



**HAL**  
open science

# Étude du mécanisme d'auto-assemblage de virus icosaédriques en microscopie de fluorescence à l'échelle de la molécule unique

Thomas Bugea

► **To cite this version:**

Thomas Bugea. Étude du mécanisme d'auto-assemblage de virus icosaédriques en microscopie de fluorescence à l'échelle de la molécule unique. Biophysique [physics.bio-ph]. Université Paris-Saclay (2020-..), 2024. Français. NNT: . tel-04922978

**HAL Id: tel-04922978**

**<https://hal.science/tel-04922978v1>**

Submitted on 31 Jan 2025

**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.

# Étude du mécanisme d'auto-assemblage de virus icosaédriques en microscopie de fluorescence à l'échelle de la molécule unique

*Visualising the self-assembly mechanism of icosahedral viruses through fluorescence microscopy at the single molecule level*

**Thèse de doctorat de l'Université Paris-Saclay**

École doctorale n°572 Ondes et Matière (EDOM)  
Spécialité de doctorat : Physique  
Graduate school : Physique. Référent : Faculté des sciences d'Orsay

Thèse préparée au sein des laboratoires LuMin (Université Paris-Saclay, ENS Paris-Saclay, CNRS, Centrale Supélec) et LPS (Université Paris-Saclay, CNRS), sous la direction de **Karen PERRONET**, chargée de recherche CNRS et le co-encadrement de **Guillaume TRESSET**, directeur de recherche CNRS.

**Thèse soutenue à Paris-Saclay, le 30/09/2024, par**

**Thomas BUGEA**

## Composition du Jury

Membres du jury avec voix délibérative

<b>Sandrine LÉVÊQUE-FORT</b> Directrice de recherche CNRS, Institut des Sciences Moléculaires d'Orsay, Orsay	Présidente
<b>Cyril FAVARD</b> Ingénieur de recherche CNRS, Institut de Recherche en Infectiologie de Montpellier, Montpellier	Rapporteur & Examineur
<b>Cendrine MOSKALENKO</b> Professeure à l'École Normale Supérieure de Lyon, Lyon	Rapporteur & Examinatrice
<b>Francesco PEDACI</b> Directeur de recherche CNRS, Centre de Biologie Structurale, Montpellier	Examineur

**Titre :** Étude du mécanisme d'auto-assemblage de virus icosaédriques en microscopie de fluorescence à l'échelle de la molécule unique

**Mots clés :** microscopie de fluorescence, molécule unique, virus, auto-assemblage, modélisation

**Résumé :** La diversité des virus qui nous entourent est telle que nous avons dû apprendre à vivre avec eux. Pour faciliter cette cohabitation et surtout lutter contre leur prolifération, il a fallu comprendre leur mode de fonctionnement. Cette nécessité a donné lieu à de nombreuses études menées sur des virus de tailles et compositions variées. Les résultats obtenus jusqu'alors ont contribué à acquérir des informations d'ensemble sur les voies d'assemblage des capsides virales et leur dépendance quant aux conditions de salinité et pH. En effet, la majorité des techniques utilisées fournissent des résultats moyennés sur toute une population, sans illustrer l'hétérogénéité inhérente au mécanisme d'assemblage. La microscopie de fluorescence par réflexion totale interne nous permet non seulement de travailler à l'échelle de la capside unique mais également de la sous-unité unique. Ce montage expérimental a été appliqué sur le virus de la marbrure chlorotique de la cornille, dont les dimères de protéines et l'ARN simple brin étaient marqués par deux fluorophores distincts. L'observation du mécanisme d'auto-assemblage de ce virus sur une lamelle de microscope a nécessité une longue optimisation de la méthode de passivation de cette der-

nière. La chimie de surface choisie et les différentes caractérisations associées nous ont prouvé que les signaux de fluorescence collectés correspondent bien aux interactions spécifiques entre sous-unités marquées et ARN simple brin. Ainsi, la mise en place de ces expériences couplée à l'utilisation d'un algorithme de détection de marches, Ruptures, nous ont permis de suivre la dynamique d'assemblage à l'équilibre du virus icosaédrique d'intérêt. Nous avons pu accéder à des informations telles que le nombre moyen de sous-unités par ARN ou la durée entre deux accroches consécutives. À la suite des premières expériences réalisées, nous avons observé, comme attendu, une augmentation de la taille des particules formées et une diminution du taux d'accroches quand la concentration totale en sous-unités augmentait. Ces résultats, combinés à ceux obtenus en photométrie de masse, nous ont indiqué que l'assemblage suivi en microscopie de fluorescence pouvait conduire à des capsides complètes. Enfin, les expériences menées hors équilibre semblaient suggérer une accélération des premières étapes d'assemblage lorsque la concentration en sel du milieu augmentait.

**Title :** Visualising the self-assembly mechanism of icosahedral viruses through fluorescence microscopy at the single molecule level

**Keywords :** fluorescence microscopy, single molecule, virus, self-assembly, modelling

**Abstract :** The diversity of viruses we are facing dictated us to learn how to live with them. To facilitate this coexistence and, above all, combat their proliferation, it has been necessary to understand how they work. This need has led to numerous studies on viruses of various sizes and compositions. Up to now, the results obtained through ensemble measurements contribute to improving our knowledge upon assembly pathways of viral capsids and their dependence on salt and pH. Indeed, most of the techniques used give average results for an entire population, without being able to access to the inherent heterogeneity of the assembly mechanism. Total internal reflection fluorescence microscopy (TIRFM) allows us to work not only at the level of the individual capsid, but also at the level of the individual subunit. This experimental setup has been used on the Cowpea Chlorotic Mottle Virus (CCMV), whose dimers and single-stranded RNA were fluorescently labelled. Observing the self-assembly mechanism of this icosahedral virus on a coverslip was achieved after a long

optimisation stage of the glass passivation method. The surface chemistry we chose, and the several characterisation steps showed that the fluorescent signals collected corresponded exactly to the specific interactions between labelled subunits and single-stranded RNA. Therefore, these experiments combined to the step-fit algorithm Ruptures enabled us to follow the self-assembly dynamics of CCMV at equilibrium. We accessed to information such as the mean number of capsid subunits per RNA and the time between two consecutive binding events. These initial experiments showed, as expected, an increase in particle size and a decrease in binding rate as the subunit concentration increased. These results, combined with those obtained by mass photometry, indicated that the self-assembly mechanism followed through fluorescence microscopy could be close to give fully formed capsids. Finally, non-equilibrium experiments seemed to suggest that the early steps of assembly were faster when the salt concentration of the medium increased.

# Remerciements

Au cours de l'été précédant le début de ma thèse, un très bon ami m'a posé cette question : "Mais trois ans c'est long ? Il y aura forcément des jours où tu n'auras rien à faire." À l'époque, sans en connaître davantage sur mon projet, je lui avais simplement répondu que ce serait étonnant, vu qu'en recherche on va toujours essayer de résoudre un problème. Aujourd'hui, à la suite de mes trois années de thèse, je peux l'affirmer : à aucun moment je ne me suis ennuyé. Il est vrai que la thématique scientifique a beaucoup aidé, mais encore plus que le projet, ce sont aussi les personnes que j'ai pu rencontrer et côtoyer qui ont permis de rendre ces trois années aussi passionnantes.

Pour commencer, mes premiers mots concernent les membres de mon jury de thèse. Je tiens à remercier Cyril Favard et Cendrine Moskalenko pour avoir pris le temps de rapporter mon travail et lire avec attention mon manuscrit en plein mois d'août. Merci également à Sandrine Lévêque-Fort et Francesco Pedaci d'avoir accepté d'évaluer mon travail. Plus globalement, merci à vous quatre pour les échanges que nous avons eus à propos de mon projet de thèse durant ma soutenance.

Ces trois années ont été menées au sein du Laboratoire de Physique des Solides et du laboratoire Lumière, Matière et Interfaces. Je souhaite notamment remercier les services de direction et administratifs des deux laboratoires qui m'ont permis de me consacrer pleinement et sans encombre à mes travaux de recherche.

Parmi les membres de ces laboratoires, il y en a deux que je souhaite tout particulièrement mettre en avant : Karen Perronet et Guillaume Tresset. Nos premiers contacts ont eu lieu par visio quand, en stage de M2 à Bordeaux, je cherchais à m'inscrire en thèse en région parisienne. Le sujet que vous proposiez correspondait tout à fait aux critères que je recherchais mais il est vrai que j'étais assez anxieux à l'idée de rejoindre un environnement inconnu. Et je pense qu'il s'agissait également d'un challenge pour vous de recruter un étudiant en thèse sans l'avoir eu en stage au préalable. Mais finalement, tout ce que je peux dire pour résumer ces trois années tient en un mot : merci.

Tout au long de ma thèse, j'ai pu en apprendre énormément sur différents domaines scientifiques grâce à vous. Vous m'avez aiguillé mais également soutenu quand les difficultés étaient présentes. Il est vrai que j'ai passé plus de temps au LuMIn qu'au LPS mais dans tous les cas, peu importe le moment, vous avez su être disponibles et à l'écoute. Karen, tu as toujours trouvé le temps pour venir me voir dans la salle de manip quand j'étais en expérience. Sur le nombre incalculable de discussions que nous avons pu avoir à ce moment, je ne compte pas le nombre de fois où nous avons fini par rire de certains problèmes rencontrés (au hasard la chimie de surface). De mon côté également, je suis souvent passé à ton bureau pour te parler des derniers résultats ou de certaines difficultés et à chaque fois j'en suis reparti avec une ou plusieurs idées que tu avais pu me suggérer. Guillaume, j'aime retenir les mails sûrement un peu trop détaillés que je n'ai pas hésité à t'envoyer mais auxquels tu as toujours pris le temps de répondre avec beaucoup de précisions. À plusieurs reprises également ces échanges de mails ont mené à de longues

---

discussions à ton bureau, sans que je ne te prévienne en amont pour autant. Tu as d'ailleurs souvent eu le dernier mot qui m'a permis de dédramatiser vis-à-vis de certains résultats un peu moins réjouissants. Les conversations que j'ai eues avec chacun d'entre vous se sont généralement poursuivies par nos habituelles "courtes" réunions mensuelles qui m'ont permis de faire le point et de repartir encore plus motivé vers de nouvelles expériences.

À tous les deux, vous m'avez permis d'évoluer dans un environnement de travail aussi serein que passionnant et je vous en suis reconnaissant. Du fond du coeur, merci à vous deux.

Au LPS, je souhaite remercier Laetitia Gargowitsch pour m'avoir aidé à plusieurs reprises durant les purifications de virions et préparations de tampons. Merci pour tes nombreux conseils, ta disponibilité et d'avoir pris le temps de m'expliquer à plusieurs reprises certaines manipulations que tu m'avais déjà montrées. Je te souhaite le meilleur pour la suite et notamment pour la thèse que tu viens d'entamer.

Toujours au LPS, j'ai pu partager un bureau avec deux étudiantes en thèse qui se sont succédé. Merci à Kalouna Kra pour les différents moments que nous avons partagés ensemble (expérience à SOLEIL, GRC en Italie, tentatives de désassemblage de la CP183EEE), c'était vraiment super. Bon courage pour la suite au Canada et j'espère qu'on aura l'occasion de se recroiser en conférence ou ailleurs. Merci également à Tamizhmalar Sundararajan pour ton enthousiasme et ta bonne humeur communicative à chaque fois que je suis venu au LPS. Good luck for your driving lessons!

Au LuMIn désormais, je tiens à remercier les différents membres de l'équipe qui m'ont accueilli. Merci notamment à François Marquier pour tes conseils scientifiques et les échanges que nous avons pu avoir sur ton expérience en tant que Professeur mais aussi pour les anecdotes, entre autres, sur l'Institut d'Optique. Merci également à François Treussart pour nos discussions au laboratoire mais aussi à MiFoBio et de m'avoir permis de participer à l'organisation de PhysBio 2024. Je souhaite aussi remercier Simon l'Horset, Sébastien Rousselot et Colin Lopez pour m'avoir apporté votre aide afin d'améliorer le montage expérimental.

Je veux désormais dire un grand merci aux différents étudiants en thèse avec qui j'ai eu la chance de partager le bureau 3S68 : Florian Semmer, Marie-Charlotte Emperauger, Baptiste Bouhet, Raphaël Rousselle et Rémi Coq. Florian, je te remercie pour ton aide, tes conseils et ta curiosité sans limite en ce qui concerne les différents sujets sur lesquels nous avons pu échanger. Bon courage pour la suite et notamment pour ce projet sur l'iSCAT. En ayant pu voir ta persévérance et ta motivation en thèse, nul doute que les résultats seront positifs. Marie-Charlotte, tu es la seule avec qui j'ai pu partager un bureau sans discontinuité pendant ces trois années, et j'en suis ravi. Je tiens à te remercier pour ta bonne humeur, ta bienveillance et pour toutes ces discussions que nous avons eues durant nos thèses respectives. Il ne me reste plus qu'à te souhaiter la meilleure des carrières dans l'enseignement, tu feras une excellente enseignante en classes préparatoires. Baptiste, je te remercie pour les très bons moments que nous avons passés ensemble en manip, en enseignement ou encore à MiFoBio. Bon courage avec les cours à l'IUT, bonne suite de thèse et en espérant que tous tes efforts sur la pince magnétique soient récompensés. Raphaël et Rémi, j'ai vraiment apprécié nos conversations, qu'elles aient eu lieu dans notre bureau ou dans le couloir avant votre arrivée. Je vous souhaite à tous les deux le meilleur pour vos derniers mois en thèse au sein du bureau 3S68 mais également pour vos projets futurs.

Par ailleurs, même si je n'ai pas partagé de bureaux avec vous, je remercie également Véronique Vienne et Baptiste Grimaud. Véronique, merci pour les conversations que nous avons pu avoir sur la manip, les enseignements, tes cours de danse et bien d'autres choses. Baptiste, merci à toi de ne pas m'avoir laissé trop seul au laboratoire les nombreuses fois où les expériences ont un peu trop duré.

