission où sont séquestrées les protéines P2 (rouge) et P3 (bleu). Reproduit de (Drucker & Then 2015).

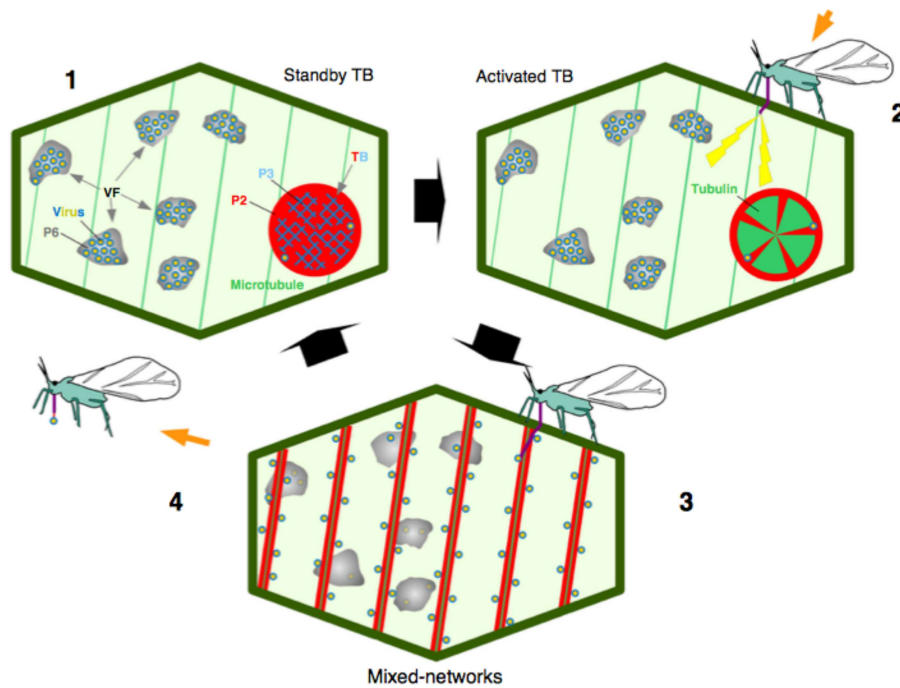
### 1.3. Étude de la transmission plante-vecteur du CaMV

Le CaMV est transmis selon le mode non circulant semi-persistant. Dans ce mode de transmission, il n'y a pas de grande spécificité du virus pour son vecteur, et un même virus peut être propagé par de nombreuses espèces de pucerons. Ainsi des pucerons non inféodés à une culture pourront jouer un rôle important dans la transmission virale par leurs simples piqûres d'essais. Le mode non circulant signifie que le virus ne circule pas à l'intérieur du système digestif du puceron, mais les particules virales restent attachées aux pièces buccales, et peuvent alors les transmettre immédiatement à une autre plante. La transmission est également qualifiée de semi-persistante, c'est-à-dire que la période de rétention du CaMV chez le puceron peut durer de quelques minutes à plusieurs jours. Le puceron vecteur devra alors se nourrir une nouvelle fois sur une plante infectée pour se « recharger » en particules virales. La transmission virale du CaMV, de la plante au vecteur, fait intervenir son facteur assistant à la transmission P2. La protéine P2 va jouer le rôle de pont moléculaire entre le récepteur présent à l'extrémité du stylet du puceron, et le complexe transmissible P4-P3 (**Figure 73**). La protéine P2 semble se lier au stylet du puceron au niveau de son extrémité N-terminale (Moreno et al. 2005; Uzest et al. 2007), alors que son extrémité C-terminale se lie à l'extrémité N-terminale de la protéine P3 (Leh et al. 1999).



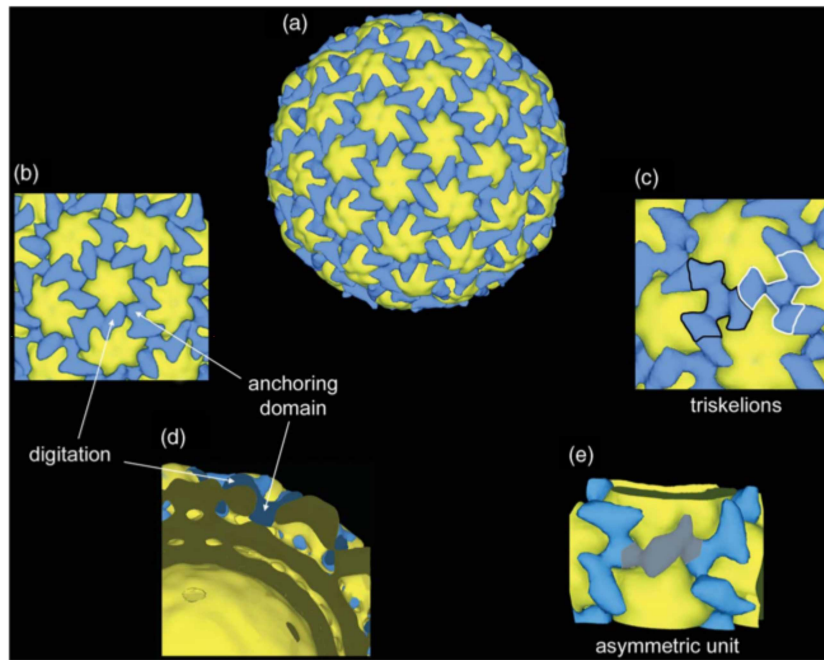
**Figure 73 : Le complexe transmissible P4:P3.** Il va s'attacher au stylet du puceron par l'intermédiaire de la protéine de transmission P2, servant de pont moléculaire entre le récepteur du puceron et le complexe transmissible. Reproduit de (Martiniere et al. 2013).

Comme énoncé précédemment, la totalité des protéines P2 se retrouve complexée avec P3 dans le corps de transmission. Cette structure va alors réagir à la présence du puceron, lorsque son stylet pénètre la cellule infectée. Un influx de tubulines arrive dans le corps de transmission qui est alors désintégré en quelques secondes (Martiniere et al. 2013), ce qui entraîne une redistribution de P2 sur le réseau de microtubules de la cellule. Au même moment, les virions sont libérés des viroplasmes et se lient aux protéines P2 présentes sur les microtubules, un réseau mixte composé de la protéine P2 et du complexe transmissible se forme sur les microtubules, permettant une dispersion homogène du virus à l'intérieur de la cellule. Le stylet du puceron entrant à des positions aléatoires dans la cellule, le réseau mixte donne alors une plus grande chance au virus d'être aspiré par le puceron. Après le départ du puceron, les protéines P2 se dissocient des microtubules et reforment un nouveau corps de transmission, de même pour les virions qui regagnent les viroplasmes, en attendant une nouvelle piqûre (**Figure 74**). Un stress mécanique ne semble pas être suffisant pour provoquer l'activation de la transmission, une combinaison entre un stress mécanique, et un stress chimique semble être nécessaire (Drucker & Then 2015).

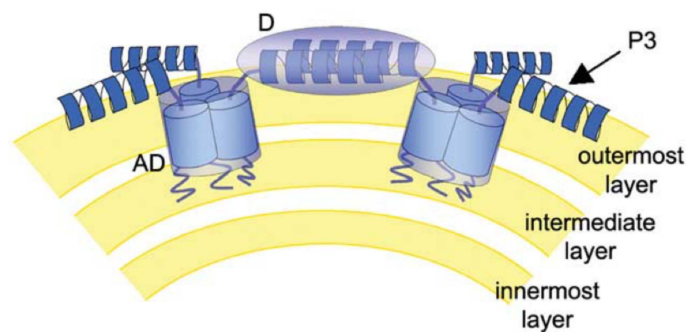


**Figure 74 : Modèle d'acquisition du CaMV.** (1) Séparation spatiale des viroplasmes et du corps de transmission après infection. (2) Un puceron vient effectuer une piqûre d'essai, déclenchant un influx de tubulines dans le corps de transmission qui se désintègre. Les virions sont libérés des viroplasmes au même moment. (3) Un réseau mixte, composé du complexe transmissible virion:P3 et de P2, se forme sur les microtubules, permettant l'absorption du virus par le puceron. (4) Après le départ du puceron, le réseau mixte se dissocie, P2 forme un nouveau corps de transmission et les virions rejoignent les viroplasmes. Reproduit de (Drucker & Then 2015).

De premières études structurales ont été menées pour mieux comprendre les interactions moléculaires régissant la transmission du CaMV. Les prédictions de structures secondaires indiquent que la partie N-terminale de P3, et la partie C-terminale de P2 seraient composées d'un enroulement d'hélices alpha formant des domaines de type superhélice. Les premières structures du CaMV nu et du complexe virion:P3 ont été résolues par cryo-ME à 18 Å et 26 Å respectivement. Une comparaison entre ces deux structures a permis de proposer une organisation de P3 qui forme des motifs triskèles (motif à trois jambes) à la surface de la capsid (Plisson et al. 2005) (**Figure 75**). Un modèle de l'interaction de P3 avec la capsid est proposé : La partie C-terminale de P3 est ancrée dans la capsid et forme un domaine d'ancrage constitué de trois monomères de P3, alors que sa partie N-terminale reste à la surface de la capsid et forme un dimère antiparallèle avec la partie N-terminale d'une protéine P3 voisine, par des interactions de type superhélice (**Figure 76**).



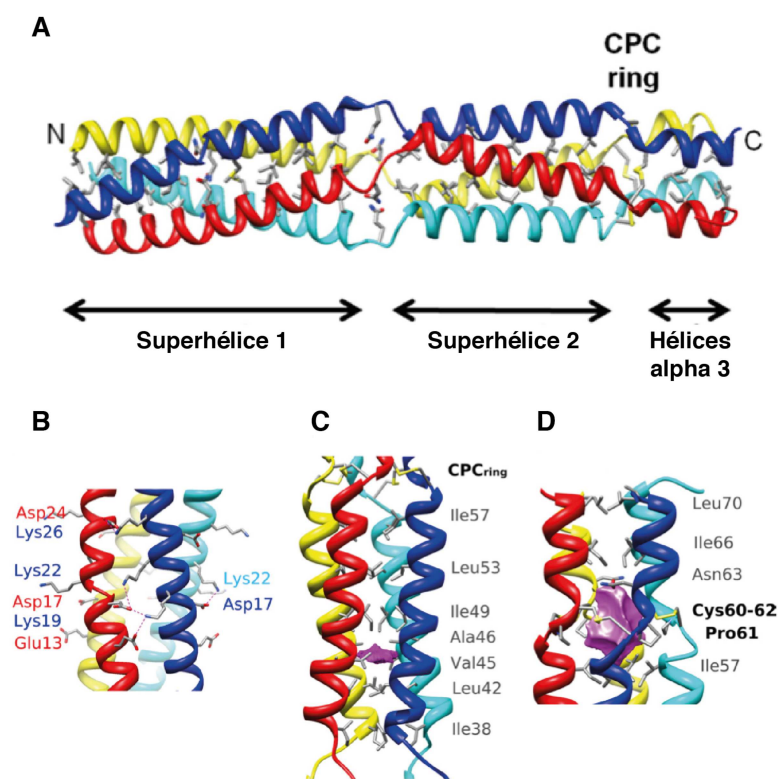
**Figure 75 : Visualisation de la carte de différence du CaMV obtenue par cryo-ME.** La capside est représentée en jaune et la protéine P3 en bleue. (a) Représentation du complexe transmissible. (b) Représentation d'un pentamère entouré par cinq hexamères. La digitation formée par P3 relie deux hexamères ou un hexamère avec un pentamère. (c) Vue en perspective de trois hexamères adjacents, montrant deux triskèles (noir et blanc). (d) Coupe du complexe transmissible montrant l'ancrage de la protéine P3 dans la couche externe de la capside. (e) Représentation de l'unité asymétrique d'une triskèle.



**Figure 76 : Modèle de l'interaction de la protéine P3 avec la capside du CaMV.** L'association de la partie C-terminale de trois monomères de P3 forment le domaine d'ancrage (AD) dans la capside, qui pénètre la couche intermédiaire interagit à la fois avec la capside et le génome du virus. À la surface de la capside, les parties N-terminale de P3 forme des dimères antiparallèles.



Plus tard, la partie N-terminale de la protéine P3 a été résolue par cristallographie aux rayons X (Hoh et al. 2010). Cette partie N-terminale est composée de trois segments d'hélices alpha, dont les deux premiers segments forment des domaines superhélices. Un anneau est également formé après la deuxième superhélice à partir d'un motif Cystéine-Proline-Cystéine (CPC). La protéine adopte une oligomérisation en tétramère parallèle en solution d'une longueur d'environ 120 Å et un diamètre d'environ 25 Å. La première superhélice (résidus 3 à 32) possède une hélicité gauche, une région hydrophobe et la formation de ponts salins et liaisons hydrogène interchaînes, qui stabilisent la formation de la superhélice. Au contraire, la deuxième superhélice (résidus 38 à 59) possède une hélicité droite et une région hydrophobe moins serrée, avec les résidus hydrophobes plus exposés au solvant par rapport au premier domaine. De plus, elle ne possède aucun pont salin ni aucune liaison hydrogène, la rendant moins stable que la première. Enfin, l'anneau formé par le motif CPC permet de créer quatre ponts disulfures interchaînes, stabilisant la partie C-terminale (**Figure 77**).

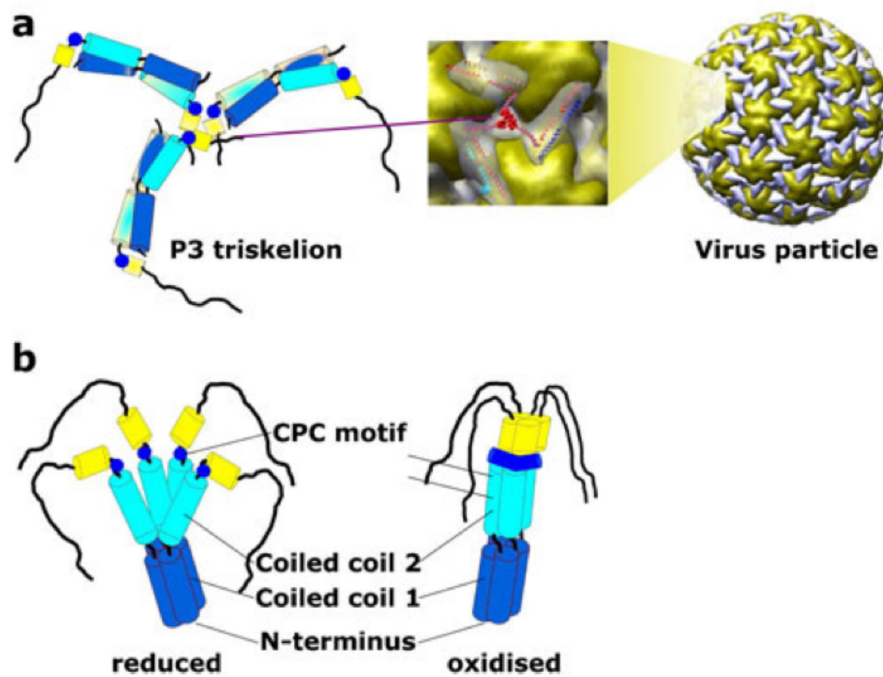


**Figure 77 : Structure de la partie N-terminale de la protéine P3 obtenue par cristallographie aux rayons X. (A)** Représentation en ruban du tétramère formé de deux domaines superhélices, et d'un troisième domaine formé par des hélices alpha. **(B)** Vue détaillée des ponts salins interchaînes (pointillés) de la superhélice 1. **(C)** Vue détaillée de la superhélice 2 avec sa région hydrophobe (magenta) et les acides aminés hydrophobes (bâtons gris). **(D)** Vue détaillée des quatre ponts disulfures formés par l'anneau CPC délimitant une région hydrophobe (magenta). Reproduit de (Hoh et al. 2010).



L'existence des deux oligomérisations différentes de P3 (en dimère quand liée au virus et en tétramère en solution) pourrait s'expliquer par l'existence d'une forme réduite et d'une forme oxydée de la protéine P3. En effet, dans l'environnement réducteur du cytoplasme de la cellule, il n'y aurait pas de formation des ponts disulfures qui stabilisent le tétramère. De cette manière, il pourrait exister une population hétérogène de la protéine P3 (**Figure 78**) :

- Une forme tétramérique partiellement parallèle, stabilisée uniquement par le premier domaine superhélice en N-terminale, puisque le deuxième domaine superhélice déjà moins stable, n'est plus stabilisé par l'anneau de ponts disulfures.
- L'absence de stabilisation par l'anneau de ponts disulfures entraîne la dissociation de certains tétramères qui se réassemble sur la capsid du virus en dimères antiparallèles.

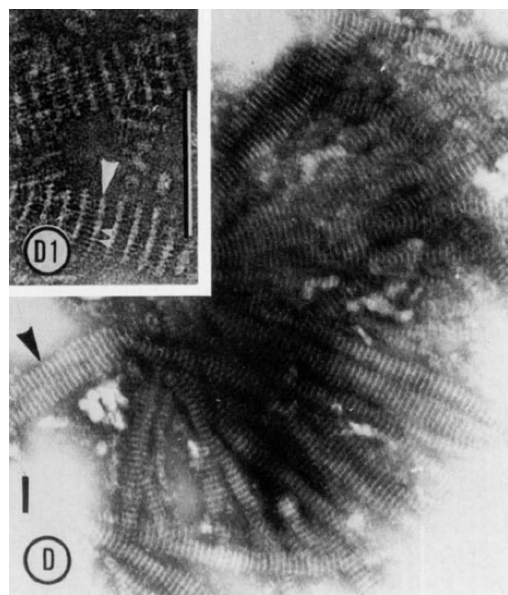


**Figure 78 : Les différentes formes de la protéine P3.** (a) P3 s'assemble en dimères antiparallèles à la surface de la capsid et l'interaction avec P2 se fait au niveau des superhélices des deux protéines. (b) P3 en solution s'assemble en tétramère parallèle. Sous forme réduite, la superhélice 2 est relâchée, permettant de libérer le site de liaison de P2. Sous forme oxydée, le motif CPC forme un anneau bloquant la superhélice 2, masquant le site de liaison de P2. Reproduit de (Bak et al. 2012).

L'existence de ces formes réduite et oxydée pourrait alors expliquer la rupture du corps de transmission. Sous forme réduite, la distorsion partielle du tétramère permet l'exposition du site de

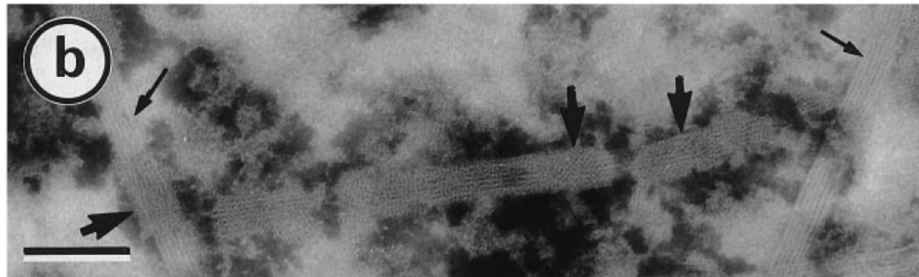
liaison pour la protéine P2, cette interaction pourrait être responsable de la formation des corps de transmission. Plus tard, le stress provoqué par la piqûre du puceron pourrait induire un environnement oxydant, entraînant la formation des ponts disulfures qui stabiliseront le tétramère. La protéine P2 n'interagit alors plus avec la protéine P3, entraînant la rupture du corps de transmission et la libération de P2 sur les microtubules (Bak et al. 2012).

Le mode de transmission non circulant a longtemps été vu comme le mode de transmission le plus simple se résume à une contamination des pièces buccales du puceron par le virus. Cependant, ces dernières études démontrent que cette transmission fait intervenir des interactions bien spécifiques entre la plante, le virus et le vecteur. Afin d'approfondir nos connaissances sur les mécanismes mis en jeu pendant la transmission du virus au vecteur, le but de ce projet est d'obtenir à la fois la structure de la capsid du CaMV et du complexe transmissible à une meilleure résolution que les précédents travaux, et également la structure de la protéine de transmission P2. La stratégie étant une nouvelle fois d'allier cryo-ME, pour la capsid du virus, et cristallographie aux rayons X, pour la protéine P2. La structure de la capsid a ainsi été étudiée de la même manière que les virus ArMV et BBSV présentés dans les chapitres précédents, à partir du virus extrait directement de plantes infectées. Quant à la protéine P2, une étude de la protéine complète n'est pas possible par cristallographie aux rayons X, car lorsqu'elle est surexprimée, P2 s'auto assemble pour former un réseau cristallin, appelé paracristaux. Que cela soit dans des cellules d'insectes Sf9 ou de plantes, ces paracristaux se forment et décoorent alors les microtubules de la cellule infectée pour former un complexe stable (**Figure 79**).



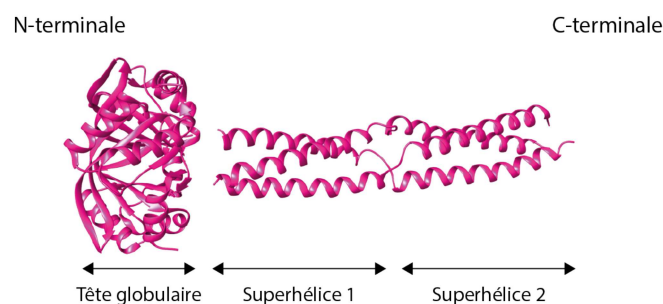
**Figure 79 : Observation des paracristaux de P2 dans des cellules Sf9 par cryo-ME en coloration négative.**  
Reproduit de (Blanc et al. 1993)

De plus, un mutant dérivé de P2 (appelé 157m), a été surexprimé dans les mêmes conditions. La substitution de la leucine 157 par une asparagine a suffi à rendre ce mutant biologiquement inactif, incapable de se lier au complexe transmissible, et incapable de former des paracristsaux. Néanmoins, ce complexe décore et stabilise également les microtubules (**Figure 80**).



**Figure 80 : Observation du mutant 157m décorant les microtubules de cellules *Sf9* par cryo-ME en coloration négative.** Les petites flèches noires indiquent les microtubules nus, les grosses flèches noires indiquent les microtubules décorés par 157m. Reproduit de (Blanc et al. 1996)

Une approche intégrée entre cristallographie aux rayons X et cryo-ME est alors utilisée pour caractériser la structure de la protéine P2. Les prédictions de structures de la protéine P2 ont montré la présence de deux domaines superhélice en C-terminale qui s'organiseraient en trimère. Une tête globulaire, correspondant aux 100 résidus présents avant les domaines superhélicés, a été modélisée arbitrairement (**Figure 81**).



**Figure 81 : Prédiction de structure de la protéine P2.**

Dans un premier temps, des mutants de la protéine P2 sont surexprimés en bactéries *E. coli*, afin d'obtenir un pseudo-modèle atomique de la protéine par cristallographie aux rayons X. Une seconde

approche par analyse de particules uniques en cryo-ME est utilisée pour étudier le complexe 157m: microtubule, dont la finalité est d'obtenir une enveloppe de la protéine P2, qui en combinaison avec la cristallographie permettra de déterminer les résidus impliqués dans l'interaction avec le microtubule, voire même d'obtenir directement un modèle atomique grâce aux progrès de la cryo-ME.

## **2. Matériel et Méthodes**

### **2.1. Étude de la protéine de capsid du CaMV**

#### **2.1.1. Purification du virus du CaMV à partir de plantes infectées**

Le virus CaMV a été purifié à partir de 100 g de feuilles de navet infectées. Les feuilles ont été broyées dans 0,5 M de tampon phosphate à pH 7,2 avec 0,75% de sulfite de sodium. La solution a ensuite été filtrée et agitée avec 1 M d'urée et 2,5 % de triton X100 pendant 24 h à 4°C. Le filtrat a ensuite été clarifié par centrifugation, et les virions ont été précipités par ultracentrifugation, puis resuspendus dans de l'eau stérile avant une nouvelle clarification. Deux gradients successifs de saccharose (10 à 40%, puis 40% à 80%) ont permis de sédimenter la bande virale, auxquels du butanol a été ajouté pour éliminer les protéines P3 et P2 qui pourraient encore être présentes sur le virus. Les virions sont une nouvelle fois précipités par centrifugation et resuspendus dans un tampon Tris 20 mM à pH 6,8 avec un détergent ionique n-octyl  $\beta$ -D-glucopyranose. La solution virale est enfin agitée à 4°C pendant plusieurs jours pour que le maximum de particule soit resuspendu. La concentration finale obtenue était de 5 mg/ml.