Je remercie également tous les doctorants et post-doctorants que j'ai pu rencontrer au laboratoire :

---

Athulya, Fabienne, Isis, Alexandre, Gizem, Liam. Pour finir sur le LuMIn, je tiens à remercier les trois stagiaires qui ont contribué à faire avancer ma thèse : Ethan Saïman, Roxanne Tamarii et Roméo Suss. Une pensée particulière pour toi Roméo à qui je passe le flambeau sur ce projet riche en futures expériences. Tu as tout ce qu'il faut pour apporter de nouvelles idées et atteindre les objectifs fixés : rendez-vous à ta soutenance !

En ce qui concerne certains des résultats décrits dans mon manuscrit, je tiens à remercier Charles Truong pour ton aide sur l'implémentation de l'algorithme Ruptures à nos données expérimentales. Merci également à Bertrand Raynal et à Maeleenn Chevreuil pour m'avoir permis de réaliser des expériences de photométrie de masse à l'Institut Pasteur.

Mes trois années en thèse ont également été agrémentées de missions d'enseignement à l'IUT d'Orsay. Je remercie les différentes personnes qui m'ont aidé durant mes séances de TD et TP : Fabienne Goldfarb, Grégory Gredat, Fabien Legrand et Jérémy Ferrand. Une attention particulière pour toi Loïc Rondin, qui m'a donné l'opportunité d'avoir cette première expérience dans l'enseignement.

J'en profite désormais pour m'éloigner de l'environnement du travail en remerciant ma famille et en commençant notamment par mes parents. Même si je sais que vous avez dû vous demander de nombreuses fois ce que je pouvais bien faire aussi tard au travail, vous m'avez soutenu et surtout accompagné jusqu'à finalement obtenir un doctorat. Je remercie également mon frère Nicolas, ma sœur Aurélie et mon beau-frère Franck pour votre curiosité à chacun sur mon travail en thèse. Merci à vous tous d'avoir pu assister à ma soutenance, cela comptait beaucoup pour moi.

Je souhaite, presque, terminer en remerciant une dernière poignée de personnes. Merci à Giulia, Anne, Robin et Clément que j'ai pu rencontrer grâce à quelqu'un dont je n'ai pas encore parlé (ton paragraphe arrive). Merci également à mon canadien préféré Zackie, notamment pour nos parties de tennis.

Cette thèse a fait suite à trois années passées à l'Institut d'Optique et durant lesquelles j'ai notamment rencontré Fabien, Inès et Blandine avec qui j'ai passé de très bons moments. Merci à vous !

Je remercie également mes trois meilleurs amis Thomas, Quentin et Fantin. Vous m'accompagnez tous les trois depuis de nombreuses années désormais et je tenais à vous dire un grand merci. Vous m'avez souvent demandé des nouvelles concernant ma thèse pour savoir si tout se passait bien mais surtout les moments passés ensemble m'ont permis de me vider l'esprit et penser à autre chose.

Pour finir, ces derniers mots sont pour toi Stella, Voilà maintenant près de six ans que nous avons décidé de faire de nos vies une aventure commune. Une étape supplémentaire vient d'être franchie et le fait d'avoir pu la passer à tes côtés ne pouvait pas me rendre plus heureux. Tout au long de ces trois années tu as été là pour m'accompagner, me rassurer et il n'y a pas un moment durant lequel je ne me suis pas senti soutenu. Merci pour tous tes conseils, tes attentions au quotidien, comme nos petits rituels qui se terminaient généralement par un traditionnel résumé de notre journée. Peut-être que j'étais moins bavard que toi dans ces moments mais c'était aussi appréciable de moins penser virus ou fluo. Une chose est sûre : mon manuscrit et mes résultats sont aussi les tiens et j'espère être à la hauteur pour t'accompagner dans ton récit final sur les trois photons. Et une fois cette nouvelle étape passée, il me tarde de voir ce que l'avenir va bien pouvoir nous réserver ?

---

# Table des matières

Table des matières	v
Table des figures	ix
Liste des tableaux	xiii
<b>1 Introduction</b>	<b>1</b>
<b>2 Les virus</b>	<b>3</b>
2.1 Généralités . . . . .	4
2.1.1 Composition d'un virus . . . . .	4
2.1.2 Comment classer les virus? . . . . .	5
2.1.2.1 La capsid virale comme classification . . . . .	5
2.1.2.2 La classification de Baltimore . . . . .	7
2.1.3 Cycle viral . . . . .	7
2.1.4 Mécanismes d'assemblage connus de la capsid virale . . . . .	10
2.1.4.1 Formation de capsides vides . . . . .	10
2.1.4.2 Formation de capsides pleines . . . . .	14
2.1.5 Conclusion . . . . .	18
2.2 Le CCMV . . . . .	19
2.2.1 La capsid protéique . . . . .	19
2.2.2 Le matériel génétique . . . . .	20
2.3 Quantifier l'assemblage de virus ... . . . .	21
2.3.1 ... à l'aide de mesures d'ensemble . . . . .	21
2.3.1.1 Diffusion dynamique de la lumière . . . . .	21
2.3.1.2 Diffusion des rayons X aux petits angles . . . . .	23
2.3.1.3 Chromatographie d'exclusion de taille . . . . .	24
2.3.2 ... à l'échelle du virus unique . . . . .	26
2.3.2.1 Méthodes de mesures indirectes . . . . .	26
2.3.2.2 Méthodes d'observations directes . . . . .	29
2.3.3 Conclusion . . . . .	32
<b>3 Imagerie d'un virus à l'échelle de la molécule unique</b>	<b>33</b>
3.1 Le TIRFM . . . . .	34
3.1.1 Principe de la technique . . . . .	34
3.1.2 Description du montage . . . . .	36



3.1.2.1	Voies d'illumination et de détection . . . . .	36
3.1.2.2	Maintien de la focalisation dans le temps . . . . .	41
3.1.2.3	Déroulement d'une acquisition . . . . .	43
3.1.3	La cellule micro-fluidique . . . . .	44
3.2	Fonctionnalisation des lamelles de microscope . . . . .	46
3.2.1	Focus sur la chimie de surface PLL-g-PEG . . . . .	48
3.2.1.1	Présentation de l'architecture . . . . .	48
3.2.1.2	Les premiers tests . . . . .	49
3.2.1.3	Est-il possible de combler les trous de la chimie de surface? . . . . .	52
3.2.1.4	La recette finale . . . . .	60
3.2.2	Comparaison avec d'autres types de chimie de surface . . . . .	64
3.2.2.1	La chimie de surface à base de BSA . . . . .	64
3.2.2.2	La chimie de surface PEG-Silane . . . . .	66
3.2.2.3	La chimie de surface de type chimie clic . . . . .	69
3.2.3	L'hybridation de l'ARN simple brin . . . . .	71
3.2.4	Conclusion . . . . .	74
3.3	Retour sur la fluorescence . . . . .	76
3.3.1	Principes et mécanismes importants . . . . .	76
3.3.2	Caractérisation de la photostabilité de l'Atto647N . . . . .	78
3.3.2.1	Photoblanchiment de l'Atto647N . . . . .	80
3.3.2.2	Clignotements de l'Atto647N . . . . .	83
3.3.3	Vers l'amélioration de la photostabilité de l'Atto647N . . . . .	85
3.3.3.1	Mécanisme général d'un système désoxygénant . . . . .	85
3.3.3.2	Application expérimentale sur l'Atto647N . . . . .	88
3.3.4	Stratégie de marquage choisie . . . . .	91
3.4	Premières étapes d'analyse des données brutes . . . . .	93
3.4.1	Traitement des films enregistrés . . . . .	94
3.4.2	Détection des fluorophores et suivi de l'intensité . . . . .	97
3.5	Conclusion . . . . .	101
<b>4</b>	<b>Détection de marches sur des données obtenues en microscopie de fluorescence</b>	<b>103</b>
4.1	Les algorithmes de détection . . . . .	104
4.1.1	Nécessité d'une telle méthode pour nos expériences . . . . .	104
4.1.2	Présentation de l'algorithme Ruptures . . . . .	105
4.1.3	Comparaison avec un autre algorithme : quickPBSA . . . . .	107
4.2	Introduction à la simulation de données . . . . .	108
4.2.1	Modélisation de la dynamique d'assemblage à l'équilibre . . . . .	109
4.2.2	Génération des jeux de données simulées . . . . .	112
4.2.3	Présentation des métriques utilisées en <i>Machine Learning</i> . . . . .	114
4.3	Analyse des performances des algorithmes de détection ... . . . .	115
4.3.1	... pour des caractéristiques photophysiques simples . . . . .	117
4.3.1.1	Ruptures : choix de la pénalité adaptée . . . . .	117
4.3.1.2	Capacité des algorithmes à bien détecter les marches . . . . .	120
4.3.1.3	Influence sur le calcul des paramètres d'assemblage . . . . .	121
4.3.2	... pour un fluorophore qui présente des clignotements . . . . .	125

---

4.4	Conclusion	131
<b>5</b>	<b>Dynamique d'auto-assemblage du CCMV à l'équilibre et hors équilibre</b>	<b>135</b>
5.1	Caractérisation de nucléocapsides en photométrie de masse	136
5.1.1	Réaliser une expérience en photométrie de masse	136
5.1.1.1	Description de la technique	136
5.1.1.2	Conditions expérimentales utilisées	137
5.1.1.3	Caractérisation de virions natifs	139
5.1.2	Suivi de la taille des objets en fonction de la concentration en sous-unités	140
5.1.2.1	À partir de sous-unités marquées	140
5.1.2.2	Comparaison avec une situation classique, sans marquage	143
5.1.2.3	Justification des objets de masse plus élevée qu'attendue	146
5.1.3	Conclusion	147
5.2	Présentation des conditions expérimentales en microscopie de fluorescence	149
5.3	Suivons-nous réellement un assemblage de capsides ?	150
5.3.1	Les sous-unités marquées interagissent peu avec la NeutrAvidine	151
5.3.2	Le marqueur Atto647N n'interagit pas avec l'ARN	153
5.4	Analyse de la dynamique d'assemblage à l'équilibre	155
5.4.1	Du microscope à l'étude des signaux	155
5.4.1.1	Description d'une expérience typique	155
5.4.1.2	Protocole d'analyse mis en place	157
5.4.2	Vérification expérimentale de la validité du modèle	158
5.4.2.1	Preuve de l'équilibre thermodynamique	158
5.4.2.2	Influence du photoblanchiment sur l'estimation de $\langle n \rangle$	160
5.4.3	Effet de la concentration en protéines sur la dynamique d'assemblage	162
5.4.3.1	Étude à concentration constante	162
5.4.3.2	Influence de la concentration totale en protéines	164
5.4.4	Conclusion	168
5.5	Étude de la cinétique d'assemblage hors équilibre	170
5.5.1	Description de l'expérience	170
5.5.2	Observation de la cinétique des premières étapes d'assemblage	171
5.5.2.1	Évolution typique des intensités de fluorescence	171
5.5.2.2	Influence de la salinité sur la cinétique d'assemblage	173
5.5.2.3	Estimation du nombre de sous-unités marquées par ARN	175
5.5.3	Conclusion	179
5.6	Conclusion du chapitre	180
<b>6</b>	<b>Vers l'implémentation d'un montage iSCAT</b>	<b>183</b>
6.1	La microscopie à diffusion interférométrique	184
6.1.1	Les principes théoriques	184
6.1.2	Exemples de configurations expérimentales	186
6.2	Stratégies de couplage avec la microscopie TIRF	189
6.2.1	Choix d'une illumination en champ large	189
6.2.2	Les premières images	191
6.2.3	Les améliorations envisagées	193

<b>7 Conclusions et perspectives</b>	<b>195</b>
<b>A Des virions natifs de CCMV aux sous-unités de capsid marquées</b>	<b>199</b>
<b>B Synthèse de l'ARN C2 du CCMV</b>	<b>207</b>
<b>C Protocole de chimie de surface</b>	<b>211</b>
<b>D Double hybridation de l'ARN simple brin</b>	<b>215</b>
<b>E Protocole d'analyse des données de photophysique</b>	<b>221</b>
<b>F Évaluation des performances de Ruptures et quickPBSA</b>	<b>225</b>
<b>G Photométrie de masse</b>	<b>231</b>
<b>H Dynamique d'assemblage à l'équilibre</b>	<b>237</b>
<b>Bibliographie</b>	<b>241</b>

# Table des figures

2.1	Différentes représentations d'un virus . . . . .	3
2.2	Schématisations d'un virus . . . . .	4
2.3	Diversité des capsides virales . . . . .	6
2.4	Classification de Baltimore . . . . .	7
2.5	Schématisation du cycle viral . . . . .	8
2.6	Observation de plusieurs étapes du cycle viral . . . . .	9
2.7	Influence du pH et de la salinité sur la structure des capsides vides . . . . .	11
2.8	Chemin d'assemblage proposé pour le virus de l'hépatite B . . . . .	12
2.9	Observations expérimentales caractérisant l'auto-assemblage de HBV . . . . .	13
2.10	Deux voies d'assemblage pour un mélange protéines de capside et matériel génétique . . .	15
2.11	Données TR-SAXS caractérisant l'assemblage des sous-unités de CCMV autour de son ARN viral . . . . .	16
2.12	Modèle d'un mécanisme d'assemblage contrôlé par des interactions spécifiques entre ARN et protéines . . . . .	18
2.13	Structures de la capside et des protéines de CCMV . . . . .	20
2.14	Le CCMV en tant que virus multipartite . . . . .	21
2.15	Principe d'une mesure par DLS et exemple expérimental . . . . .	22
2.16	Principe d'une mesure par SAXS et exemple expérimental . . . . .	24
2.17	Principe d'une mesure par SEC et exemple expérimental . . . . .	25
2.18	Schéma de l'illumination en FCS et exemple expérimental . . . . .	27
2.19	Schéma d'un spectromètre de masse et exemple expérimental . . . . .	28
2.20	Observations de capsides de CCMV en TEM . . . . .	29
2.21	Principe d'une mesure par AFM et exemple expérimental . . . . .	30
2.22	Schéma de l'illumination en iSCAT et exemple expérimental . . . . .	31
3.1	Vers l'imagerie d'un virus à l'échelle de la molécule unique . . . . .	33
3.2	Propagation de rayons lumineux à travers une interface verre-eau . . . . .	34
3.3	Excitation de marqueurs fluorescents en illumination TIRF . . . . .	35
3.4	Schéma du montage optique permettant une illumination TIRF . . . . .	37
3.5	Représentation de l'état de polarisation après passage par une lame demi-onde et un PBS	39
3.6	Exemples d'images expérimentales acquises en microscopie TIRF . . . . .	41
3.7	Schéma explicatif du maintien de la focalisation au cours du temps . . . . .	42
3.8	Cycle de synchronisation entre la caméra et l'obturateur de faisceau durant une acquisition	43
3.9	Schéma d'une cellule micro-fluidique . . . . .	44
3.10	Cellule micro-fluidique sur le microscope . . . . .	45