#### **2.1.2. Préparation des grilles**

Des grilles Quantifoil R2/2 ont été recouvertes d'un film de carbone afin d'augmenter la concentration du virus dans les trous. Ce film d'abord préparé en évaporant du carbone (Cressington Carbon Coater 208 Carbon, Ted Pella, Inc.) sur une surface de mica fraîchement clivée. Le film de carbone obtenu est d'une épaisseur d'environ 10 à 12 nm. Ce dernier est délicatement décollé de la surface du mica à l'interface air-eau, puis déposé sur les grilles Quantifoil R2/2.

### 2.1.3. Congélation de l'échantillon

L'échantillon a été congelé sur des grilles Quantifoil R2/2 (*Quantifoil Instruments GmbH*), avec le même protocole que décrit dans le sous-chapitre 2.3.2 du chapitre 2.

### 2.1.4. Acquisition des films

Les films ont été collectés sur un microscope Polara (FEI) à 300 kV (Institut de Biologie Structurale à Grenoble), équipés d'une caméra à détection directe K2 Summit (Gatan). 178 films ont été collectés à un grossissement x 20000 et une taille de pixel de 0,97 Å, en mode super-résolution. Durant chaque acquisition, 40 vues sont collectées pendant 4 secondes avec une dose de  $1e^{-}/\text{Å}^2$  par vue, empilées dans un fichier. En parallèle, ces 40 vues sont également alignées et sommées avec un logiciel propriétaire du constructeur Gatan, la somme est ensuite enregistrée dans un deuxième fichier.

### 2.1.5. Traitement d'image

À partir des 178 films enregistrés, 15037 coordonnées de particules avec une taille de boîte de 680 pixels ont été enregistrées avec le programme e2boxer (EMAN2). La fonction de transfert de contraste a été corrigée par le logiciel Ctfind4 par l'intermédiaire de l'interface graphique du programme Relion. Les 15037 particules ont été extraites par Relion à partir de la micrographie enregistrée pendant 4 secondes, et compressées par 2 pour réduire le temps de calcul. Les particules extraites ont donc une nouvelle taille de boîte de 340 pixels et une nouvelle taille du pixel de 1,94 Å. Une classification 2D a permis de sélectionner un nouveau jeu de 14170 particules qui a permis d'obtenir une première carte de densité.

## 2.2. Étude de la protéine P2 exprimée dans *E. coli*

Trois mutants de la souche Cabb-JI de la protéine P2 ont été conçus. Chaque clone comprend la partie globulaire en N-terminale, et une coupure située à différentes positions au niveau des domaines superhélices. Les tests d'expression et de solubilité se font faits avec les mêmes paramètres que décrits dans le Chapitre 2, c'est-à-dire 4 différents plasmides avec différentes

protéines porteuses (MBP, TRX, NUS, GST), 5 souches bactériennes (BL21, BL21-Rosetta, BL21-Star, BL21-C41, et Origami), 2 milieux de culture (LB et ZYM) et différentes températures de culture.

## **2.3. Étude du complexe 157m:microtubules**

### **2.3.1. Purification du complexe à partir de cellules d'insectes *Sf9* infectées par des baculovirus**

Le système d'expression baculovirus AcSLP10 a été utilisé pour exprimer le gène du mutant 157m dans les cellules *Sf9*. Deux jours après infection, Taxol est ajouté à culture cellulaire à une concentration finale à 2  $\mu$ M, la boîte est ensuite incubée 3 jours supplémentaires à 29°C. La culture est ensuite centrifugée, et le culot est resuspendu dans un tampon 100 mM Pipes, 1 mM EGTA, 1 mM MgCl<sub>2</sub>, 10  $\mu$ M de Taxol à pH 6,6. La solution est ensuite agitée 30 secondes avec un vortex et centrifugée à 800 g à température ambiante pendant 1 minute. 20  $\mu$ l du surnageant est ensuite dilué avec 20  $\mu$ l de Tween 0,1 %, la préparation est prête à être congelée.

### **2.3.2. Congélation de l'échantillon**

3  $\mu$ l de l'échantillon ont été déposés sur une grille Quantifoil R2/2, préalablement recouverte d'un film de carbone et rendue hydrophile. La grille a été placée sur la guillotine CP3 de Gatan, l'excès de solution a été retiré après application de papiers filtres des deux côtés de la grille pendant deux secondes, dans une chambre avec une humidité relative de 99%. La grille a ensuite été plongée dans de l'éthane liquide refroidi par de l'azote liquide.

### **2.3.3. Acquisition des micrographies**

Les micrographies ont été collectées sur un microscope JEOL 2200FS à 200 kV (Centre de Biochimie Structurale), équipé d'un filtre d'énergie oméga et d'une caméra CCD 4x4k (Gatan). 75 micrographies ont été collectées à un grossissement x 50000 avec une taille de pixel de 2,3 Å.

### 2.3.4. Traitement d'image

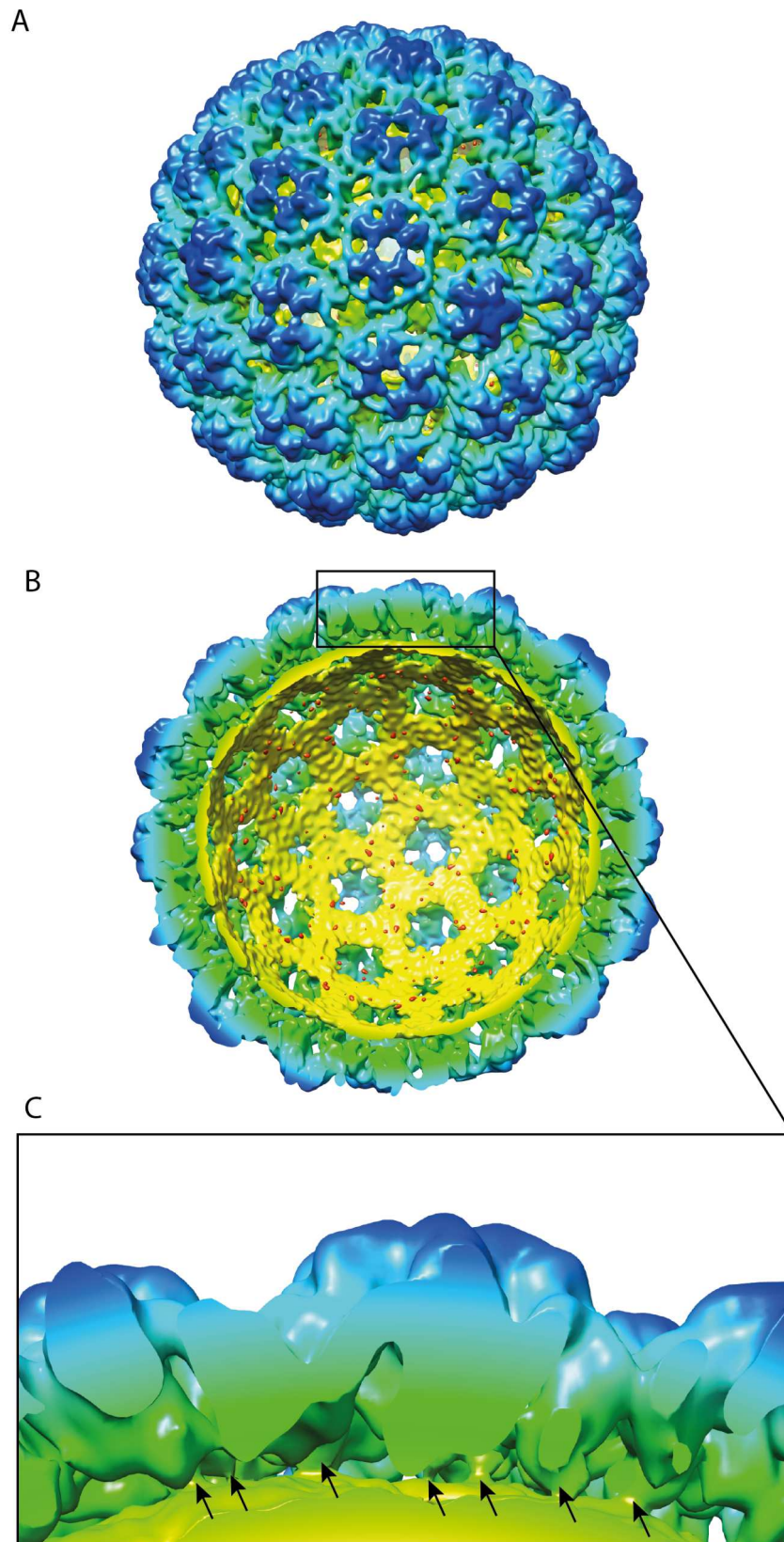
Les micrographies ont été triées avec le programme `e2projectmanager.py`, puis l'estimation des paramètres FTC des a été faite avec `Ctffind 3` (Mindell & Grigorieff 2003). À partir des images corrigées par la FTC, les vignettes des particules ont été extraites avec le programme `helixboxer` (EMAN2). Les vignettes correspondent chacune à une boîte contenant une partie du microtubule, elles sont répétées le long du microtubule, et chaque vignette chevauche à 98% la vignette précédente. Pendant l'extraction, les vignettes ont été compressées par 2 et un filtre passe-bas de 20 Å a été appliqué. Un total de 37 118 vignettes de dimensions 200 x 200 pixels a pu être extrait. La classification 2D des particules s'est faite avec le programme `Imagic` (van Heel & Keegstra 1981), qui permet, au travers d'une approche appelée analyse multi statistique (Multi-statistical alignment ou MSA), de réunir des images similaires en classes. Il est ensuite possible de choisir certaines classes comme références, et d'aligner les particules contre ces références, au travers de l'approche d'alignement contre de multiples références (Multi-Reference Alignment ou MRA). La carte de densité a été reconstruite avec le logiciel de reconstruction hélicoïdale `IHRSR++`, basé sur le logiciel `IHRSR` (Egelman 2007).

## 3. Résultats

### 3.1. Étude de la protéine de capsid du CaMV

Une reconstruction de 14170 particules extrait des films entiers, a permis de reconstruire une carte de densité de la capsid à une résolution de 9,2 Å, auquel un aiguisement gaussien d'une valeur de -200 Å<sup>2</sup> a été appliqué (**Figure 82**).

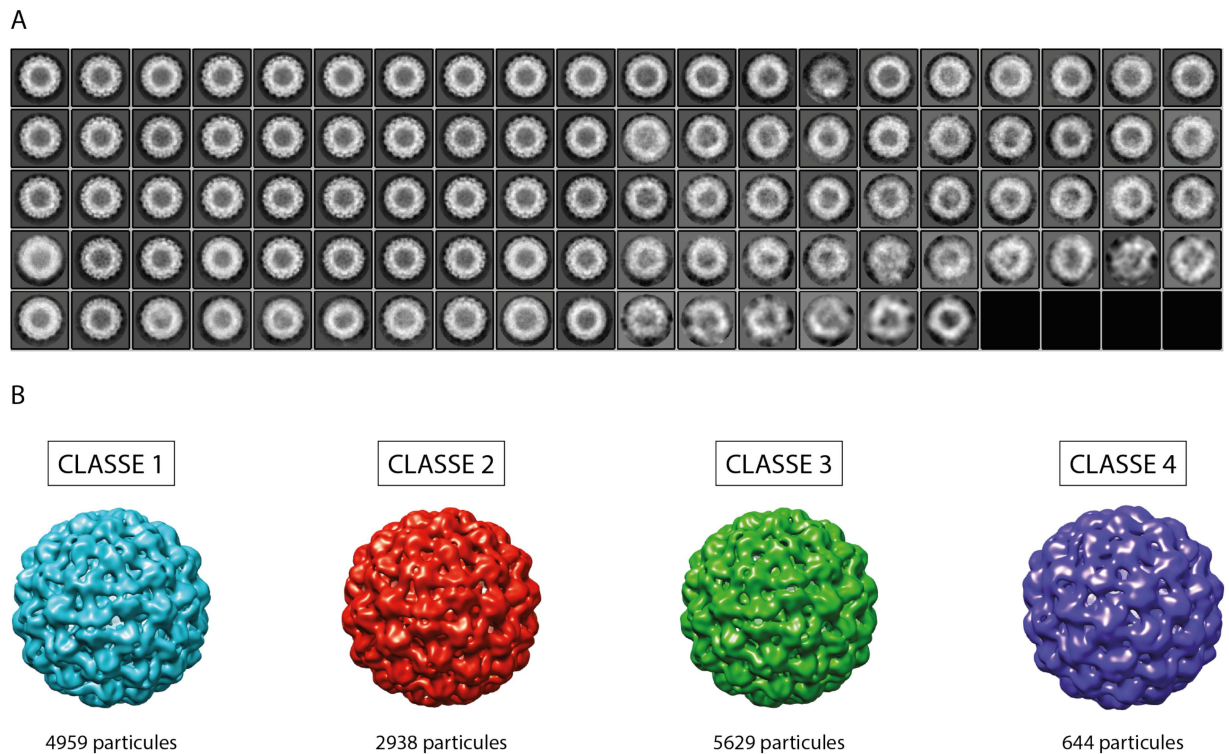




**Figure 82 : Reconstruction 3D de la capsid du CaMV. (A) Structure externe de la capsid du CaMV. (B) Coupe interne de la capsid du CaMV. (C) Zoom sur les interactions entre les protéines de la capsid et la couche externe du génome du CaMV.**



Il semble que la capside du CaMV possède un certain degré de flexibilité qui avait déjà limité la résolution de la précédente reconstruction publiée dans (Plisson et al. 2005). La présence de particules de différentes tailles était supposée, mais les classifications 2D et 3D n'ont pas montré de différences visibles (**Figure 83**).



**Figure 83 : Classification 2D et 3D de la capside du CaMV.** (A) Classes 2D obtenues après 30 cycles itératifs. (B) Classes 3D obtenue après 25 cycles itératifs.

Cette reconstruction n'intègre pas l'utilisation des vues, et aucune reconstruction à partir d'un sous-jeu issu de la classification 3D n'a été réalisée. Cependant, il reste difficile d'envisager combler un écart de 5 Å pour pouvoir accéder aux chaînes latérales avec seulement ces approches. La caractérisation du virus à haute résolution devient plus difficile que les études des virus ArMV et BBSV et une investigation de la flexibilité locale sera prochainement réalisée au travers de nouvelles approches permettant de concentrer l'alignement des particules sur une seule région. En réalisant des reconstructions localisées, c'est-à-dire en se concentrant seulement sur un axe de symétrie du virus à la fois, il est possible d'aligner les particules seulement sur cette région sans prendre en compte ce qu'il existe en dehors. Cette nouvelle approche devrait permettre de reconstruire séparément les différentes régions de la capside à plus haute résolution.

## 3.2. Étude de la protéine P2 exprimée dans *E.coli*

Les mutants de P2 ont été conçus de manière à conserver la partie globulaire en N-terminale de la protéine, et ajouter différentes longueurs de résidus en C-terminale comme décrit ci-dessous :

Séquence de la protéine P2 entière de la souche Cabb-J1 :

MSITGQPHVYKKDTIIRLKPLSLNSNNRSYVFSSSKGNIQNIINHLNNLNEIVGRSLLGIWKINSYFGLSKDPSESCKSNPSVF  
NTAKTIFKSGGVDYSSQLKEIKSLLESQNTRIKSLEKAIQSLDDKIEPEPLTKEEVKELKESINSIKEGLKNIIG

Mutant P2A – Partie N-terminale et une partie de la tête globulaire (8 kDa) :

MSITGQPHVYKKDTIIRLKPLSLNSNNRSYVFSSSKGNIQNIINHLNNLNEIVGRSLLGIWKINSYFGLSKDPSESCKSN  
PSVFNTAKTIFKSGGVDYSSQLKEIKSLLESQNTRIKSLEKAIQSLDDKIEPEPLTKEEVKELKESINSIKEGLKNIIG

Mutant P2B – Partie N-terminale jusqu'à la fin de la tête globulaire (11 kDa) :

MSITGQPHVYKKDTIIRLKPLSLNSNNRSYVFSSSKGNIQNIINHLNNLNEIVGRSLLGIWKINSYFGLSKDPSESCKSN  
PSVFNTAKTIFKSGGVDYSSQLKEIKSLLESQNTRIKSLEKAIQSLDDKIEPEPLTKEEVKELKESINSIKEGLKNIIG

Mutant P2C – Partie N-terminale jusqu'à la fin de la première superhélice (14,5 kDa) :

MSITGQPHVYKKDTIIRLKPLSLNSNNRSYVFSSSKGNIQNIINHLNNLNEIVGRSLLGIWKINSYFGLSKDPSESCKSN  
PSVFNTAKTIFKSGGVDYSSQLKEIKSLLESQNTRIKSLEKAIQSLDDKIEPEPLTKEEVKELKESINSIKEGLKNIIG

Les meilleurs résultats de purification de la protéine P2 ont été obtenus avec le mutant P2C fusionné à la protéine porteuse MBP, transformé dans la souche BL21 mise en culture dans le milieu d'auto-induction ZYM à une température de 37°C pendant 8 heures, puis à 30°C pendant une nuit. Les bactéries ont été lysées dans un tampon 50 mM MES, 500 mM NaCl, 2 mM  $\beta$ -mercaptoéthanol, 1 mg/ml lysozyme, 1 mg/ml DNase. La protéine P2C a été purifiée par chromatographie d'affinité sur colonne de nickel (**Figure 84A**), puis séparée de la protéine porteuse MBP après coupure avec la protéase TEV pendant 3 h à température ambiante, suivie d'une séparation des protéines sur billes de nickel pour enlever les protéines MBP et TEV qui ont toutes deux une étiquette histidine, la protéine P2C n'est donc plus retenue par les billes (**Figure 84B**). Une chromatographie d'exclusion par la taille a permis d'enlever les protéines MBP et TEV résiduelles issues de la séparation sur billes de nickel, et de récupérer un volume de 6 ml de la protéine P2C (**Figure 84C**). Enfin, 2 fractions de 1 ml ont été utilisées pour des tests de concentration par centrifugation sur deux types de membranes hydrophiles : une membrane de polyethersulfone (PES) et une membrane de cellulose (Hydrosart®). La concentration d'un volume de 1 ml vers un volume de 200  $\mu$ l sur chacune des deux membranes

s'est ainsi montrée concluante. (**Figure 84D**). La membrane PES a été choisie pour concentrer le total de volume restant de protéines (4 ml), car la concentration s'était révélée être plus rapide sur cette membrane. Cependant, la concentration d'un plus gros volume semble entraîner une perte de la protéine P2C, puisqu'après être passée d'un volume de 4 ml vers un volume de 2ml, la quantité de protéine diminue (**Figure 84E**).

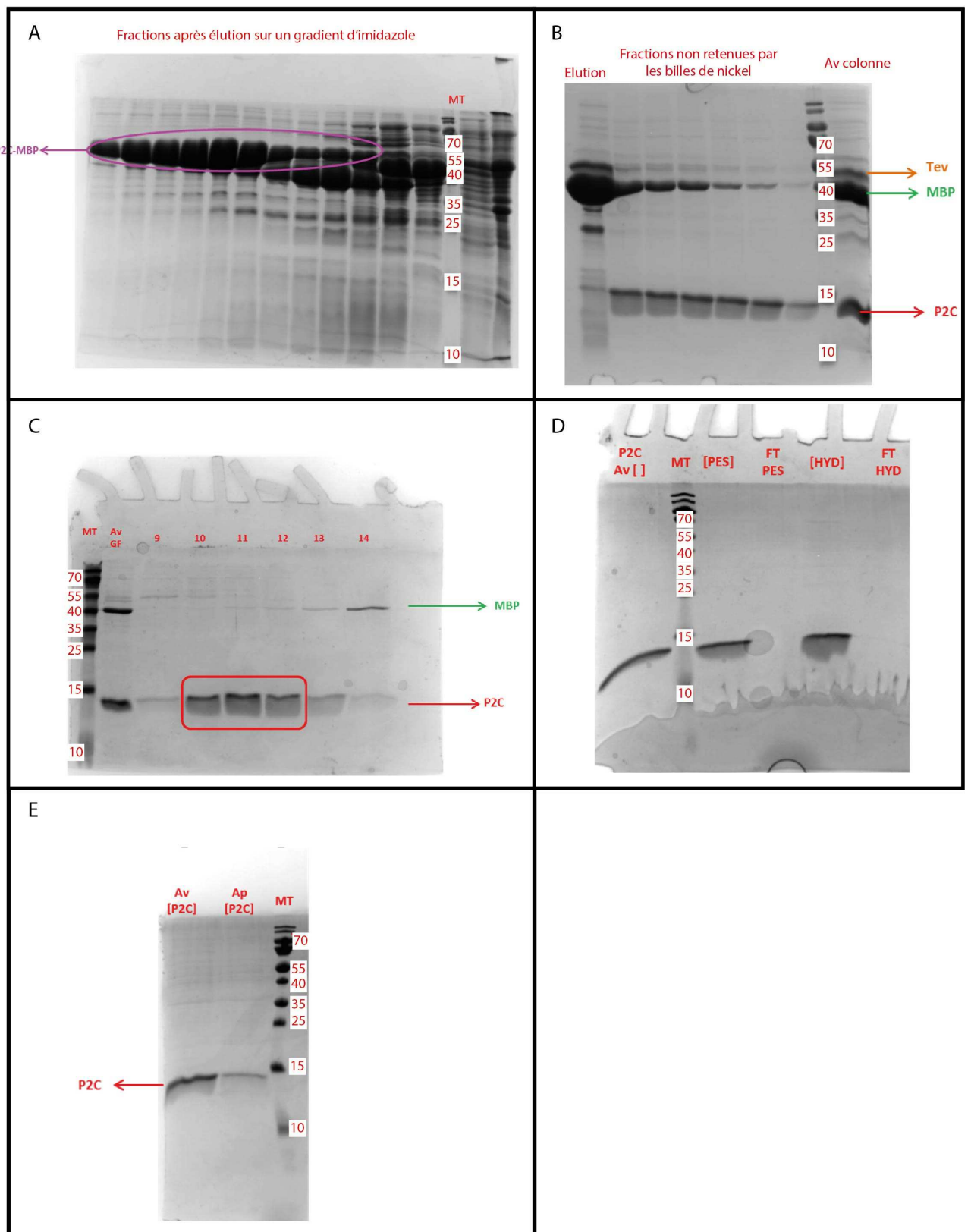


Figure 84 : Gels de polyacrylamide des étapes de purification de la protéine mutante P2C. Av=avant, Ap=après. (A) Fractions (1 ml) obtenues après chromatographie sur colonne histidine (MT = Marqueur des