3.11	Mise en évidence de la distinction entre molécules uniques . . . . .	46
3.12	Observation des accroches non spécifiques des dimères marqués sur la lamelle de verre . .	47
3.13	Formule semi-développée d'une chaîne de PLL-biotine . . . . .	49
3.14	Influence du couple biotine-NeutrAvidine sur les accroches non spécifiques en protéines . .	50
3.15	Mise en lumière des accroches non spécifiques de la NeutrAvidine . . . . .	51
3.16	Effet d'une passivation supplémentaire à partir de sous-unités . . . . .	53
3.17	Effet d'une passivation supplémentaire à partir de lysozyme . . . . .	55
3.18	Effet d'une passivation supplémentaire à partir de BSA et BSA-biotinylée . . . . .	56
3.19	Évaluation des performances d'une passivation à partir de PLL-g-PEG et BSA-biotinylée	57
3.20	Effet de la lysine sur le nombre d'accroches non spécifiques en protéines marquées . . . .	59
3.21	Effet de la concentration en PLL-biotine sur les accroches en protéines marquées et Atto565- biotinylé . . . . .	61
3.22	Observation des accroches en ARN et protéines durant une expérience d'assemblage . . .	62
3.23	Amélioration de la statistique en ARN marqués au cours d'une expérience d'assemblage .	63
3.24	Fonctionnalisation PLL-g-PEG et PLL-biotine utilisée en expérience . . . . .	64
3.25	Adsorptions non spécifiques sur une chimie de surface de type BSA . . . . .	65
3.26	Fonctionnalisation de surface PEG-Silane et PEG-biotine . . . . .	67
3.27	Adsorptions non spécifiques sur une chimie de surface de type PEG-Silane et PEG-biotine	68
3.28	Fonctionnalisation de surface de type chimie clic . . . . .	69
3.29	Adsorptions non spécifiques sur une chimie de surface de type chimie clic . . . . .	70
3.30	Illustration de l'assemblage sous forme de <i>cherry bomb</i> . . . . .	71
3.31	Comparaison des constructions d'hybridation . . . . .	72
3.32	Spécificité de la fonctionnalisation de surface pour l'ARN hybridé . . . . .	73
3.33	Diagramme de Jablonski . . . . .	77
3.34	Illustration d'une couverture en Atto647N pour une expérience de photophysique . . . . .	79
3.35	Observation du photoblanchiment sur l'Atto647N . . . . .	80
3.36	Caractérisation du photoblanchiment sur l'Atto647N . . . . .	81
3.37	Comparaison de la durée moyenne avant photoblanchiment de l'Atto647N et du Cy5 . . .	82
3.38	Observation de clignotements avec l'Atto647N . . . . .	83
3.39	Caractérisation des durées de passages en OFF pour l'Atto647N . . . . .	84
3.40	Fonctionnement du système ROXS . . . . .	86
3.41	Mécanisme réactionnel entre le PCA et l'oxygène catalysé par le PCD . . . . .	87
3.42	Évolution de $F_\infty$ dans le tampon ROXS-A en fonction de la durée d'incubation . . . . .	89
3.43	Mécanisme réactionnel entre une fonction NHS-Ester et une lysine . . . . .	92
3.44	Comparaison des membranes de dialyse pour deux lots de protéines marquées différemment	93
3.45	Observation du fond de fluorescence sur une image brute . . . . .	94
3.46	Simulation de l'intensité d'une image brute . . . . .	95
3.47	Résultat du traitement d'images . . . . .	97
3.48	Tableau Python regroupant les intensités de fluorescence de sites d'assemblage . . . . .	98
3.49	Caractérisation de la dérive mécanique observée durant une acquisition . . . . .	99
3.50	Suivi de l'intensité de fluorescence provenant de l'accroche de protéines marquées sur un ARN simple brin . . . . .	100
4.1	Détection de marches dans un signal bruité . . . . .	103
4.2	Phase d'entraînement de Ruptures . . . . .	106

4.3	Exemple de détection par quickPBSA . . . . .	108
4.4	Illustration de la dynamique d'assemblage d'une capsid à l'équilibre . . . . .	109
4.5	Illustration des trois séries de données simulées . . . . .	116
4.6	Comparaison d'une fonction de coût sur des données expérimentales et simulées . . . . .	117
4.7	Représentation des métriques d'analyse pour le jeu de données D2 . . . . .	119
4.8	Distributions des temps entre deux marches montantes consécutives . . . . .	122
4.9	Distributions du nombre de sous-unités marquées par ARN pour la série D2 . . . . .	124
4.10	Illustration des clignotements sur la série D2 . . . . .	126
4.11	Influence des clignotements sur la distribution des $\Delta t_{on}$ . . . . .	128
4.12	Application de Ruptures et quickPBSA aux données expérimentales . . . . .	134
5.1	Assemblage de capsides de CCMV vu sous deux conditions différentes . . . . .	135
5.2	Principe de la photométrie de masse . . . . .	137
5.3	Photométrie de masse : caractérisation de virions natifs de CCMV . . . . .	139
5.4	Photométrie de masse : caractérisation d'assemblages à partir de dimères marqués . . . . .	141
5.5	Photométrie de masse : caractérisation de la population obtenue à partir de 300 nM en protéines dont 14,2 % sont marquées . . . . .	142
5.6	Photométrie de masse : caractérisation d'assemblages à partir de dimères non marqués . . . . .	143
5.7	Photométrie de masse : caractérisation de la population obtenue à partir de 300 nM en protéines sans marquage . . . . .	144
5.8	Photométrie de masse : caractérisation de l'ARN utilisé en expérience d'assemblage . . . . .	148
5.9	Comparaison des interactions spécifiques entre sous-unités marquées et ARN et non spécifiques entre sous-unités marquées et NeutrAvidine . . . . .	152
5.10	Interactions non spécifiques entre ARN hybridés et Atto647N-NHS-Ester . . . . .	154
5.11	Dynamique à l'équilibre : observation expérimentale de l'atteinte de l'équilibre thermodynamique . . . . .	159
5.12	Dynamique à l'équilibre : influence de la concentration totale en protéines sur l'apparition de l'équilibre thermodynamique . . . . .	160
5.13	Dynamique à l'équilibre : influence du photoblanchiment sur l'estimation de $\langle n \rangle$ au cours d'une expérience . . . . .	161
5.14	Dynamique à l'équilibre : distributions des paramètres d'assemblage mesurés expérimentalement pour 50 nM en protéines . . . . .	163
5.15	Dynamique à l'équilibre : caractérisation de la reproductibilité des mesures pour 50 nM en protéines . . . . .	164
5.16	Dynamique à l'équilibre : influence de la concentration totale en protéines sur $\langle n \rangle$ . . . . .	165
5.17	Dynamique à l'équilibre : caractérisation de la dynamique des sous-unités de CCMV en fonction de la taille des objets observés . . . . .	166
5.18	Dynamique à l'équilibre : influence du lot de protéines étudié sur l'estimation des paramètres d'assemblage . . . . .	167
5.19	Cinétique hors équilibre : exemple de couvertures en ARN et dimères marqués . . . . .	171
5.20	Cinétique hors équilibre : évolutions typiques de l'intensité de fluorescence émise par des sous-unités marquées sur un site d'ARN . . . . .	172
5.21	Cinétique hors équilibre : influence du sel sur la distribution des instants où $I_{max}$ est atteinte . . . . .	173
5.22	Cinétique hors équilibre : évolution de $t_{median}$ en fonction de la concentration en sel . . . . .	174
5.23	Cinétique hors équilibre : évolution de $\langle n \rangle$ pour 100 mM en NaCl . . . . .	175

---

5.24	Cinétique hors équilibre : influence de la concentration en sel sur $\langle n \rangle$ . . . . .	176
5.25	Cinétique hors équilibre : évaluation de l'influence du photoblanchiment sur la décroissance de $\langle n \rangle$ . . . . .	178
6.1	Illustration du montage expérimental iSCAT couplé au TIRFM . . . . .	183
6.2	Deux exemples de configurations de microscopie par diffusion . . . . .	185
6.3	Observation de virus avec un montage iSCAT à partir d'une illumination en champ large . . . . .	187
6.4	Observation de virus avec un montage iSCAT à partir d'une illumination à balayage . . . . .	188
6.5	Schéma du montage iSCAT en champ large couplé au TIRFM . . . . .	190
6.6	Observation expérimentale de sphères fluorescentes en iSCAT et TIRFM . . . . .	192
A.1	Feuilles de cornille . . . . .	199
A.2	Absorbances d'une solution de virions . . . . .	201
A.3	Absorbances d'une solution de protéines marquées avec un Atto647N . . . . .	205
C.1	Résumé de l'optimisation de la chimie de surface PLL-g-PEG et PLL-biotine, partie 1 . . . . .	213
C.2	Résumé de l'optimisation de la chimie de surface PLL-g-PEG et PLL-biotine, partie 2 . . . . .	214
D.1	Interactions non spécifiques issues des oligo-Atto565 . . . . .	217
F.1	Fonctions de coût pour des données d'entraînement avec et sans clignotements . . . . .	227
F.2	Évolution du $F_{score}$ en fonction du seuil quickPBSA . . . . .	229
G.1	Photométrie de masse : calibration expérimentale de l'appareil <i>Refeyn TwoMP</i> . . . . .	232
G.2	Photométrie de masse : caractérisation de dimères marqués et non marqués . . . . .	234
G.3	Photométrie de masse : illustration de la différence de masse des objets assemblés avec et sans marquage . . . . .	236
H.1	Dynamique à l'équilibre : distributions de $n$ et $\Delta t_{on}$ pour 200 et 400 nM en protéines . . . . .	238

# Liste des tableaux

3.1	Influence des conditions d'illumination sur les paramètres de photoblanchiment de l'Atto647N	82
3.2	Influence des conditions d'illumination sur les paramètres de clignotements de l'Atto647N	85
3.3	Système ROXS : composition des solutions tampons utilisées	88
3.4	Système ROXS : conditions expérimentales d'illumination	88
3.5	Système ROXS : résultats expérimentaux	90
4.1	Hétérogénéité des variations d'intensité détectées au cours d'une expérience d'assemblage	105
4.2	Paramètres impliqués dans le modèle de dynamique d'assemblage à l'équilibre	110
4.3	Paramètres de simulation de signaux mimant la dynamique d'assemblage de sous-unités marquées	115
4.4	Données d'entraînement : recherche de la meilleure pénalité	118
4.5	Données d'entraînement et de validation : comparaison des métriques selon la pénalité choisie	120
4.6	Données tests : comparaison des métriques obtenues avec quickPBSA et Ruptures	120
4.7	Données tests : estimation des constantes d'association $k_{on}$	123
4.8	Données tests : estimation du nombre moyen de sous-unités marquées $\langle n \rangle$	125
4.9	Données tests : métriques caractérisant quickPBSA et Ruptures avec et sans clignotements	127
4.10	Données tests : comparaison des $k_{on}$ avec et sans clignotements	129
4.11	Données tests : comparaison des $\langle n \rangle$ avec et sans clignotements	130
5.1	Photométrie de masse : caractérisation des populations observées avec et sans marquage	145
5.2	Composition des solutions tampons essentielles aux expériences de molécule unique	149
5.3	Analyse quantitative des interactions spécifiques et non spécifiques impliquant les sous-unités marquées	153
5.4	Analyse quantitative des interactions non spécifiques entre ARN et Atto647N-NHS-Ester	155
A.1	Composition des tampons de purification	200
A.2	Valeurs des coefficients d'extinction molaire utilisés en purification	203
D.1	Effet du rinçage sur les accroches en oligo-Atto565 et ARN hybridés	218
F.1	Données d'entraînement : comparaison des pénalités avec et sans clignotements	226
F.2	Seuils utilisés avec quickPBSA	228
H.1	Dynamique à l'équilibre : résumé des valeurs des paramètres d'assemblage pour les trois lots de protéines	239





# Chapitre - 1

---

## Introduction

---

Notre histoire commune avec les virus ne date pas de leur découverte. En effet, bien avant que la communauté scientifique ne comprenne que ces particules invisibles à l'oeil nu sont à l'origine de certaines maladies, l'Homme a dû apprendre à coexister avec. C'est au fur et à mesure des différentes découvertes scientifiques que la compréhension de leur mode de prolifération a pu être affinée et que les traitements se sont améliorés. Néanmoins, la récente pandémie du SARS-CoV-2 nous rappelle que les virus peuvent toujours continuer à nous surprendre, et ce malgré plusieurs décennies de recherche les concernant.

Grâce à ces nombreux travaux, il nous a été possible de comprendre à quel point le cycle viral est un mécanisme remarquable. De l'entrée d'un virus au sein d'une cellule hôte, à la sortie de ses différentes copies, les étapes mises en jeu ont toutes une importance primordiale. Parmi celles-ci, l'auto-assemblage des nouvelles nanoparticules est à mettre en avant. En effet, c'est à ce stade du cycle, où le virus est le plus vulnérable, que les protéines virales et le matériel génétique de ce dernier vont venir s'associer pour former de nouvelles particules infectieuses. Ce processus d'auto-assemblage, qui a lieu au sein d'un environnement encombré et complexe, conduit à la formation d'un très grand nombre de copies toutes identiques en taille et en composition. De nombreuses études *in vitro* ont été réalisées afin de comprendre les mécanismes impliqués au cours de cette étape. Les résultats décrits dans la littérature ont permis de mettre en évidence plusieurs voies d'assemblage, qui peuvent varier d'un virus à un autre. Par ailleurs, certains paramètres, tels que la salinité ou le pH du milieu, ont une influence sur les interactions entre les composants d'une particule virale et les structures des objets finaux. Néanmoins, la grande majorité des travaux menés vis-à-vis d'une meilleure compréhension du mécanisme d'auto-assemblage l'ont été au travers de mesures d'ensemble. Les techniques concernées ne permettent d'accéder qu'à des résultats moyennés sur toute une population, sans avoir accès à l'hétérogénéité inhérente d'un tel processus. De plus, même si certaines études

---

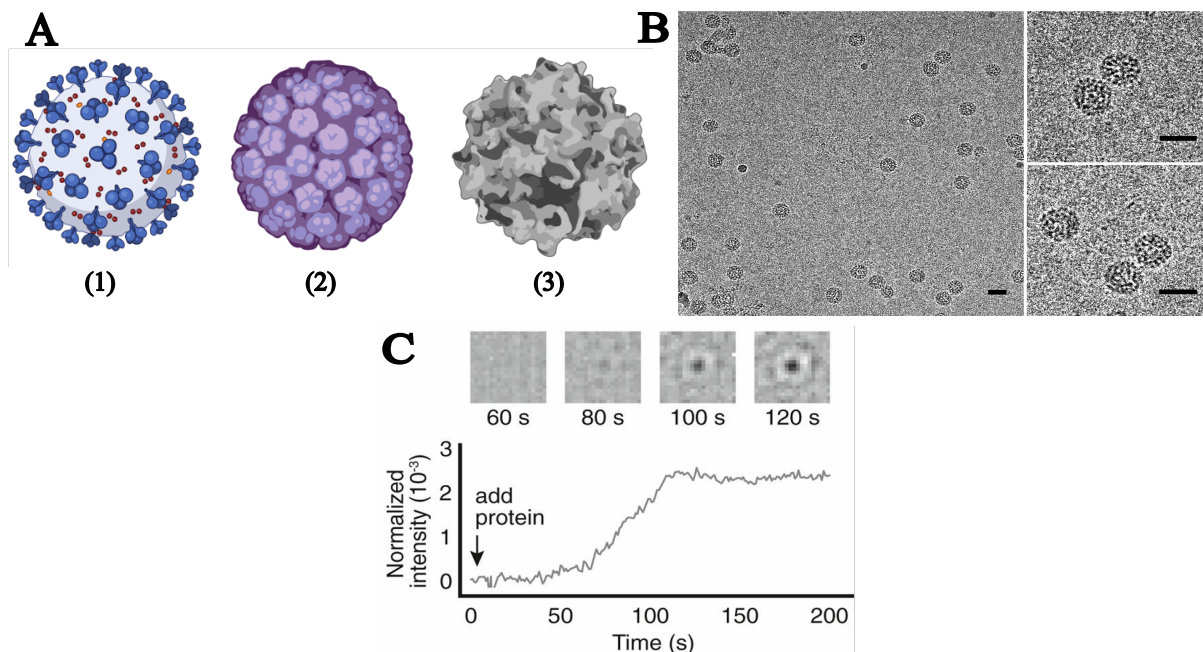
ont montré qu'elles peuvent suivre la formation de virus en temps réel, aucune ne peut accéder à la dynamique d'assemblage de protéines virales individuelles autour de leur matériel génétique. L'objectif de cette thèse est d'apporter des informations temporelles sur ce point en se plaçant non seulement à l'échelle du virus unique mais également de l'accroche unique. Les travaux détaillés ici ont principalement été menés dans deux laboratoires : le Laboratoire de Physique des Solides (LPS) et le Laboratoire Lumière, Matière et Interfaces (LuMIn).

Le chapitre 2 du manuscrit constitue une introduction générale sur les virus et leur caractérisation *in vitro*. En premier lieu, j'adopte un point de vue général sur la composition d'une particule virale et les principaux mécanismes d'auto-assemblage, avant de faire une description plus détaillée du virus sur lequel j'ai travaillé pendant trois ans : le virus de la marbrure chlorotique de la cornille (*Cowpea Chlorotic Mottle Virus* ou CCMV). J'insiste également sur les différences entre les techniques expérimentales qui effectuent des mesures d'ensemble et celles permettant de travailler à l'échelle de la molécule unique. Dans le chapitre 3, je décris les briques essentielles du projet de thèse. Il s'agit notamment de présenter le montage de microscopie de fluorescence par réflexion totale interne (*Total Internal Reflection Fluorescence Microscopy* ou TIRFM), les principes photophysiques qui en découlent mais aussi l'optimisation de l'étape de fonctionnalisation de surface qui rend possible l'observation à l'échelle de la molécule unique. Ensuite, le chapitre 4 du manuscrit concerne le fonctionnement et les performances de Ruptures, un algorithme de détection de "marches d'escalier" à partir duquel je me suis appuyé pour étudier une grande partie des données expérimentales. J'introduis également le modèle de dynamique d'assemblage sur lequel se basent l'analyse des expériences réalisées à l'équilibre et la simulation des signaux qui permettent d'évaluer les performances de Ruptures. Dans le chapitre 5, je décris les principaux résultats obtenus à l'échelle de la molécule unique. Les analyses présentées concernent les tests de photométrie de masse menés à l'Institut Pasteur, mais surtout les expériences réalisées en microscopie TIRF pour rendre compte de la dynamique d'assemblage de capsides du CCMV à l'équilibre et hors équilibre. Pour terminer, le chapitre 6 me permet de décrire plus en détails le principe de la microscopie à diffusion interférométrique (*Interferometric Scattering* ou iSCAT). J'explique quelles ont été les stratégies que nous avons envisagées afin de coupler cette technique optique au montage TIRF et je présente les premières images obtenues.

# Chapitre - 2

## Les virus

Dans ce chapitre, je commence par une introduction générale sur les virus avant de présenter plus particulièrement celui sur lequel j'ai travaillé pendant trois ans : le CCMV. Dans la dernière partie, un point de vue plus centré sur les techniques expérimentales est adopté via notamment la mise en lumière de différentes méthodes d'observation de virus, que ce soit durant l'assemblage ou une fois la particule formée. Des premiers exemples de représentation sont illustrés par la Figure 2.1.



**FIGURE 2.1** – **Différentes représentations d'un virus.** A) Représentations faites par Biorender.com de trois virus différents : (1) SARS-CoV-2, (2) JC Polyomavirus, (3) Virus Adéno-associé. B) Images en cryo-microscopie électronique de capsides de CCMV ré-assemblées autour de leur ARN, les barres d'échelle font 30 nm, figure adaptée de Panahandeh et al. (2020). C) Suivi d'un assemblage d'une capside de bactériophage MS2 autour de son ARN en microscopie iSCAT, figure adaptée de Garmann et al. (2019).

## 2.1 Généralités

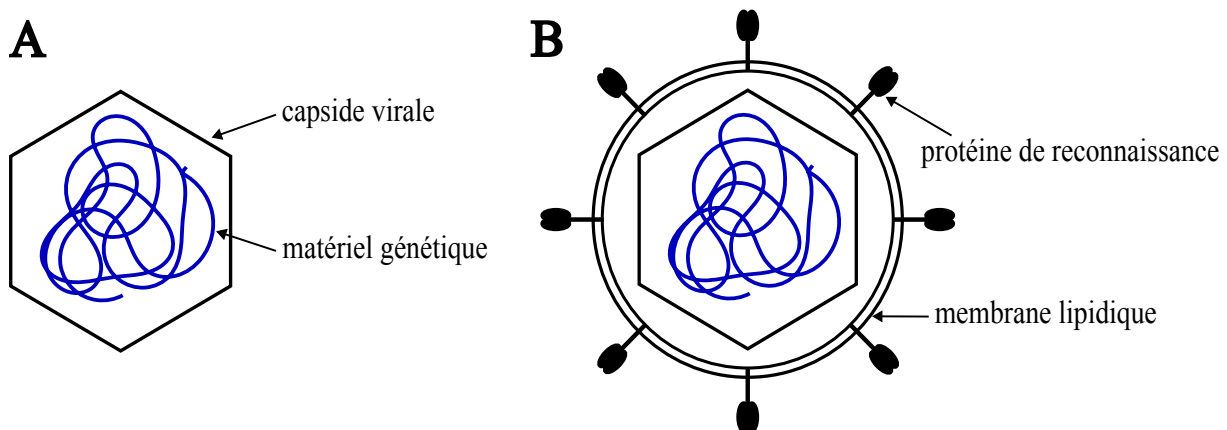
Dans cette première partie, je présente les principales informations qu'il est important d'avoir en tête lorsque l'on parle de virus. De quoi est composé un virus? En existe-t-il différentes familles? Comment un virus se forme-t-il? Telles sont les questions abordées au cours de cette partie.

### 2.1.1 Composition d'un virus

Quand on cherche naïvement dans la littérature le nombre de virus présents sur la planète, la réponse a de quoi surprendre. En effet, près de  $10^{31}$  types de virus différents ont été estimés comme étant présents sur Terre (Dance (2021)). Même si une très faible partie de ces particules peut infecter l'Homme, il peut paraître difficile d'étudier simplement un virus sans avoir de points de repères particuliers. Fort heureusement, il est possible de retrouver plusieurs caractéristiques communes :

- Un matériel génétique de type ADN ou ARN.
- Une capsid virale, composée de plusieurs protéines qui s'associent entre elles afin d'entourer et donc de protéger le matériel génétique.
- Une enveloppe virale, composée généralement d'une membrane lipidique et de protéines virales. Ces dernières sont souvent utilisées par le virus pour reconnaître des cellules cibles et ainsi se fixer à leur surface.

Cette dernière n'est pas toujours présente chez un virus, ce qui permet de faire une première distinction entre les virus dits nus ou non enveloppés (Figure 2.2.A) et les virus enveloppés (Figure 2.2.B).



**FIGURE 2.2** – Schématisations d'un virus. A) Virus nu ou non enveloppé. B) Virus enveloppé.

Par exemple dans la Figure 2.1, seul le virus du SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*) est un virus enveloppé. Les quatre autres virus illustrés, dont fait partie le CCMV, sont des virus non enveloppés. Cette première classification est assez primaire puisque certains virus peuvent être assez complexes avec, par exemple,

des capsides imbriquées les unes dans les autres ou encore un matériel génétique défini par plusieurs segments distincts d'ARN ou d'ADN. D'autres classifications, plus précises, sont utilisées en virologie et c'est ce qui est décrit dans la prochaine section.

### 2.1.2 Comment classer les virus ?

En cherchant dans la littérature, nous pouvons voir que deux classifications, plus officielles que celle que nous avons adoptée précédemment, sont à considérer :

- La classification de Baltimore, qui va s'appuyer sur les différences au niveau du matériel génétique pour distinguer les particules virales (Baltimore (1971)).

- La classification du Comité International de Taxonomie des Virus (ICTV), qui se base sur la même façon de classer les êtres vivants à savoir en famille, genre et espèce. Par ailleurs, il est important de rappeler que la question "*Un virus est-il vivant ?*" continue de susciter des débats en virologie. Cette classification ne sera pas étudiée ici.

Avant de directement s'intéresser aux virus du point de vue du matériel génétique, il est tout à fait possible de remonter d'un échelon et de se positionner à l'échelle de la capside. En effet, comme avec l'enveloppe, les virus peuvent être différenciés selon le type de capside qu'ils possèdent.

#### 2.1.2.1 La capside virale comme classification

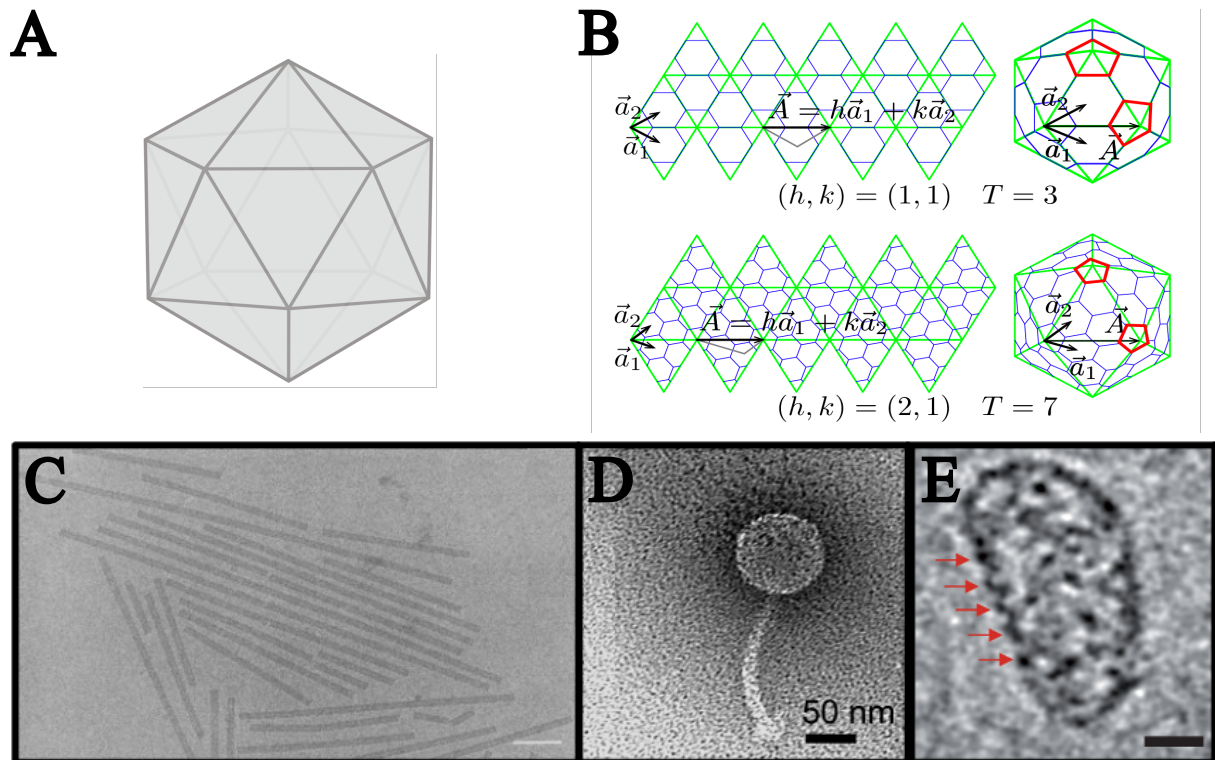
On peut recenser trois grandes catégories de capsides :