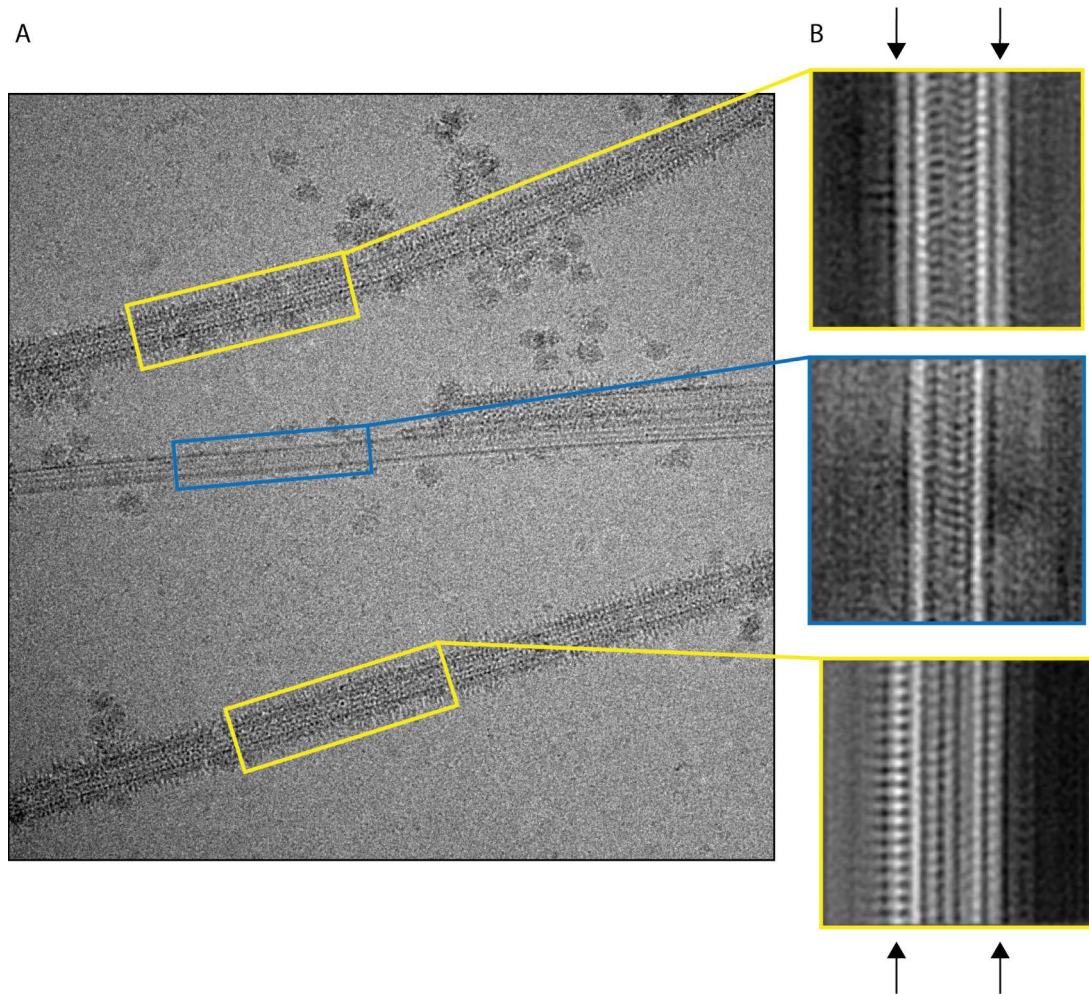
tailles moléculaires, MBP-P2C = 56 kDa). **(B)** Fractions (1 ml) obtenues après coupure avec la protéase TEV et chromatographie sur billes de nickel (TEV = 52 kDa, MBP = 42 kDa). **(C)** Fractions (2 ml) obtenues après chromatographie d'exclusion la taille. **(D)** Fractions obtenues après concentration d'un échantillon de 1 ml sur membranes PES et Hydrosart®, vers un volume de 200 µl. **(E)** Fraction obtenue après concentration du volume restant (4 ml) sur membrane PES vers un volume de 2 ml.

Aucun précipité ne s'est formé pendant la concentration de P2, il se pourrait qu'une augmentation de la quantité de la protéine P2 favorise les interactions au niveau des domaines superhélices qui se colle à la membrane PES. Cette membrane avait été choisie arbitrairement après les tests de concentration et il reste à purifier une nouvelle fois la protéine pour un test de concentration sur une membrane Hydrosart®. Même si protéine P2 n'est pas encore purifiée à une concentration suffisante pour des tests de cristallisation, les résultats restent très encourageants puisque c'est la première fois que la protéine est obtenue sous forme soluble à partir de bactéries *E. coli*.

### 3.3. Étude du complexe 157m:microtubules

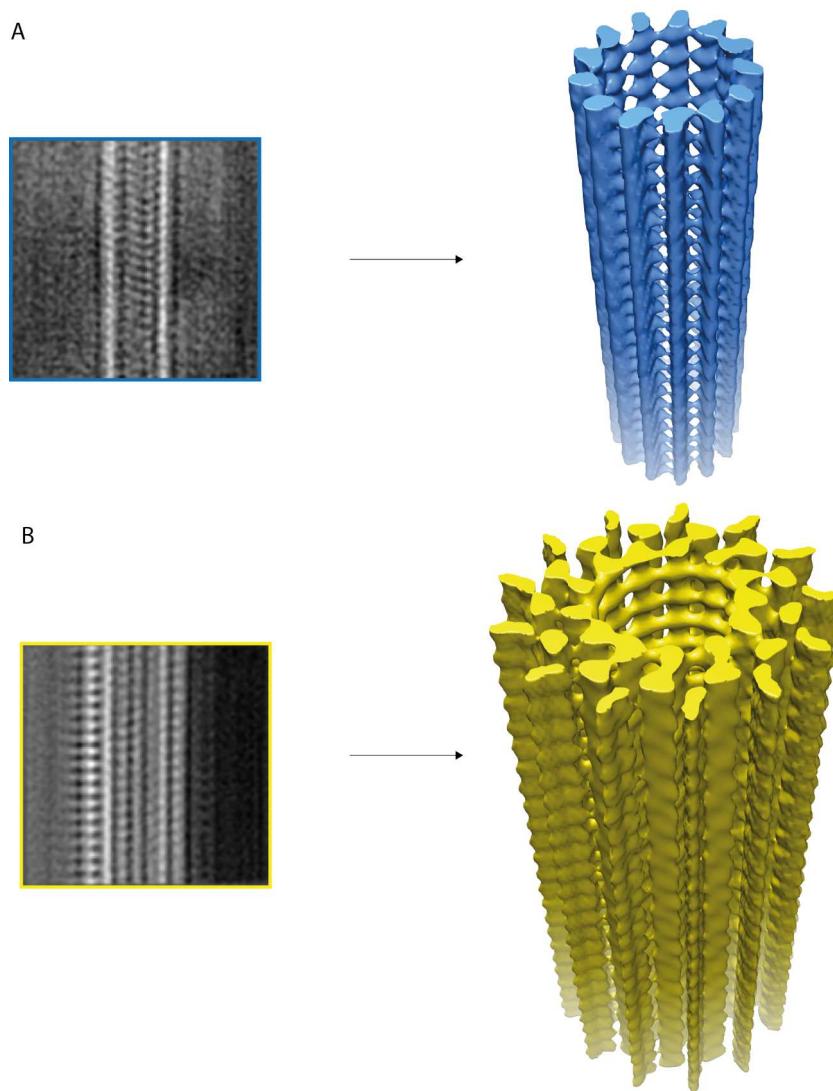
Les micrographies du complexe étaient composées à la fois des microtubules décorés, discernables par les stries formées autour des microtubules, mais également des microtubules partiellement décorés, qui ne montre aucune strie (**Figure 85A**). Des vignettes de microtubules décorés et nus ont été extraites et classifiées. L'alignement des vignettes s'est uniquement fait au niveau du centre de chaque vignette, en utilisant un masque formant un rectangle horizontal de dimensions 100 x 50 pixels. La classification a permis de différencier les parties de microtubules nus et de microtubules décorés. La distribution du mutant 157m sur les microtubules a pu être observée par comparaison entre les deux types de classes. La partie globulaire du mutant devient visible autour des microtubules, et nous pouvons discerner une traînée correspondant aux domaines superhélices (**Figure 85B**). La régularité de la décoration autour des microtubules indiquant que le mutant semble interagir avec le microtubule en suivant son hélicité.





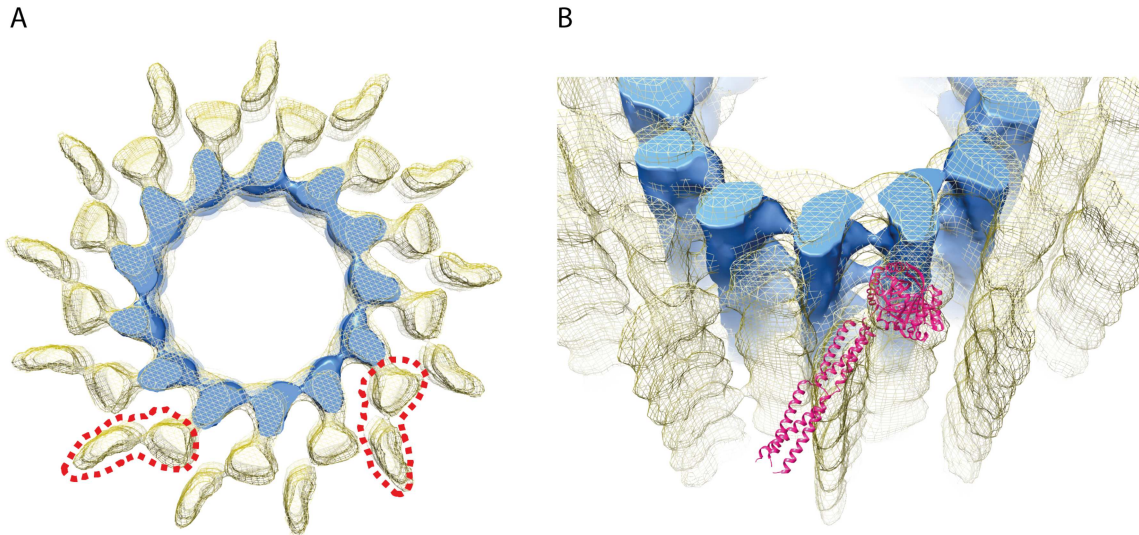
**Figure 85 : Classification 2D du complexe 157m:microtubule par Imagic.** (A) Micrographie de l'échantillon, comportant des microtubules décorés par le mutant 157m (jaune) et des portions de microtubules non décorés (bleu). (B) Exemples de classes obtenues après extraction des vignettes. Deux types de classes ont été obtenues, le microtubule nu (bleu) ou le microtubule décoré (jaune) ou les densités correspondant au mutant sont indiquées par des flèches noires.

Après la classification, les 975 vignettes correspondantes au microtubule nu ont été retenues, ainsi que 17656 vignettes provenant des classes du microtubule décoré montrant le meilleur détail. Les paramètres de symétrie des microtubules *in vivo*, c'est-à-dire le microtubule à 13 protofilaments, ont été utilisés pour reconstruire les microtubules nus et décorés : C13 (axe de rotation 13),  $\Delta Z = 9,2 \text{ \AA}$  (distance entre deux sous-unités successives),  $\Delta\phi = 27,69^\circ$  (angle de rotation entre deux sous-unités successives) (**Figure 86**). Par ailleurs, en utilisant des paramètres de symétrie différents, puis en affinant ces paramètres, les valeurs finales convergent vers les paramètres d'hélicité des microtubules.



**Figure 86 : Reconstruction 3D du complexe 157m:microtubule. (A)** Reconstruction 3D à partir des vignettes de microtubules nus. **(B)** Reconstruction 3D à partir des vignettes de microtubules décoré par le mutant.

La carte de densité du complexe semble montrer une décoration du mutant 157m qui interagit avec un ratio 1:1 avec la sous-unité du microtubule, sans pouvoir discerner s'il s'agit de l'alpha ou la beta tubuline cependant. Une carte de différence entre ces deux reconstructions a ensuite permis de délimiter les densités correspondantes au mutant 157m. Nous pouvons alors distinguer la partie globulaire du mutant qui interagit avec le microtubule (**Figure 87A**). En recalant la prédiction de structure de la protéine P2 dans la carte, nous pouvons voir que seul le premier domaine superhélice est visible, le deuxième probablement absent à cause de sa flexibilité (**Figure 87B**).



**Figure 87 :** Carte de différence entre le microtubule décoré et le microtubule nu. (A) Vue du dessus de la carte de différence. La contribution du microtubule nu est en bleu, et le complexe est en jaune. Le mutant 157m est ainsi délimité par des pointillés rouges. (B) Vue en perspective de la carte de différence avec recalage de la prédiction de structure de P2.

La taille des densités observées autour des microtubules semble cohérente avec une association de la protéine en trimère comme prédit par la prédiction des domaines superhélices. Ces premiers résultats montrent qu'une approche par cryo-ME est possible pour déterminer l'organisation structurale de P2 à travers l'étude du complexe 157m: microtubule, combinée avec les dernières avancées de la microscopie électronique, il est maintenant envisageable d'amener ce projet à plus haute résolution.