- **Les capsides icosaédriques** : un icosaèdre régulier comporte 12 faces triangulaires et 20 sommets (Figure 2.3.A). Les travaux de Caspar et Klug en 1962 ont permis de codifier de manière claire l'agencement des protéines pour des capsides ayant cette symétrie. La capside icosaédrique la plus simple est composée de 60 copies identiques d'une seule protéine présente dans la même conformation. Il est possible d'obtenir des capsides icosaédriques avec un plus grand nombre de protéines mais ces dernières, bien qu'ayant les mêmes propriétés chimiques, auront des conformations légèrement différentes ce qui conduira à des associations différentes au sein de la capside : c'est le principe de quasi-équivalence (Caspar and Klug (1962)). De cette façon, pour qu'une capside soit icosaédrique, il faut qu'elle possède  $60T$  copies d'une protéine avec  $T$ , appelé *nombre de triangulation*, qui correspond au nombre de conformations de cette dernière. Il est important de préciser que  $T$  ne peut pas être choisi au hasard et qu'il peut être calculé en utilisant la relation suivante :

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

où  $h$  et  $k$  sont des entiers qui permettent de caractériser la géométrie des capsides icosaédriques comme l'illustre la Figure 2.3.B. Par ailleurs, cette même figure met en valeur des pentagones et des hexagones comme sous-motifs caractéristiques des capsides icosa-

édriques. En effet, pour former un icosaèdre régulier, les protéines de capsid peuvent s'associer entre elles soit par groupes de cinq et former des pentamères, soit par groupes de six et conduire à des hexamères. De cette façon, une capsid icosaédrique sera composée de 12 pentamères, tous positionnés sur les sommets du solide géométrique (illustrés en rouge sur la Figure 2.3.B), et de  $10(T - 1)$  hexamères. Par exemple, la capsid du CCMV, qui possède une symétrie icosaédrique  $T = 3$ , est formée de 12 pentamères et 20 hexamères, soit 180 monomères.



**FIGURE 2.3** – **Diversité des capsides virales.** A) Icosaèdre régulier obtenu par Biorender.com. B) Illustration du lien entre le nombre de triangulation  $T$  et la géométrie d'une capsid icosaédrique, figure adaptée de Bruinsma and Klug (2015). C) Observation en cryo-microscopie électronique du TMV, barre d'échelle de 100 nm et figure adaptée de Adrian et al. (1998). D) Observation en microscopie électronique du bactériophage 9NA, figure adaptée de Andres et al. (2012). E) Observation en microscopie électronique d'une capsid de HIV-1, barre d'échelle de 20 nm et flèches rouges pointant un réseau d'hexamères, figure adaptée de Zhao et al. (2013).

- **Les capsides hélicoïdales** : ces dernières peuvent être représentées sous forme de tubes ou de bâtonnets. C'est en regardant de plus près l'agencement propre des protéines de capsid qu'il est possible de retrouver une constitution semblable à celle d'une hélice. L'exemple le plus connu est le virus de la mosaïque du tabac (TMV), qui est le premier virus à avoir été découvert à la fin du XIX<sup>e</sup> siècle (Figure 2.3.C).

- **Les capsides mixtes ou doubles** : leur particularité est d'avoir à la fois une capsid de symétrie icosaédrique et hélicoïdale. C'est le cas notamment des phages, comme par exemple le 9NA qui est représenté en Figure 2.3.D.

Bien évidemment, certains virus font figure d'exception et possèdent une capsidie ayant une forme plus exotique. C'est le cas notamment du virus de l'immunodéficience humaine de type 1 (HIV-1) dont la capsidie est plutôt conique comme nous pouvons le voir sur la Figure 2.3.E.

### 2.1.2.2 La classification de Baltimore

En 1971, David Baltimore a proposé une méthode de classification des virus en tenant compte du matériel génétique présent au sein de la particule (ADN ou ARN) mais également des étapes nécessaires à l'obtention de l'ARN messenger. Ce dernier sera ensuite traduit par le ribosome de la cellule hôte afin d'obtenir les protéines de capsidie. La Figure 2.4 illustre les sept groupes qui forment cette classification.

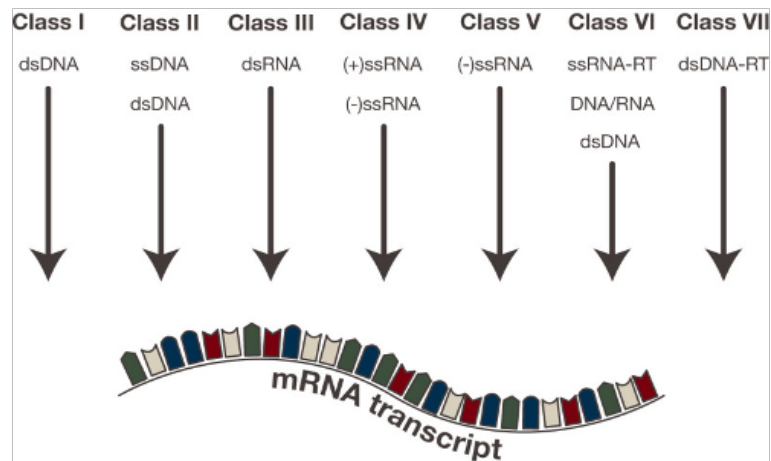


FIGURE 2.4 – Classification de Baltimore. Figure adaptée de Green and Reid (2013).

Si l'on reprend les virus représentés au sein de la Figure 2.1, il ressort que le JC Polyomavirus appartient au groupe I. Ce virus à ADN double brin a besoin que son génome entre dans le noyau de la cellule hôte afin d'être répliqué puis transcrit. Ce groupe de virus est donc généralement dépendant du cycle de division cellulaire de l'hôte. Les virus adéno-associés sont eux des virus à ADN simple brin et font partie du groupe II. Ils ont besoin de la formation d'un ADN double brin intermédiaire avant d'être transcrits. Enfin, le SARS-CoV-2, le MS2 ainsi que le CCMV appartiennent au groupe IV et possèdent donc un ARN simple brin de polarité positive qui peut être directement traduit par le ribosome de la cellule hôte.

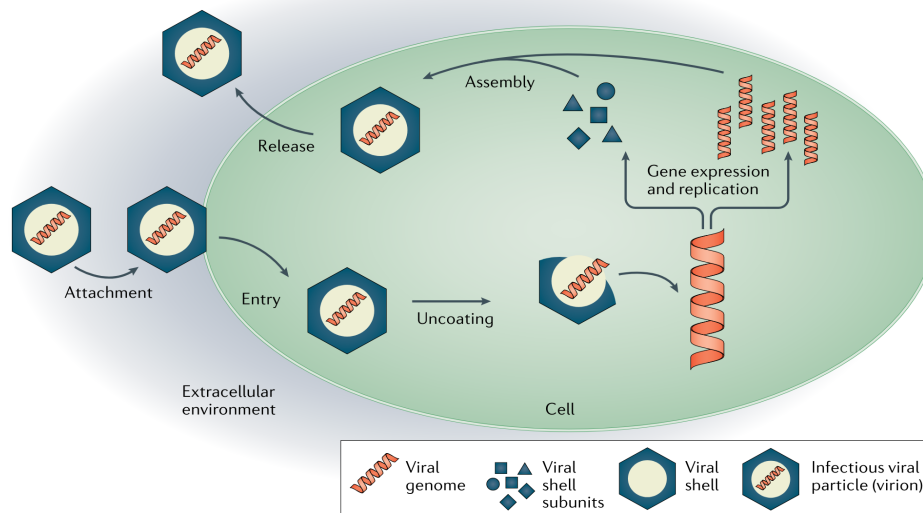
### 2.1.3 Cycle viral

La présentation de la classification de Baltimore met en lumière une caractéristique bien connue des virus qui est de détourner la machinerie de répllication de la cellule hôte. Ce processus est primordial pour qu'un virus puisse proliférer et infecter de nouvelles cellules de son organisme hôte. Bien que la classification précédente ait mis en valeur la



diversité des étapes mises en jeu dans l'expression des protéines de capsid, le cycle viral suit un cheminement assez général illustré par la Figure 2.5. Dans un premier temps, je vais faire une rapide description de chacune des étapes. Je reviendrai plus en détails sur les principaux mécanismes d'assemblage connus pour une capsid virale dans la sous-partie suivante.

Les étapes décrites ci-dessous retracent le parcours d'une particule virale depuis son entrée dans une cellule hôte jusqu'à la sortie de plusieurs de ses copies :



**FIGURE 2.5 – Schématisation du cycle viral.** Cette représentation, adaptée de Bruinsma et al. (2021), rend compte des différentes étapes du cycle viral d'une manière générale.

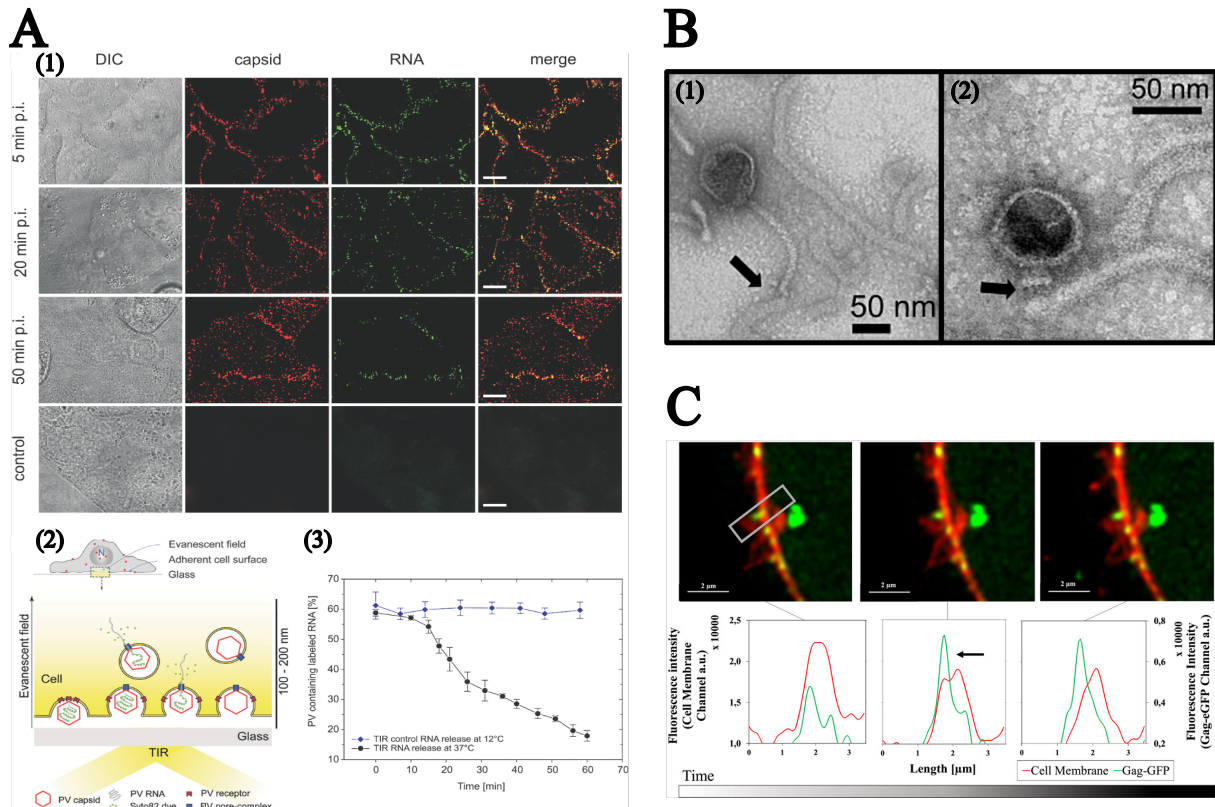
- **L'adsorption à la cellule hôte** : une particule virale, qu'elle soit nue ou enveloppée, va venir se fixer au niveau de certains récepteurs de la cellule prochainement infectée. Par exemple, le HIV-1, qui est un virus enveloppé, est connu pour se fixer au niveau des récepteurs CD4 présents sur les lymphocytes T.

- **La pénétration dans la cellule hôte** : une fois fixé à la membrane de la cellule, un virus va pouvoir entrer au sein de cette dernière selon deux mécanismes principaux. Les virus nus vont être incorporés à la cellule par endocytose, c'est-à-dire en formant une vésicule à partir de la membrane cellulaire. Pour les virus enveloppés, une fusion sera opérée entre leur propre enveloppe et la membrane de la cellule hôte afin de libérer la capsid dans le cytoplasme. À noter que les phages sont connus pour perforer la membrane de leur hôte afin de seulement y injecter leur ADN viral. La Figure 2.6.B montre deux phages appartenant à des familles différentes accrochés à une bactérie de salmonelle.

- **La décapSIDation** : la capsid est dégradée, dans la majorité des cas dans le cytoplasme, afin de permettre la libération du matériel génétique viral. L'équipe de Brandenburg et al. (2007) a mené une étude sur l'entrée de polyovirus dans des cellules HeLa. Ils ont mis en lumière que la libération de l'ARN se produit rapidement après l'international-

lisation des virus, à une distance comprise entre 100 et 200 nm de la membrane plasmique (Figure 2.6.A).

- **La réplication du matériel génétique et la traduction des protéines** : cette étape acte la prise en otage du mécanisme de réplication de la cellule hôte par le virus. La réplication virale peut avoir lieu dans le noyau ou dans le cytoplasme selon le matériel génétique du virus. À ce stade, plusieurs copies du génome et des protéines virales sont obtenues.



**FIGURE 2.6 – Observation de plusieurs étapes du cycle viral.** A) Mise en lumière de la libération de l'ARN de polyovirus (PV) dans des cellules HeLa en microscopie de fluorescence : (1) Détection en microscopie TIRF des capsides (en rouge) et de l'ARN (en vert), barres d'échelle de 10 μm, (2) Schéma du montage expérimental, (3) Évolution des colocalisations entre ARN et capsides au cours du temps, figure adaptée de Brandenburg et al. (2007). B) Illustration de l'accroche (flèches noires) de phages 9NA (1) et P22 (2) sur des bactéries de salmonelle en microscopie électronique à coloration négative, figure adaptée de Andres et al. (2012). C) Visualisation en temps réel du bourgeonnement d'une capside de HIV-1 (encadré blanc) en microscopie super-résolue, barres d'échelle de 2 μm, figure adaptée de González-Domínguez et al. (2020).

- **L'assemblage de la capside** : certaines protéines du virus s'agencent autour du matériel génétique afin de former la capside et préparer la sortie de la cellule hôte.

- **La sortie de la cellule hôte** : les centaines de copies de la particule virale quittent la cellule principalement selon deux mécanismes. Les virus nus seront généralement libérés lors de la lyse cellulaire, c'est-à-dire lors de la destruction de la membrane de la cellule. Pour les virus enveloppés, divers mécanismes existent pour retrouver l'enveloppe. Certains

simples, dus notamment à une concentration élevée en protéines d'enveloppe, d'autres plus complexes, nécessitant une réorganisation de la membrane où se retrouvent les différentes protéines (de capsidite et d'enveloppe). Le bourgeonnement, *i.e.* la sortie, des particules nouvellement enveloppées, est illustré dans le cas du HIV-1 sur la Figure 2.6.C.

#### 2.1.4 Mécanismes d'assemblage connus de la capsidite virale

Le matériel génétique d'un virus ne pouvant pas survivre seul en dehors de la cellule hôte, l'assemblage de la capsidite s'avère être une étape cruciale dans le cycle viral. Mais alors que l'entrée et la sortie de plusieurs virus ont déjà pu être étudiées *in vivo* (Figure 2.6), faire de même avec l'assemblage des protéines de capsidite autour du matériel génétique s'avère être très compliqué à réaliser à l'heure actuelle. Fort heureusement, des études théoriques mais également de nombreuses expériences réalisées *in vitro* ont pu être menées afin d'en apprendre davantage sur les mécanismes mis en jeu lors de l'assemblage des capsidites. Les conclusions obtenues sont souvent différentes d'un virus à un autre, voire pour un même spécimen selon les conditions expérimentales choisies, mais des idées directrices ont pu émerger de ces travaux.

Dans la suite de cette sous-partie, je vais m'intéresser aux capsidites formées selon un processus dit d'auto-assemblage. Ce terme a été illustré pour la première fois en 1955 quand Heinz Fraenkel-Conrat et Robley C. Williams ont observé la formation spontanée d'un virus de TMV infectieux en mélangeant *in vitro* ses protéines de capsidite et son ARN simple brin (Fraenkel-Conrat and Williams (1955)). Après cette découverte, plusieurs études ont montré que d'autres virus avaient aussi la capacité de s'auto-assembler *in vitro*. C'est le cas notamment de virus ayant une capsidite icosaédrique comme le virus simien 40 (SV40, Tsukamoto et al. (2007)), les phages MS2 et  $Q\beta$  (Ling et al. (1969)) mais également le CCMV (Bancroft and Hiebert (1967), Adolph and Butler (1976)), et de deux autres virus du même genre, ceux de la mosaïque du brome (BMV) et de la marbrure des fèves (BBMV) (Bancroft (1970)). C'est donc grâce à ces travaux initiaux que de nouvelles investigations ont pu être menées afin d'étudier, par exemple, les cinétiques d'assemblage des capsidites et de voir à quel point ces dernières dépendent des conditions expérimentales.

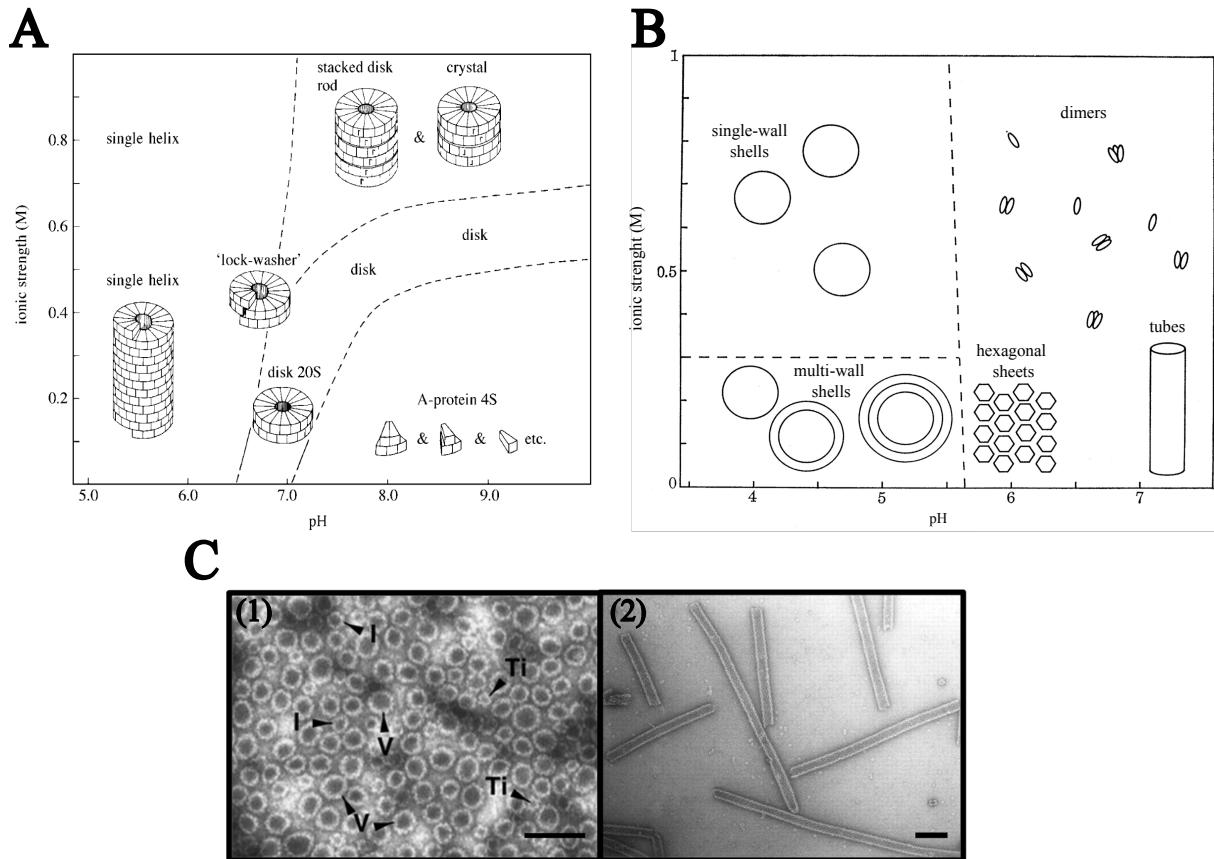
Je vais désormais présenter les principaux mécanismes d'assemblage connus pour des capsidites avec et sans matériel génétique.

##### 2.1.4.1 Formation de capsidites vides

- Influence du pH et de la salinité :

Alors que le principe même d'une capsidite virale est de se former autour de son génome afin de protéger ce dernier, plusieurs équipes ont observé que des sous-unités de capsidite peuvent s'auto-assembler entre elles, sans matériel génétique. Dans ces conditions, seules les interactions entre protéines permettent la formation de la capsidite. Généralement,

une compétition est observée entre des interactions d'origines répulsives à longue portée, facilitant ainsi le désassemblage du virus, mais également des interactions attractives à courte distance, permettant l'assemblage et le maintien de la structure de la capside. Dans la section 2.2.1, je précise l'origine de ces interactions dans le cas du CCMV. Plus globalement, les recherches menées sur plusieurs virus ont permis de montrer que les conditions de pH et salinité vont perturber ces interactions et ainsi modifier la forme des objets auto-assemblés (Figure 2.7).



**FIGURE 2.7** – Influence du pH et de la salinité sur la structure des capsides vides. A) Diagramme de phase du TMV, figure adaptée de Klug (1999). B) Diagramme de phase du CCMV, figure adaptée de Adolph and Butler (1976). C) Structures de capsides vides de SV40 observées en microscopie électronique (barres d'échelle de 100 nm) : (1) Majorité de capsides de tailles semblables aux virions natifs (pH = 7,2 et 2 M  $(\text{NH}_4)_2\text{SO}_4$ , 2 mM  $\text{CaCl}_2$  et V = capside semblable à un virion, I = structure intermédiaire, Ti = petite capside), (2) Longues structures tubulaires (pH = 5 et 150 mM NaCl, 2 mM  $\text{CaCl}_2$ ), figure adaptée de Kanesashi et al. (2003).

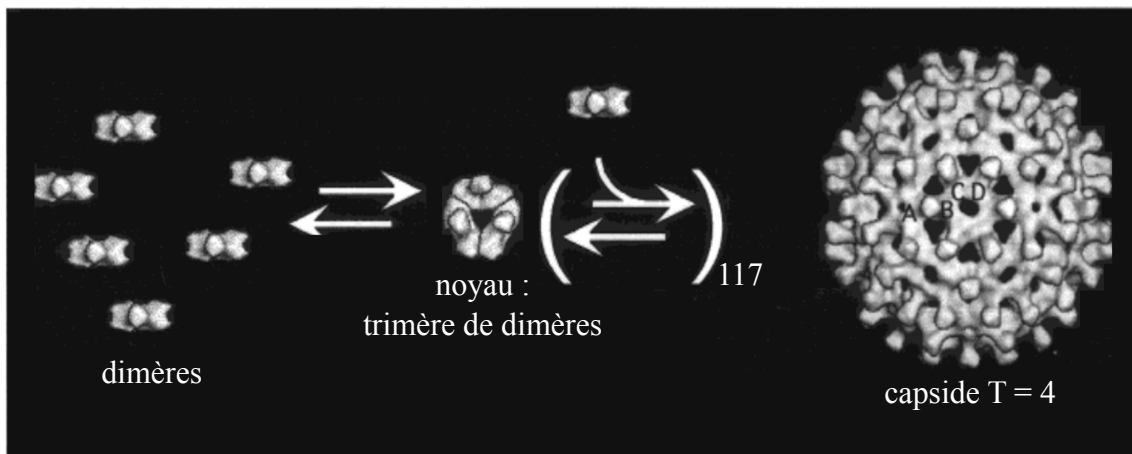
En effet, Klug (1999) a établi le diagramme de phase du TMV (Figure 2.7.A) et a montré que pour avoir des capsides hélicoïdales proches de virions natifs, il faut travailler à faible pH et forte salinité. En revanche, en se plaçant dans des conditions physiologiques (pH = 7,2 et une force ionique de 0.1 M), il est uniquement possible d'obtenir des oligomères ou deux disques de protéines l'un sur l'autre. Selon le même principe, Adolph and Butler (1976) ont démontré que des capsides vides de CCMV se forment à faible

pH et forte salinité (Figure 2.7.B) mais qu'en conditions physiologiques, on favorisera la formation de tubes. Par ailleurs, le CCMV possède 180 copies de la même protéine mais Speir et al. (1995) ont souligné que la sous-unité de capsidie la plus stable en solution est le dimère de protéines du fait des fortes interactions entre monomères adjacents. Enfin, Kaneshashi et al. (2003) ont observé qu'à pH physiologique mais forte salinité (pH = 7,2 et 2 M  $(\text{NH}_4)_2\text{SO}_4$ , 2 mM  $\text{CaCl}_2$ ), des capsides de SV40 semblables aux virions natifs sont obtenues mais qu'à pH acide et salinité physiologique (pH = 5 et 150 mM NaCl, 2 mM  $\text{CaCl}_2$ ), de longues structures tubulaires sont formées (Figure 2.7.C.1 et 2 respectivement).

Obtenir des capsides vides est donc tout à fait possible mais préférentiellement en se plaçant dans des conditions extrêmes, assez éloignées des conditions physiologiques. L'explication avancée pour décrire ce phénomène serait qu'en travaillant à pH acide et forte salinité, les interactions électrostatiques entre sous-unités de capsides sont réduites, facilitant l'auto-assemblage (Bruinsma et al. (2021)).

- Mécanisme de nucléation-croissance :

En plus de démontrer l'influence du pH et de la salinité sur la formation de capsides vides, il a également été possible d'identifier un chemin d'auto-assemblage particulier : le mécanisme de nucléation-croissance. Comme son nom l'indique, l'assemblage se déclenche une fois qu'un complexe de protéines d'une certaine taille (noyau) est obtenu, avant de se poursuivre par l'ajout progressif de sous-unités.

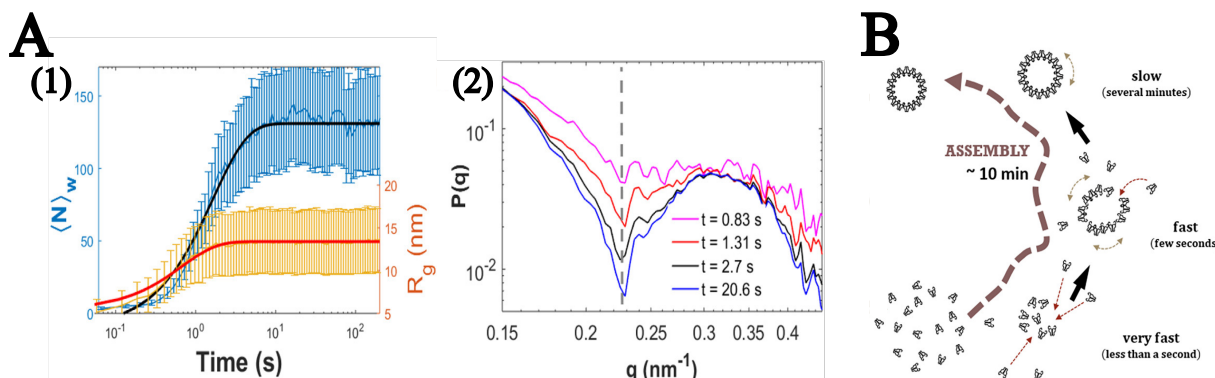


**FIGURE 2.8** – Chemins d'assemblage proposés pour le virus de l'hépatite B. Mécanisme de nucléation-croissance pour HBV où 117 correspond au nombre de dimères à ajouter au noyau trimérique pour former une capsidie  $T = 4$  et A,B,C,D correspondent aux quatre conformations des protéines de HBV, figure adaptée de Zlotnick et al. (1999).