## 4. Conclusion

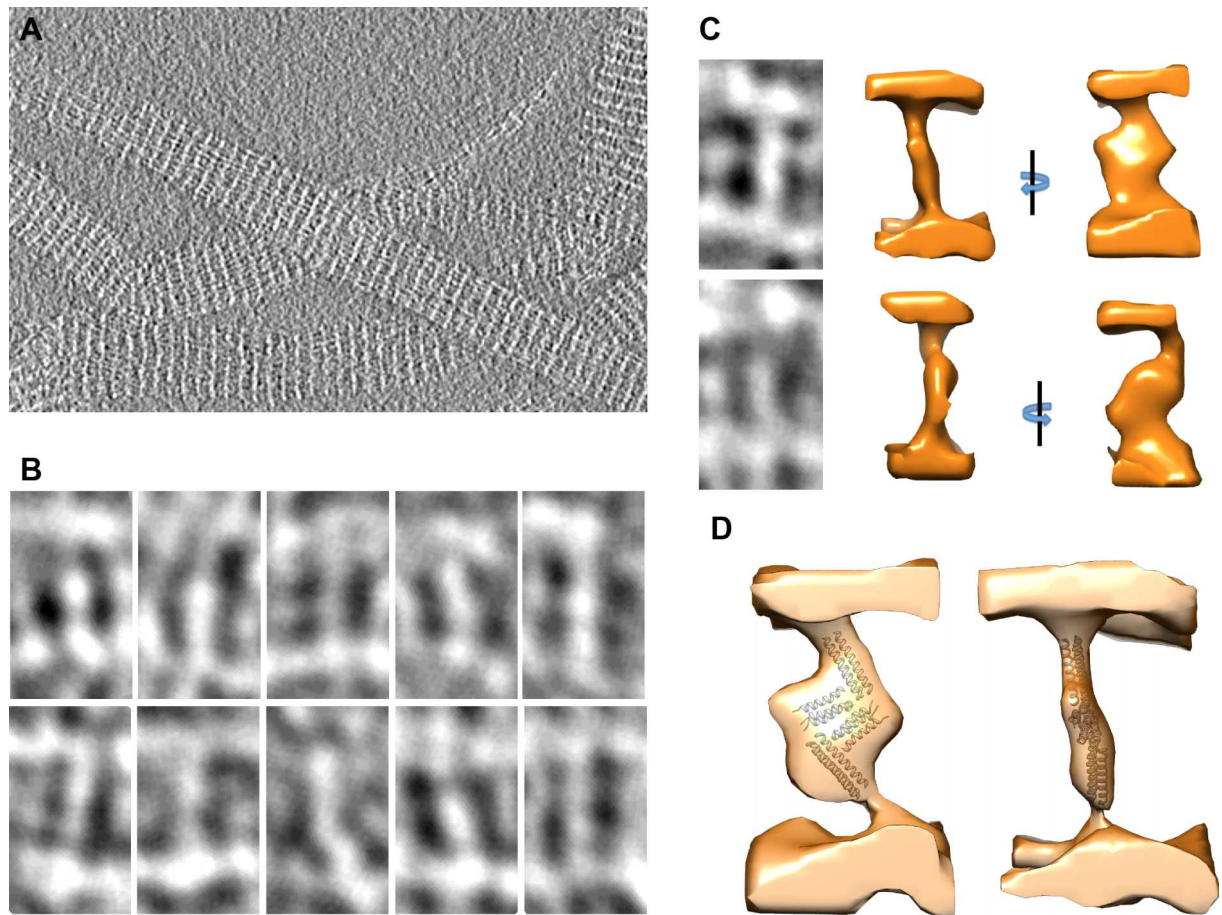
Les trois projets concernant les études structurales de la capsid du CaMV et de sa protéine de transmission sont toujours en cours. La purification de la protéine P2 à partir de bactérie *E.coli* avance à grands pas. Un temps important avait été nécessaire pour la partie biologie moléculaire, avec l'insertion des séquences des protéines dans les différents plasmides et la transformation dans les souches bactériennes. Une combinaison plasmide/bactérie/température de culture a permis de purifier la protéine de fusion MBP-P2C, qui a ensuite été séparée pour obtenir la protéine virale seule. Maintenant, seule l'étape de concentration de la protéine reste à être optimisée en essayant



d'autres types de membranes pour concentrer la protéine virale et les essais de cristallisation pourront débiter.

La reconstruction de la capsid du CaMV n'a été réalisée qu'avec les films entiers compressés par deux, et les vues n'ont encore pas été incorporés. Cependant, la résolution atteinte reste encore loin de la fréquence Nyquist qui est de 3,88 Å pour les images compressées deux fois. À ce niveau de la reconstruction, même l'utilisation des vues ne permettra pas de combler les 5-6 Å manquant pour accéder aux chaînes latérales de la protéine de capsid. Il existe maintenant des approches permettant de travailler localement sur différentes régions de la protéine d'intérêt pour essayer d'améliorer la résolution. Ces méthodes sont envisagées pour poursuivre la caractérisation de la capsid virale du CaMV et plus tard du complexe transmissible. Deux méthodes sont aujourd'hui décrites dans la littérature, la méthode de classification concentrée (X. C. Bai et al. 2015), et la reconstruction localisée (Ilca et al. 2015). Le principe des deux méthodes est similaire, un masque binaire est d'appliquer autour de la région d'intérêt, l'affinement se fera alors par rapport au signal à l'intérieur de ce masque, en ignorant les mouvements qui peuvent se faire en dehors du masque. Les deux approches sont intégrées au logiciel Relion, la différence entre celles-ci est que la première se concentre sur une seule région, alors que la deuxième permet de prendre en compte la symétrie dans le virus. De cette manière, les deux approches sont envisagées pour reconstruire séparément chaque axe de symétrie pour tenter d'améliorer la résolution de la reconstruction. Si cela ne suffit pas, des sessions d'acquisitions automatiques sur un microscope Titan Krios seraient envisageables pour enregistrer un nombre plus important de particules virales, qui pourrait permettre de classifier plus efficacement les particules, et générer des lots plus homogènes.

L'étude du complexe 157m: microtubule a permis de valider l'approche par cryo-ME comme voie possible dans la caractérisation de la structure de la protéine P2. L'utilisation des paramètres d'hélicité des microtubules *in vivo* a permis de reconstruire les cartes de densité correspondant au microtubule nu et au microtubule décoré par le mutant. Par ailleurs, la décoration du mutant semble régulière le long du microtubule et les densités compatibles avec les dimensions d'un trimère. Cette information a été validée par une étude des paracristsaux de P2 par cryo-tomographie électronique, étude réalisée par le Dr Patrick Bron avec l'étudiant en Master 2 Julien Bous, qui montre des interactions en différentes protéines P2 également compatibles avec une forme trimère de la protéine (**Figure 88**).



**Figure 88 : Étude des paracristaux de P2 par cryo-tomographie électronique.** (A) Image de paracristaux x 100000. (B) Différentes projections de tomogrammes montrant la connexion des protéines P2. (C) Reconstruction 3D de deux connexions P2:P2. (D) Recalage de la prédiction de structure de P2 (trimère) dans les densités des paracristaux.

Cette étude a mis en avant les interactions P2:P2 formant les paracristaux, leur flexibilité empêche cependant d'améliorer la résolution par une approche de moyennage des sous-volumes. C'est pourquoi l'étude du complexe 157m: microtubule semble être l'approche la plus adaptée pour obtenir des données hautes résolutions par cryo-ME de la protéine P2. De nouvelles images du complexe ont ainsi été enregistrées sur un microscope Polara (IBS, Grenoble) afin de recommencer le traitement d'image avec des meilleures données et de se diriger vers la haute résolution.

# CONCLUSION GÉNÉRALE

# Conclusion générale

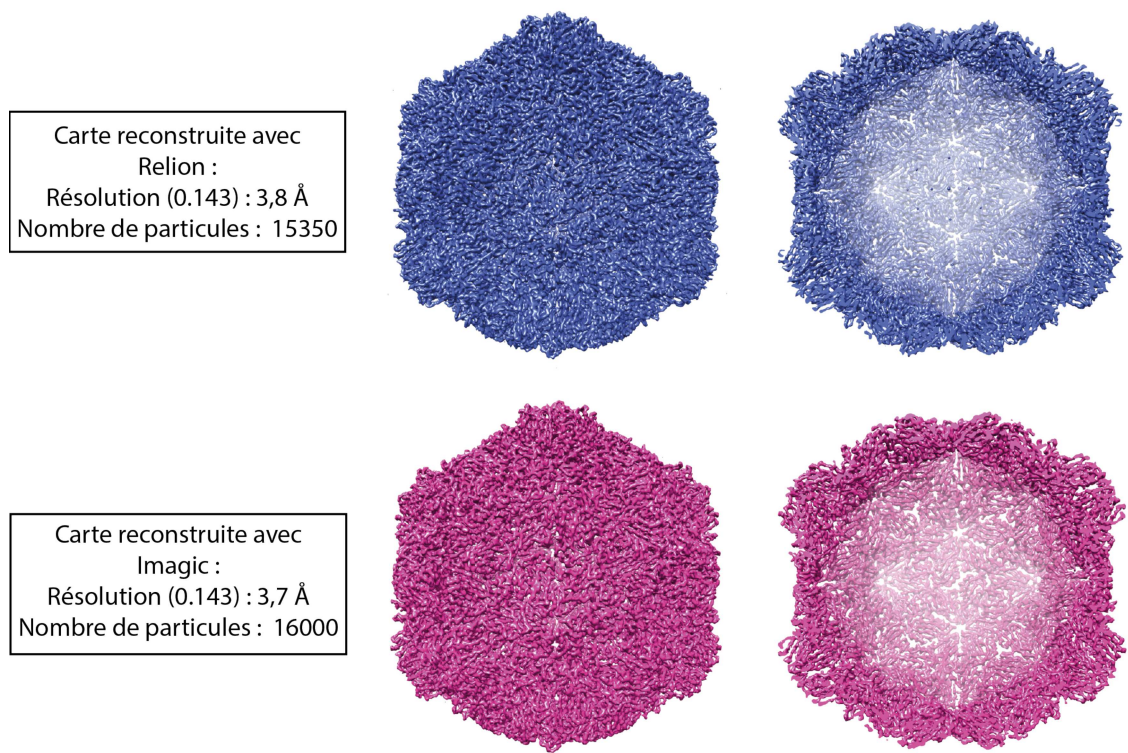
---

La microscopie électronique en tant que technique de biologie structurale est arrivée à maturité. Il est maintenant possible d'obtenir des informations structurales de protéines ou de complexes biologiques à un niveau atomique. Cette révolution est liée à des évolutions technologiques importantes, tant au niveau des microscopes, que des caméras ou des logiciels de traitement d'image. Grâce à l'arrivée de microscopes électroniques plus stables mécaniquement et électroniquement, il est possible d'enregistrer plusieurs milliers d'images de façon automatique. Les nouvelles caméras, dites à détection directe d'électrons ont une sensibilité 10 fois supérieures à celle des anciennes caméras, ce qui permet d'obtenir des niveaux de détail dans les images jusqu'alors impossibles à visualiser auparavant. De plus, il ne s'agit plus d'images au sens strict, mais de films. En effet, ces caméras enregistrent, non pas une image, mais une série d'images au cours de l'exposition. L'analyse de ces séries d'images a permis de montrer que, bien que piégées dans une mince pellicule de glace, les particules protéiques pouvaient bouger sous l'effet du faisceau d'électrons, et donc que ces mouvements pouvaient être corrigés afin d'améliorer la qualité finale de la reconstruction tridimensionnelle de la protéine ou du complexe protéique étudié. De même, l'apparition de programmes dédiés au traitement des images « plus abordables pour les utilisateurs » est une des clés qui a permis d'ouvrir la cryo-ME au plus grand nombre.

Au cours de mes travaux de thèse, j'ai exploré beaucoup d'aspects du traitement d'image notamment afin d'optimiser la résolution de la reconstruction finale de l'ArMV. Ce virus fait partie de la famille des *Nepovirus* et est impliqué dans la maladie du court-noué de la vigne. Cette maladie est liée à la transmission par nématodes, de deux virus très proches, l'ArMV et le GFLV. Or il existe une spécificité de transmission de ces virus par les nématodes. Ne disposant d'aucune information sur le récepteur des nématodes permettant la fixation des virus, la comparaison de la structure atomique des capsides de ces deux virus est une stratégie visant à identifier au niveau moléculaire, les motifs structuraux impliqués dans la sélectivité de transmission de ces deux virus. LE GFLV a été cristallisé et sa structure déterminée par cristallographie aux rayons X. Le traitement des images de l'ArMV enregistrées sur un microscope Titan Krios (IGBMC, Strasbourg) équipé d'une caméra à détection directe d'électrons de type Facon II, m'a permis de calculer une carte de densité de l'ArMV à 3.8 Å de résolution (FSC 0,143). Nous avons alors défini une stratégie basée sur des programmes dédiés à la cristallographie pour calculer une structure atomique de l'ArMV. Le traitement d'image a été réalisé en combinant plusieurs programmes. La partie principale concernant l'affinement de la structure a

été réalisée à l'aide du package Relion Version 1.4. La particularité de ce programme, et qui a fait son succès est son côté « automatique ». Contrairement aux autres suites de logiciels où l'intervention de l'utilisateur est importante, Dr Scheres a développé Relion dans le but qu'il puisse être utilisé par le plus grand nombre, donc par des personnes n'ayant pas des connaissances poussées en traitement d'image. Ceci implique que le programme optimise lui-même les paramètres afin d'améliorer l'affinement de la structure. Dans le monde de la microscopie électronique, l'arrivée de Relion a été vécue comme une véritable révolution, et par conséquent n'a pas été vu d'un bon œil par les concepteurs des autres logiciels de traitement d'image comme Imagic (van Heel & Keegstra 1981), Xmipp (Sorzano et al. 2004), Eman2 (Tang et al. 2007) ou Frealign (Grigorieff 2007). Il est souvent reproché à Relion de ne pas tenir suffisamment compte de certaines corrections ou que l'approche bayésienne induise des pertes d'information à des fréquences données. De même Relion est souvent comparé à une boîte noire alors que les autres logiciels restent plus associés à des boîtes à outils.