Un premier modèle développé par Zlotnick et al. (1999) a permis de caractériser la nature du noyau critique dans le cas du virus de l'hépatite B (HBV), dont la capsidie a la particularité d'être polydisperse : la grande majorité des objets ont une symétrie  $T = 4$ , mais il est également possible d'observer des capsides de symétrie  $T = 3$ . En

effet, en se basant sur un noyau de type trimère de dimères puis un ajout progressif de dimères (Figure 2.8), les simulations effectuées pour différentes concentrations en sel et en protéines ont confirmé les observations expérimentales.

Casini et al. (2004), en se basant sur le modèle décrit par Zlotnick, ont pu estimer la nature du noyau critique du papillomavirus humain (HPV) à un dimère de pentamères. Cette estimation paraît être cohérente avec la structure de la capside de HPV puisqu'elle possède la particularité, comme celle de SV40, d'être composée exclusivement de pentamères, au nombre ici de 72. En effet, les familles des papillomavirus et polyomavirus (dont fait partie SV40) sont caractérisées par des capsides ayant une symétrie icosaédrique  $T = 7$  mais sans respecter exactement les règles énoncées par Casper et Klug puisqu'on ne retrouve pas d'hexamères. Par ailleurs, les mesures expérimentales réalisées par Casini et al. (2004) ont également confirmé le mécanisme de nucléation-croissance menant à des capsides vides de HPV. De plus, une réduction de la durée avant l'initiation de l'assemblage tout comme une augmentation de la proportion de capsides entièrement formées ont pu être observées pour une concentration croissante en protéines. Ces deux dernières observations avaient déjà été mises en lumière par Zlotnick et al. (1999) pour HBV.



**FIGURE 2.9** – Observations expérimentales caractérisant l'auto-assemblage de HBV. A) Données obtenues en TR-SAXS pour l'assemblage de 0,6 g/L de dimères de HBV dans 1 M NaCl à pH = 9 : (1) Évolution de la taille  $R_g$  et de la masse  $\langle N \rangle_w$  des objets au cours du temps, (2) Évolution du facteur de forme  $P(q)$  traduisant la structuration des objets vers des capsides icosaédriques  $T = 4$ . B) Schématisation du mécanisme de nucléation-croissance-relaxation observé pour HBV. Figures adaptées de Chevreuil et al. (2020).

Le mécanisme de nucléation-croissance prévaut donc bien pour la formation de capsides vides et notamment pour HBV. Cela étant, plusieurs précisions ont été apportées sur ce processus et ce notamment en ce qui concerne les événements ayant lieu en fin d'assemblage. En effet, Chevreuil et al. (2020) ont montré qu'en plus des phases de nucléation et de croissance, une troisième phase dite de relaxation permet de tendre vers des capsides complètes (Figure 2.9).B. Ces trois phases sont illustrées par des données expérimentales acquises en diffusion de rayons X aux petits angles résolue en temps (TR-SAXS, Figure

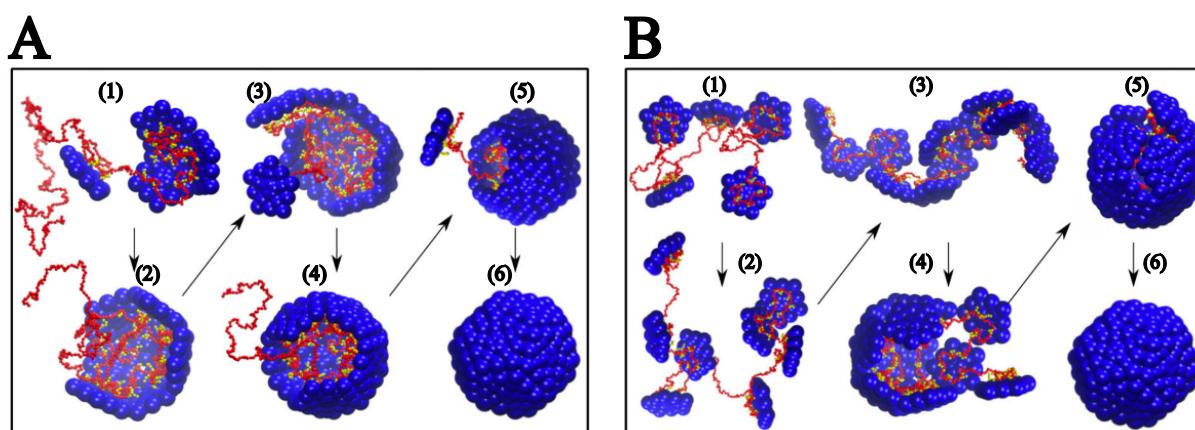
2.9.A). La Figure 2.9.A.1 montre d’abord qu’à la fois la taille  $R_g$  et la masse  $\langle N \rangle_W$  des objets augmentent conjointement, synonyme d’une phase de nucléation en sous-unités. Ensuite, la taille des objets se stabilise alors que la masse continue d’augmenter : nous sommes donc dans la phase de croissance où des dimères de protéines continuent de participer à la formation de la capsid. Enfin, lorsqu’à la fois la masse et la taille des objets n’évoluent plus, la Figure 2.9.A.2 nous montre que même après 20 secondes d’acquisition, le facteur de forme  $P(q)$  continue d’évoluer vers des arches de plus en plus creusées, synonyme d’une structure qui tend à se rapprocher progressivement d’une capsid icosaédrique de type  $T = 4$ . Cette phase de relaxation correspond donc à une lente restructuration de la capsid vers sa forme finale.

Ce mécanisme en trois étapes a été obtenu pour plusieurs conditions expérimentales mais Chevreuil et al. (2020) ont observé qu’en cas de forte concentration en sel, la phase de relaxation est absente, ce qui peut signifier que les structures finales possèdent des défauts, sans possibilité de se restructurer. Ce phénomène qui avait déjà été modélisé, notamment par Hagan et al. (2011), et constaté expérimentalement par Lutomski et al. (2018), porte le nom de piège cinétique. L’origine de ce processus a été décrite comme suit : une forte salinité diminuerait de manière trop importante les répulsions entre sous-unités, conduisant à une trop forte attraction entre ces dernières et à une grande difficulté de réarrangement. C’est également en se basant sur les interactions attractives entre sous-unités que Panahandeh et al. (2020) ont généralisé ce mécanisme à l’auto-assemblage de capsides icosaédriques vides de symétrie  $T = 3$ . En présence de trop fortes attractions entre sous-unités, les objets obtenus sont cinétiquement piégés sous forme de structures aberrantes, alors qu’en diminuant les interactions attractives, un mécanisme de nucléation-croissance permet de former des capsides de symétrie icosaédrique  $T = 3$  comme le CCMV. Les auteurs ont également précisé que les erreurs d’assemblage visibles sur les structures aberrantes peuvent se corriger grâce au franchissement d’une barrière d’énergie, mais seulement au bout d’un certain temps.

#### 2.1.4.2 Formation de capsides pleines

Les recherches menées sur des capsides vides ont permis de travailler avec des objets relativement simples et pour lesquels les interactions mises en jeu au cours de l’assemblage sont facilement modulables à l’aide du pH et de la concentration en sel. Néanmoins, il apparaît important d’inclure le matériel génétique dans les modèles et expériences d’assemblage si l’on souhaite comprendre les mécanismes à l’origine de la formation de capsides pleines et donc de virus infectieux. En ajoutant une longue séquence de nucléotides chargée négativement, cela modifie forcément le paysage des interactions mises en jeu. C’est le cas notamment des protéines de SV40 qui, en conditions physiologiques (150 mM NaCl à pH = 7,2), ne forment que des agrégats sans structure particulière alors qu’avec leur

ADN double brin, des capsides de 40 nm sont majoritairement formées (Tsukamoto et al. (2007)). En comparaison, le CCMV suit à peu près le même comportement : le diagramme de phase de la Figure 2.7.B indique que les capsides vides ne sont plus présentes au-delà d'un pH de 5,5 et une force ionique inférieure à 0,3 mais Adolph and Butler (1976) ont pu observer de nombreuses capsides semblables aux virions natifs en présence d'ARN simple brin pour une force ionique de 0,2 et un pH de 6,2. En prenant en compte ces observations, il est possible de parler d'assemblage coopératif entre les protéines de capsidite et le génome viral, dans lequel le matériel génétique facilite la formation de l'enveloppe protéique.



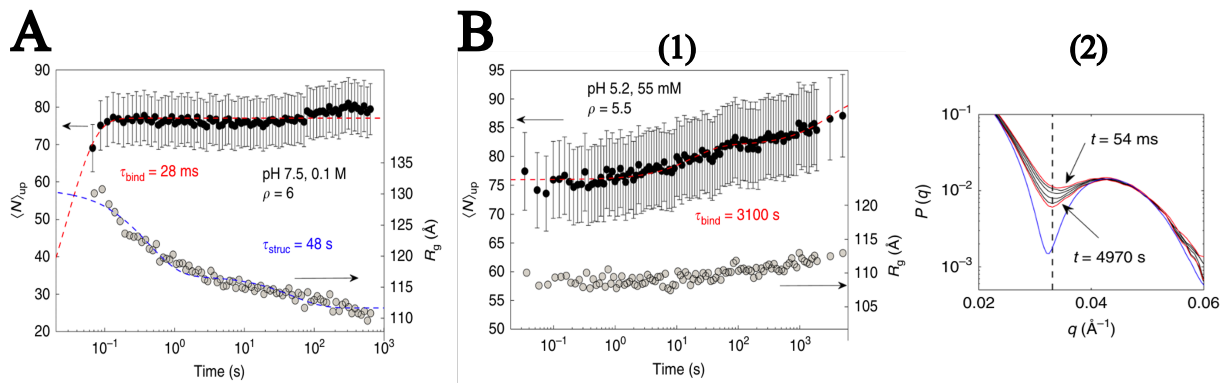
**FIGURE 2.10** – Deux voies d'assemblage pour un mélange protéines de capsidite et matériel génétique. A) Mécanisme d'assemblage de nucléation-croissance. B) Mécanisme d'assemblage dit en masse. Figures adaptées de Perlmutter et al. (2014).

Pour le mécanisme d'assemblage, il est toujours possible de parler de nucléation-croissance comme cela a été observé pour la formation de capsides de MS2 en présence de son ARN simple brin (Garmann et al. (2019)). En effet, plus la concentration en protéines augmente, plus la durée avant l'initiation de l'assemblage se réduit, ce qui indique bien une diminution de la barrière d'énergie d'activation à dépasser pour enclencher l'assemblage. Une nucléation de protéines est donc toujours possible sauf que cette fois-ci, elle peut se produire directement au contact du génome viral (Figure 2.10.A). Toutefois, le matériel génétique modifiant le comportement des protéines, un autre processus d'assemblage a pu être identifié.

- Mécanisme d'assemblage dit en masse :

Grâce à la présence du matériel génétique du virus, les protéines peuvent venir interagir avec ce dernier à plusieurs endroits à la fois. Il est donc possible de considérer un assemblage désordonné, au cours duquel les protéines se fixent non spécifiquement au matériel génétique avant de progressivement s'organiser et conduire à une capsidite complète. Cette voie d'assemblage est représentée en Figure 2.10.B.





**FIGURE 2.11** – Données TR-SAXS caractérisant l’assemblage des sous-unités de CCMV autour de son ARN viral. A) Évolution de la taille  $R_g$  et de la masse  $\langle N \rangle_{up}$  des objets au cours du temps pour un assemblage réalisé avec un rapport en masse protéines/ARN  $\rho$  de 6 dans une solution de force ionique 0.1 M et un pH de 7,5. B) Passage des objets pré-assemblés dans une solution de force ionique 55 mM et un pH de 5,2 ( $\rho$  vaut 5,5) : (1) Évolution de la taille  $R_g$  et de la masse  $\langle N \rangle_{up}$  des objets au cours du temps, (2) Évolution du facteur de forme  $P(q)$  (de  $t = 54$  ms à 4970 s) traduisant la structuration des objets vers des capsides icosaédriques  $T = 3$  semblables aux virions purifiés (courbe bleue). Figures adaptées de Chevreuil et al. (2018).

Les expériences menées par Chevreuil et al. (2018) ont corroboré cette voie d’assemblage pour la capsid du CCMV comme cela est illustré au sein de la Figure 2.11. En effet, en se plaçant dans un tampon favorisant les interactions entre protéines et matériel génétique (force ionique de 0,1 M mais un pH légèrement supérieur à la neutralité), les données acquises en TR-SAXS ont montré une rapide augmentation du nombre moyen de dimères par ARN autour de 75 unités (Figure 2.11.A). Ensuite, la taille des objets diminue plus lentement et tend à se stabiliser au bout de plusieurs minutes alors que le nombre de dimères reste constant. Ces deux premières phases sont parfaitement illustrées respectivement par les étapes (1) - (2) puis (3) - (4) de la Figure 2.10.B : un complexe désordonné est rapidement formé avant d’observer une lente structuration de la capsid. Enfin, la Figure 2.11.B.1 montre qu’une acidification du milieu, *i.e.* une augmentation des interactions entre sous-unités, permet d’augmenter le nombre de dimères par objet pré-assemblé de 75 à 85-90, soit la taille réelle d’une capsid de CCMV. Durant cette troisième phase, la taille des objets évolue très peu mais une lente structuration de la capsid vers une symétrie icosaédrique  $T = 3$  a lieu comme l’indique l’évolution du facteur de forme vers celui de virions natifs du CCMV (Figure 2.11.B.2). Cette dernière phase de réarrangement correspond bien aux étapes (5) - (6) du mécanisme d’assemblage en masse. Ces expériences ont donc mis en valeur le rôle des interactions protéines - protéines et protéines - matériel génétique étudiées par Perlmutter et al. (2014).