Dans ce contexte, mon équipe d'accueil a entrepris une collaboration avec le Dr Marin Van Heel (NECEN, Netherlands) qui est le concepteur du logiciel Imagic. Les images de l'ArMV ayant été faites sur le microscope Titan Krios du NECEN, cette collaboration avec le Dr Van Heel est venue naturellement en discutant du traitement haute résolution des images. Les logiciels comme Imagic ont dû adapter leur méthodologie afin d'intégrer la manière de traiter les films provenant des caméras à détection directe. L'ancien étudiant en thèse du Dr Van Heel, Dr Pavel Afanasyev, a donc traité les images de l'ArMV à l'aide d'Imagic en intégrant les derniers développements qui avaient été définis à partir du traitement d'un jeu d'images de l'hémocyanine (en soumission dans le journal *Scientific Reports*). On peut résumer les dernières améliorations d'Imagic en 4 parties : 1°) correction du bruit contenu dans les images (Afanasyev et al. 2015), 2°) alignement des images composant les films, 3°) optimisation de l'alignement des images par rapport à une référence, 4°) correction anisotropique du grandissement des images (Grant & Grigorieff 2015a). À titre de comparaison, Relion ne corrige pas le « background » et le grandissement anisotropique des images. Une autre différence concerne la partie dite de « post-traitement d'image » qui consiste à rehausser les données hautes résolutions par rapport aux données basses résolutions dans la carte de densité. Cette étape est automatique dans Relion et totalement manuelle dans Imagic. Dans la **Figure 89**, les cartes de densité de l'ArMV obtenues par Relion et par Imagic sont présentées. La conclusion de cette comparaison est que les deux cartes sont extrêmement similaires et à des résolutions équivalentes, à savoir entre 3.7 et 3.8 Å. De même, le modèle atomique calculé à partir de ces deux cartes est équivalent.



**Figure 89 : Reconstructions 3D de l'ArMV obtenues par Relion (bleu) et Imagic (rose)**

On peut donc conclure que les corrections supplémentaires présentes dans Imagic n'induisent pas d'amélioration significative jusqu'à cette résolution de 3.8 Å, et qu'entre nos mains, nous n'avons pas vu de différence structurale entre les cartes obtenues par Relion ou par Imagic. Dans ce contexte, le traitement automatique proposé par Relion répond donc parfaitement aux exigences souhaitées, à savoir obtenir une carte de densité d'un complexe protéique à une résolution quasi atomique. Ces résultats résument parfaitement la mutation qui s'est produite dans le domaine du traitement d'image en cryo-microscopie. Nous sommes passés d'un domaine d'experts à un domaine ouvert à tous. Ceci explique en grande partie pourquoi un très grand nombre de cristallographes se sont mis à la cryo-microscopie électronique pour résoudre la structure atomique de complexes biologiques d'intérêts.

Dans le contexte de ma thèse, après avoir réalisé que le virus contaminant n'était pas le CPMV, mais le BBSV, j'ai repris l'analyse de l'ensemble des images et films. J'ai ainsi pu obtenir une carte de densité du BBSV à 3.8 Å de résolution et calculer un modèle atomique de ce virus.

Comment se fait-il que malgré les classifications 3D et 2D de nos images, il n'ait pas été possible de descendre en dessous de 3.8 Å de résolution pour l'ArMV et le BBSV ?

En théorie, la taille du pixel étant de 1.13 Å, il devrait être possible de calculer des cartes de densité jusqu'à 2.3 Å. Il est évident que nous nous retrouvons face à un problème d'hétérogénéité structurale. Alors que pendant de nombreuses années, l'utilisation d'objets ordonnés (particules icosaédriques, hélices) était un gage de « haute résolution », nous sommes passés à un stade maintenant où cela pose finalement un problème. Il est en effet possible d'affiner localement la structure d'une protéine ou d'un complexe avec peu de symétrie alors que cela devient très difficile pour un objet compact et répété. Dans le cas des virus icosaédriques, ou encore des objets ayant une symétrie hélicoïdale, il faudra donc développer de nouvelles stratégies afin de pouvoir traiter plus finement l'hétérogénéité structurale si on veut obtenir des reconstructions de meilleure qualité.

C'est cette hétérogénéité structurale qui nous a limités dans le cas du CaMV. Les travaux précédents réalisés par (Plisson et al. 2005) avaient permis d'obtenir une enveloppe du CaMV autour des 20 Å de résolution. Ils avaient aussi montré que les particules virales affichaient une hétérogénéité de forme, avec des particules qui pouvaient avoir des aspects ovoïdes. Nous sommes partis du postulat qu'il devait être possible de trier les particules selon leur forme et affiner la structure du virus à partir de lots d'images homogènes de particules. Mais il fallait pour cela avoir un nombre important de particules. Nous avons donc décidé de déposer un film de carbone sur les grilles Quantifoil R2/2 afin d'augmenter artificiellement la concentration locale des virus. Dans ces conditions, il est possible d'avoir un tapis de particules de CaMV après congélation. La première information est que finalement l'aspect visuel des particules montrait moins d'hétérogénéité que ce qui avait été observé précédemment. L'analyse des images a permis d'obtenir une reconstruction du CaMV autour des 9 Å de résolution. C'est un progrès important en termes de gain de résolution, mais assez décevant par rapport à ce que nous espérions. D'autre part les classes 2D des images ne m'ont pas permis de trier plus finement les particules, ni même les classes 3D. Afin d'améliorer la résolution de la carte de densité, il sera nécessaire de faire une reconstruction localisée autour des axes de symétrie afin d'essayer de caractériser chaque région séparément. Si cela ne suffit pas, il sera nécessaire d'avoir beaucoup plus d'images initiales, ce qui implique de pouvoir faire des acquisitions automatiques de plusieurs jours sur un microscope Titan Krios. Dans un tel projet, l'idéal serait d'avoir plus d'un million de particules afin de pouvoir être extrêmement sélectif pour générer des lots de particules homogènes conséquents et ainsi être capable de calculer une carte de densité du CaMV à une résolution inférieure à 4 Å.

Dans l'étude du CaMV et de la protéine P2 impliquée dans la transmission par le puceron, nous avons montré qu'il était possible d'exploiter le fait que le mutant P2 157m pouvait décorer les microtubules, et montrer que la protéine P2 devait être associée sous forme de trimère. Les premiers



résultats sont très encourageants et les difficultés rencontrées, forçant à combiner plusieurs techniques différentes, rendent ce travail d'autant plus intéressant. Le protocole de purification de la protéine soluble demande seulement un dernier ajustement de la concentration de la protéine, avant de pouvoir commencer les premiers essais de cristallisation. Le traitement de nouvelles images du mutant 157m acquises sur un microscope Polara équipé d'une caméra de dernière génération, et le développement de programmes intégrant la reconstruction hélicoïdale devrait permettre d'obtenir une meilleure structure du complexe très prochainement.

Ce travail de thèse montre clairement le potentiel et l'apport de la cryo-microscopie électronique dans l'étude de la structure des virus de plante et de leur transmission. L'évolution de la cryo-microscopie est telle ces dernières années qu'il est possible d'obtenir des cartes de densité de complexes biologiques à des résolutions atomiques permettant de calculer la structure atomique de ces complexes. Cependant, les challenges futurs de la cryo-ME sont encore nombreux. Bien que des programmes comme Relion ont permis d'ouvrir la cryo-ME à des utilisateurs issus de la cristallographie, il reste encore de nombreuses améliorations à faire dans la préparation des échantillons et dans le traitement des images. La manière de traiter de l'hétérogénéité sera sûrement un des challenges les plus importants de ces prochaines années puisque l'analyse de l'hétérogénéité des échantillons congelés nous renseignera également sur leur dynamique. Il faudra donc encore améliorer les programmes de traitement d'image afin d'intégrer plus finement l'analyse de l'hétérogénéité des protéines ou des complexes biologiques, aussi bien en particules isolées que sur des objets présentant une symétrie élevée comme les particules icosaédriques ou les particules à symétrie hélicoïdale. De même, il faudra également améliorer l'étape qui consiste à obtenir le premier modèle qui sera utilisé pour l'étape d'affinement de la structure. En effet, il est possible à partir des images des particules de générer des moyennes de classe 2D qui présentent un contraste élevé et de très nombreux détails et pourtant ne pas être capable de générer un premier modèle 3D fiable à basse résolution. La cryo-tomographie électronique peut être une approche à considérer pour générer un premier modèle stable, mais elle reste longue et fastidieuse à mettre en place. De même, comme il faut traiter de plus en plus d'images, l'automatisation de la plupart des étapes doit être poursuivie et intégrée dans une interface simple et évolutive. Il faudrait également pouvoir introduire un ou des critères pour automatiser la sélection des images qu'il faut retenir pour le traitement, et améliorer l'exploitation des films. On peut également se demander quelle sera la taille des plus petits objets qu'il sera possible d'imager et de reconstruire aux vues des récents progrès obtenus par les systèmes « phase plate » qui améliorent très fortement le contraste des images. Ces

évolutions feront sans conteste de la cryo-ME la technique majeure pour l'obtention de structure de protéines et de complexes biologiques dans les années à venir.

# BIBLIOGRAPHIE

# Bibliographie

---

- Abraham, A.D. et al., 2010. Two distinct nanovirus species infecting Faba bean in Morocco. *Archives of Virology*, 155(1), pp.37–46.
- Adrian, M. et al., 1984. Cryo-electron microscopy of viruses. *Nature*, 308, pp.32–36.
- Afanasyev, P. et al., 2015. A posteriori correction of camera characteristics from large image data sets. *Scientific reports*, 5, p.10317.
- Bai, X. et al., 2015. An atomic structure of human  $\gamma$ -secretase. *Nature*, 525(7568), pp.212–7.
- Bai, X.C. et al., 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife*, 2013(2), pp.1–12.
- Bai, X.C. et al., 2015. Sampling the conformational space of the catalytic subunit of human  $\gamma$ -secretase. *eLife*, 4, p.e11182.
- Bak, A. et al., 2012. Host cell processes to accomplish mechanical and non-circulative virus transmission. *Protoplasma*, 249(3), pp.529–539.
- Blanc, S. et al., 1993. Paracrystalline structure of cauliflower mosaic virus aphid transmission factor produced both in plants and in a heterologous system and relationship with a solubilized active form. *Virology*, 197(1), pp.283–292.
- Blanc, S. et al., 1996. The aphid transmission factor of cauliflower mosaic virus forms a stable complex with microtubules in both insect and plant cells. *Proceedings of the National Academy of Sciences of the United States of America*, 93(26), pp.15158–15163.
- Blomme, G. et al., 2013. A historical overview of the appearance and spread of Musa pests and pathogens on the African continent: Highlighting the importance of clean Musa planting materials and quarantine measures. *Annals of Applied Biology*, 162(1), pp.4–26.
- Booth, C. & Mooney, P., 2013. Applications of electron-counting direct-detection cameras in high-resolution cryo-electron microscopy. *Microscopy and Analysis*, 27(6), pp.13–21.
- Brenner, S. & Horne, R.W., 1959. A negative staining method for high resolution electron microscopy of viruses. *Biochimica et biophysica acta*, 34, pp.103–110.
- Briggs, J.A.G., 2013. Structural biology in situ—the potential of subtomogram averaging. *Current Opinion in Structural Biology*, 23(2), pp.261–267.
- Brilot, A.F. et al., 2012. Beam-induced motion of vitrified specimen on holey carbon film. *Journal of Structural Biology*, 177(3), pp.630–637.
- Brown, A. et al., 2015. Tools for macromolecular model building and refinement into electron cryo-microscopy reconstructions. *Acta Crystallographica Section D: Biological Crystallography*, 71, pp.136–153.
- Callaway, E., 2015. The Revolution Will Not Be Crystallized. *Nature*, 525, pp.172–174.
- Campbell, M.G. et al., 2015. 2.8 Å Resolution Reconstruction of the Thermoplasma Acidophilum 20 S Proteasome Using Cryo-Electron Microscopy. *eLife*, 2015(4), pp.1–22.
- Campbell, M.G. et al., 2012. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure*, 20(11), pp.1823–1828.
- Chen, S. et al., 2013. High-resolution noise substitution to measure overfitting and validate resolution

- in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy*, 135(C), pp.24–35.
- Chen, V.B. et al., 2010. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66(1), pp.12–21.
- Demangeat, G. et al., 2005. Survival of Xiphinema index in Vineyard Soil and Retention of Grapevine fanleaf virus Over Extended Time in the Absence of Host Plants. *Phytopathology*, 95(10), pp.1151–1156.
- Demangeat, G., 2007. Transmission des Nepovirus par les nématodes longidoridae. *Virologie*, 11(4), pp.309–321.
- Drucker, M. et al., 2002. Intracellular distribution of viral gene products regulates a complex mechanism of cauliflower mosaic virus acquisition by its aphid vector. *Proc Natl Acad Sci U S A*, 99(4), pp.2422–2427.
- Drucker, M. & Then, C., 2015. Transmission activation in non-circulative virus transmission: A general concept? *Current Opinion in Virology*, 15, pp.63–68.
- Durand, A., Papai, G. & Schultz, P., 2013. Structure, assembly and dynamics of macromolecular complexes by single particle cryo-electron microscopy. *Journal of Nanobiotechnology*, 11(Suppl 1), p.S4.
- Egelman, E.H., 2007. The iterative helical real space reconstruction method: Surmounting the problems posed by real polymers. *Journal of Structural Biology*, 157(1), pp.83–94.
- Emsley, P. & Cowtan, K., 2004. Coot -model building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography*, D60, pp.2126–2132.
- Grant, T. & Grigorieff, N., 2015a. Automatic estimation and correction of anisotropic magnification distortion in electron microscopes. *Journal of Structural Biology*, 192(2), pp.204–208.
- Grant, T. & Grigorieff, N., 2015b. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife*, 4, p.e06980.
- Grigoras, I. et al., 2014. Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe. *Journal of General Virology*, 95(PART 5), pp.1178–1191.
- Grigoras, I. et al., 2009. Reconstitution of authentic nanovirus from multiple cloned DNAs. *Journal of virology*, 83(20), pp.10778–87.
- Grigorieff, N., 2007. FREALIGN: High-resolution refinement of single particle structures. *Journal of Structural Biology*, 157(1), pp.117–125.
- Haas, M. et al., 2002. Cauliflower mosaic virus : still in the news. *Molecular Plant Pathology*, 3, pp.419–429.
- van Heel, M. & Keegstra, W., 1981. IMAGIC: A fast, flexible and friendly image analysis software system. *Ultramicroscopy*, 7(2), pp.113–129.
- Henderson, R., 2013. Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proceedings of the National Academy of Sciences of the United States of America*, 110(45), pp.18037–41.
- Hesketh, E.L. et al., 2015. Mechanisms of assembly and genome packaging in an RNA virus revealed by high-resolution cryo-EM. *Nature Communications*, 6, p.10113.
- Hoh, F. et al., 2010. Structural insights into the molecular mechanisms of cauliflower mosaic virus transmission by its insect vector. *Journal of virology*, 84(9), pp.4706–4713.

- Hohn, T., 2013. Plant pararetroviruses: Interactions of cauliflower mosaic virus with plants and insects. *Current Opinion in Virology*, 3(6), pp.629–638.
- Ilca, S.L. et al., 2015. Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nature Communications*, 6, p.8843.
- Jiang, J. et al., 2015. Atomic structure of anthrax protective antigen pore elucidates toxin translocation. *Nature*, 521(7553), pp.545–549.
- Jiang, W. & Chiu, W., 2001. Web-based Simulation for Contrast Transfer Function and Envelope Functions. *Microscopy and microanalysis : the official journal of Microscopy Society of America, Microbeam Analysis Society, Microscopical Society of Canada*, 7(4), pp.329–334.
- Kleywegt, G.J. & Jones, T.A., 1996. xDLMAPMAN and xDIDATAMAN - Programs for reformatting, analysis and manipulation of biomacromolecular electron-density maps and reflection data sets. *Acta Crystallographica Section D: Biological Crystallography*, 52(4), pp.826–828.
- Kremer, J.R., Mastrorade, D.N. & McIntosh, J.R., 1996. Computer visualization of three-dimensional image data using IMOD. *Journal of structural biology*, 116(1), pp.71–6.
- Kumari, S.G. et al., 2010. Detection and partial characterization of Milk vetch dwarf virus isolates from faba bean (*Vicia faba* L.) in Yunnan province, China. *Journal of Phytopathology*, 158(1), pp.35–39.
- Lai-Kee-Him, J. et al., 2013. The backbone model of the Arabis mosaic virus reveals new insights into functional domains of Nepovirus capsid. *Journal of Structural Biology*, 182(1), pp.1–9.
- Leh, V. et al., 1999. Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. *The EMBO Journal*, 18(24), pp.7077–7085.
- Leschziner, A., 2010. The orthogonal tilt reconstruction method. *Methods in Enzymology*, 482(C), pp.237–262.
- Leschziner, A.E. & Nogales, E., 2006. The orthogonal tilt reconstruction method: An approach to generating single-class volumes with no missing cone for ab initio reconstruction of asymmetric particles. *Journal of Structural Biology*, 153(3), pp.284–299.
- Li, X. et al., 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature methods*, 10(6), pp.584–90.
- Liao, M. et al., 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*, 504(7478), pp.107–12.
- Lin, T. et al., 1999. The Refined Crystal Structure of Cowpea Mosaic Virus at 2.8 Å Resolution. *Virology*, 265(1), pp.20–34.
- Martelli, G.P., 2014. Infectious degeneration. *Journal of Plant Pathology*, 96(1 SUPPL.), pp.1–4.
- Martiniere, A. et al., 2013. A virus responds instantly to the presence of the vector on the host and forms transmission morphs. *eLife*, 2013(2), pp.1–26.
- McMullan, G. et al., 2014. Comparison of optimal performance at 300keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy*, 147, pp.156–163.
- McMullan, G. et al., 2009. Experimental observation of the improvement in MTF from backthinning a CMOS direct electron detector. *Ultramicroscopy*, 109(9), pp.1144–1147.
- Merk, A. et al., 2016. Breaking Cryo-EM Resolution Barriers to Facilitate Article Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell*, 165, pp.1–10.
- Mindell, J.A. & Grigorieff, N., 2003. Accurate determination of local defocus and specimen tilt in

- electron microscopy. *Journal of Structural Biology*, 142(3), pp.334–347.
- Moreno, A. et al., 2005. A single amino acid position in the helper component of cauliflower mosaic virus can change the spectrum of transmitting vector species. *Journal of Virology*, 79(21), pp.13587–13593.
- Murshudov, G.N. et al., 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, 67(4), pp.355–367.
- Nature Methods, 2016. Method of the Year 2015: Cryo-EM. *Nature Methods*, 13(1), p.1.
- Nicholls, R.A., Long, F. & Murshudov, G.N., 2012. Low-resolution refinement tools in REFMAC5. *Acta Crystallographica Section D: Biological Crystallography*, 68(4), pp.404–417.
- Orlova, E. V. & Saibil, H.R., 2011. Structural analysis of macromolecular assemblies by electron microscopy. *Chemical Reviews*, 111(12), pp.7710–7748.
- Plisson, C. et al., 2005. Structure of the mature P3-virus particle complex of cauliflower mosaic virus revealed by cryo-electron microscopy. *Journal of Molecular Biology*, 346(1), pp.267–277.
- Radermacher, M. et al., 1987. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *Journal of Microscopy*, 146(2), pp.113–136.
- Razinkov, I. et al., 2016. A new method for vitrifying samples for cryoEM. *Journal of Structural Biology*, 195(2), pp.190–198.
- Rohou, A. & Grigorieff, N., 2015. *CTFFIND4: Fast and accurate defocus estimation from electron micrographs*,
- Rosenthal, P.B. & Henderson, R., 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology*, 333(4), pp.721–745.
- Russo, C.J. & Passmore, L.A., 2016. Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens. *Journal of Structural Biology*, 193(1), pp.33–44.
- Russo, C.J. & Passmore, L.A., 2014. Ultrastable gold substrates for electron cryomicroscopy. *Science (New York, N.Y.)*, 346(6215), pp.1377–80.
- Saibil, H.R., 2000. Macromolecular structure determination by cryo-electron microscopy. *Acta Crystallographica Section D: Biological Crystallography*, 56(10), pp.1215–1222.
- Schellenberger, P. et al., 2010. A stretch of 11 amino acids in the betaB-betaC loop of the coat protein of grapevine fanleaf virus is essential for transmission by the nematode *Xiphinema index*. *Journal of virology*, 84(16), pp.7924–7933.
- Schellenberger, P. et al., 2011. Structural insights into viral determinants of nematode mediated grapevine fanleaf virus transmission. *PLoS Pathogens*, 7(5), p.e1002034.
- Scheres, S.H.W., 2014. Beam-induced motion correction for sub-megadalton cryo-EM particles. *eLife*, 3, p.e03665.
- Scheres, S.H.W., 2012. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, 180(3), pp.519–530.
- Scheres, S.H.W. & Chen, S., 2012. Prevention of overfitting in cryo-EM structure determination. *Nature Methods*, 9(9), pp.853–854.
- Sicard, A. et al., 2015. Circulative Nonpropagative Aphid Transmission of Nanoviruses: an



- Oversimplified View. *Journal of virology*, 89(19), pp.9719–26.
- Sicard, A. et al., 2013. Gene copy number is differentially regulated in a multipartite virus. *Nature communications*, 4, p.2248.
- Sorzano, C.O.S. et al., 2004. XMIPP: A new generation of an open-source image processing package for electron microscopy. *Journal of Structural Biology*, 148(2), pp.194–204.
- Stavolone, L. et al., 2005. A coiled-coil interaction mediates *Cauliflower mosaic virus* cell-to-cell movement. *Proceedings of the National Academy of Sciences of the United States of America*, 102(17), pp.6219–6224.
- Studier, F.W., 2005. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification*, 41(1), pp.207–234.
- Tan, Y.Z. et al., 2015. Automated data collection in single particle electron microscopy. *Microscopy (Oxford, England)*, 65(December 2015), p.dfv369-.
- Tang, G. et al., 2007. EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology*, 157(1), pp.38–46.
- Timchenko, T. & Bernardi, F., 2007. Les nanovirus, petits virus de plantes: Similitudes et différences avec les géminivirus. *Virologie*, 11(1), pp.27–42.
- Trapani, S. & Navaza, J., 2007. AMoRe: Classical and modern. In *Acta Crystallographica Section D: Biological Crystallography*. pp. 11–16.
- Uzest, M. et al., 2007. A protein key to plant virus transmission at the tip of the insect vector stylet. *Proceedings of the National Academy of Sciences of the United States of America*, 104(46), pp.17959–17964.
- Veesler, D. et al., 2013. Atomic structure of the 75 MDa extremophile *Sulfolobus* turreted icosahedral virus determined by CryoEM and X-ray crystallography. *Proceedings of the National Academy of Sciences of the United States of America*, 110(14), pp.5504–9.
- Vigne, E., Marmonier, A. & Fuchs, M., 2008. Multiple interspecies recombination events within RNA2 of Grapevine fanleaf virus and Arabis mosaic virus. *Archives of Virology*, 153(9), pp.1771–1776.
- Winn, M.D. et al., 2011. Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D: Biological Crystallography*, 67(4), pp.235–242.
- Yan, X., Sinkovits, R.S. & Baker, T.S., 2007. AUTO3DEM-an automated and high throughput program for image reconstruction of icosahedral particles. *Journal of Structural Biology*, 157(1), pp.73–82.