La nécessité d’une modification en deux étapes de la force ionique et du pH de la solution afin d’obtenir des capsides de CCMV complètes avait déjà été mentionnée par Garmann et al. (2014). Les travaux effectués par cette équipe ont été réalisés en assemblant des dimères de protéines de CCMV autour de l’ARN viral du BMV afin d’exclure de

potentielles interactions spécifiques entre protéines et génome provenant du même virus. Une preuve supplémentaire que l'auto-assemblage de la capsid de CCMV n'est pas régi par des interactions propres entre ses protéines et son ARN viral a été donnée par Comas-Garcia et al. (2012). En effet, ils ont montré que l'assemblage des dimères de CCMV était plus efficace autour de l'ARN viral du BMV plutôt qu'avec son propre matériel génétique. Cette découverte semble indiquer que la structure secondaire de l'ARN simple brin influence l'empaquetage de ce dernier lorsque des interactions non spécifiques entre protéines et matériel génétique sont mises en jeu.

Enfin, comme pour le mécanisme de nucléation-croissance, la concentration en protéines joue un rôle dans la formation de capsides complètes. Plus particulièrement, c'est le rapport entre la concentration en protéines et celle en matériel génétique qu'il faut surveiller. D'après Cadena-Nava et al. (2019), un rapport en masse de 4,3<sup>1</sup> est requis pour former des capsides de CCMV, traduisant ainsi une sur-stoechiométrie en faveur des protéines.

#### - Influence des "packaging signals" :

Les expériences réalisées sur le CCMV ont mis en lumière l'absence d'interactions spécifiques entre ses protéines et son matériel génétique pouvant faciliter l'auto-assemblage. Chevreuil et al. (2018) ont d'ailleurs montré que la capsid de CCMV pouvait s'assembler autour d'une chaîne synthétique de polyélectrolyte chargée négativement avec une dynamique comparable à celle illustrée sur la Figure 2.11.A. Dans ce cas, la formation de capsides pleines se base principalement sur des interactions électrostatiques.

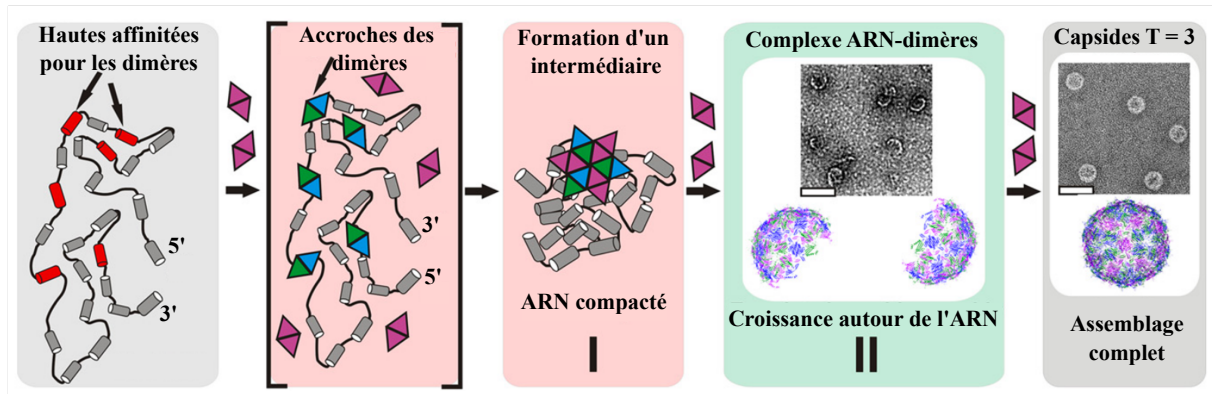
Cependant, d'autres virus sont connus pour s'assembler à partir de *packaging signals*, synonymes d'interactions spécifiques entre des séquences du matériel génétique et les protéines. C'est notamment le cas du virus du MS2 que j'ai présenté plus haut, pour lequel Garmann et al. (2019) ont observé un assemblage caractéristique du mécanisme de nucléation-croissance. En réalité, la formation du noyau critique de dimères pourrait provenir d'interactions spécifiques entre l'ARN et les dimères.

En effet, Borodavka et al. (2012) ont montré que le rayon apparent de l'ARN viral de MS2 subit une forte diminution directement après l'ajout des dimères de capsid avant d'augmenter progressivement au fur et à mesure de l'assemblage. Ces observations n'ont pas été faites quand l'ARN de MS2 a été remplacé par un ARN viral d'un autre virus ou des ARN non viraux. De plus, seuls les milieux contenant l'ARN et les sous-unités de MS2 ont donné des capsides icosaédriques complètes de symétrie  $T = 3$ . Borodavka et al. (2012) parlent donc d'un mécanisme d'assemblage initié en partie par des sites d'interactions spécifiques entre l'ARN de MS2 et ses dimères (Figure 2.12). L'assemblage commencerait donc par l'accroche de plusieurs sous-unités de MS2 à des positions spécifiques de l'ARN,

---

1. Le rapport en concentration est donc d'environ 188 en faveur des protéines.

créant un effondrement de ce dernier et donc un rapprochement des différents dimères présents à ce stade. Ces interactions pourraient donc amener à la formation du noyau de nucléation mentionné par Garmann et al. (2019) dans leurs travaux. Ensuite, la capsid va progressivement croître au contact des dimères libres en solution, jusqu'à se former entièrement.



**FIGURE 2.12** – **Modèle d'un mécanisme d'assemblage contrôlé par des interactions spécifiques entre ARN et protéines.** Les deux derniers encadrés sont chacun illustrés par une image prise en microscopie électronique. Figure adaptée de Borodavka et al. (2012).

Plus tard, de nouveaux éléments ont été apportés par Rolfsson et al. (2016), confirmant ces interactions particulières. En effet, cette équipe a pu établir les séquences peptidiques et nucléotidiques aux interfaces entre l'ARN et les dimères. Ils se sont aperçus que les mêmes interfaces pouvaient être identifiées sur les différentes particules virales étudiées. Enfin, les positions de certaines de ces interfaces correspondent aux prédictions de localisation faites par Dykeman et al. (2013). Ces différents résultats montrent encore plus le rôle prépondérant que le matériel génétique peut avoir lors de la formation de la capsid virale.

### 2.1.5 Conclusion

Cette partie introductive a permis de rendre compte des informations nécessaires à assimiler lorsque l'on s'intéresse aux virus au sens large. Nous avons pu voir que ces derniers peuvent être différenciés par plusieurs aspects de leur structure (enveloppe, capsid, matériel génétique), facilitant ainsi leur classification. De plus, nous sommes revenus sur le cycle viral dans sa globalité, en mettant l'accent sur l'étape qui nous intéresse le plus, à savoir le mécanisme d'auto-assemblage. Les différentes interactions entre les protéines de capsid et le matériel génétique ont permis d'identifier plusieurs voies d'assemblage désormais assez bien reconnues mais qui demandent encore à être approfondies. Par ailleurs, je me suis principalement intéressé à l'assemblage de virus non enveloppés ayant un ARN simple brin mais les virus présentant d'autres structures sont également documentés. Dé-

sormais, je vais me concentrer plus particulièrement sur le CCMV.

## 2.2 Le CCMV

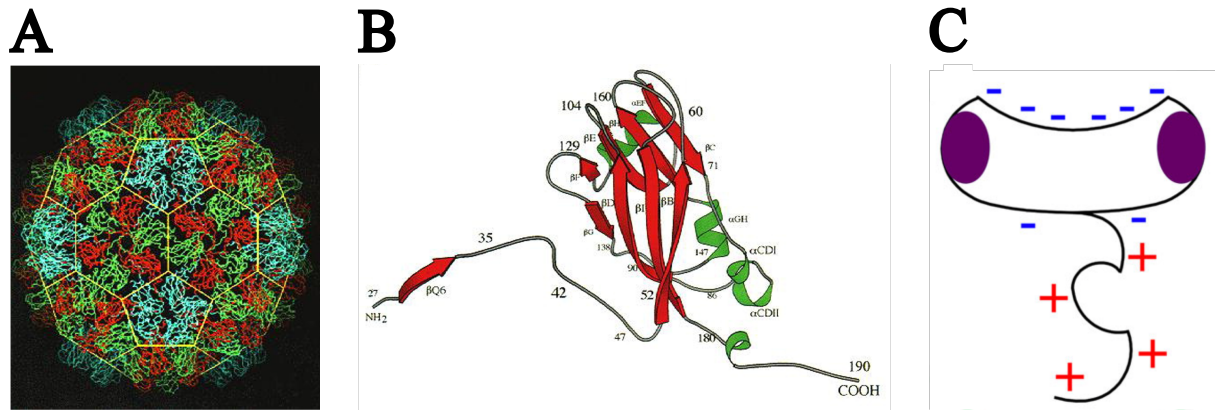
Le virus de la marbrure chlorotique de la cornille (*Cowpea Chlorotic Mottle Virus* ou CCMV) a d'ores et déjà été étudié à plusieurs reprises depuis sa première caractérisation dans les années 60. En effet, ce virus de plante qui est connu pour infecter une espèce de haricot des pays d'Amérique du Sud, est assez simple en terme de composition puisqu'il s'agit d'un virus non enveloppé, ne possédant donc qu'une capsidie associée autour de son matériel génétique. De plus, il constitue un parfait candidat d'étude en laboratoire de part sa propension à s'auto-assembler facilement. En effet, comme je l'ai souligné durant la partie précédente, il peut aussi bien s'assembler avec mais également sans matériel génétique et son comportement est facilement modulable au travers de la force ionique du milieu et du pH. Dans cette seconde partie, je vais revenir plus en détails sur la structure de ce virus sans m'arrêter sur son mécanisme d'assemblage. Les protocoles que nous utilisons au LPS afin de purifier les virions et protéines de CCMV et synthétiser l'ARN simple brin sont décrits aux annexes A et B.

### 2.2.1 La capsidie protéique

La capsidie virale du CCMV est caractérisée par une symétrie icosaédrique  $T = 3$  et mesure 28 nm de diamètre à pH = 4,8 (Speir et al. (1995)). En se basant sur les règles énoncées par Caspar et Klug, la capsidie du CCMV est composée de 180 copies de la même protéine réparties en 12 pentamères et 20 hexamères. Les protéines sont identiques d'un point de vue chimique mais possèdent trois conformations différentes notées A, B et C, comme représentées respectivement en bleu, rouge et vert sur la Figure 2.13.A. Les monomères A construisent les pentamères et les monomères B et C permettent la formation des hexamères. Par ailleurs, comme je l'avais mentionné plus tôt, Speir et al. (1995) ont souligné les fortes interactions entre certaines associations de monomères, conduisant les dimères de CCMV à être la sous-unité la plus stable en solution. Enfin, prise dans son ensemble et sans tenir compte du matériel génétique, la masse molaire de la capsidie se situe aux alentours de 3,6 MDa en considérant une masse par protéine d'environ 20 kDa.

En outre, si l'on souhaite comprendre les interactions régissant la formation de la capsidie et son maintien au cours du temps, il est intéressant de se pencher sur la structure de la protéine en elle-même. En effet, chaque monomère possède une séquence de 190 acides aminés en partant du domaine dit N-terminal situé à l'intérieur de la capsidie, au domaine C-terminal positionné à l'extérieur de la nanoparticule (Figure 2.13.B). La partie N-terminale est caractérisée par 25 acides aminés dont 10 d'entre eux portent une charge positive (Figure 2.13.C). Cette extrémité chargée positivement va donc préférentiellement

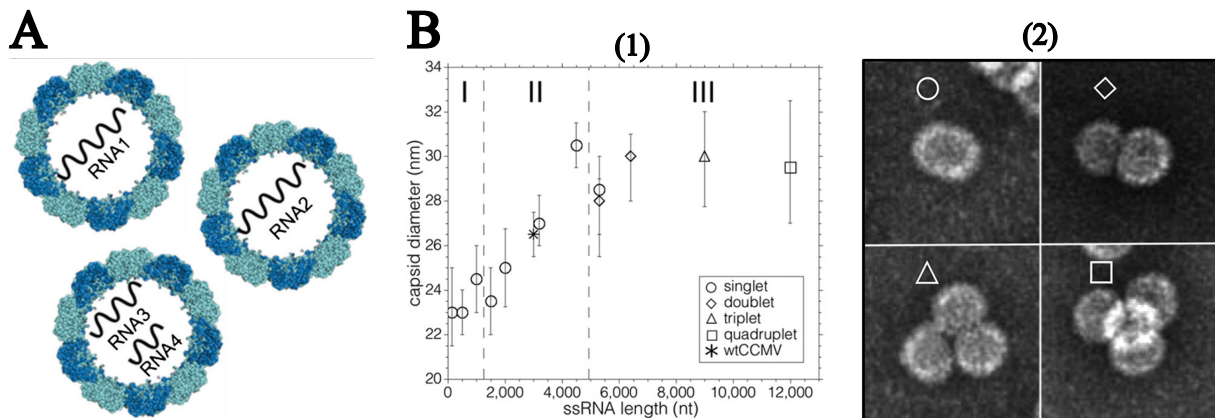
interagir avec le matériel génétique qui est chargé négativement. La Figure 2.13.C illustre également les deux autres interactions que j’avais mentionnées en section 2.1.4.1 : celles d’origine répulsive provenant des groupements carboxyles chargés négativement et les interactions attractives issues des sites hydrophobes de la protéine.



**FIGURE 2.13** – Structures de la capsid et des protéines de CCMV. A) Symétrie icosaédrique  $T = 3$  du CCMV avec les monomères A, B et C représentés respectivement en bleu, rouge et vert, figure adaptée de Speir et al. (1995). B) Diagramme en ruban de la structure tertiaire d’une protéine de CCMV, figure adaptée de Speir et al. (1995). C) Représentation schématique d’une protéine de CCMV avec les charges positives du domaine N-terminal, les charges négatives des groupements carboxyles et les sites hydrophobes illustrés en violet, figure adaptée de Tresset et al. (2017).

### 2.2.2 Le matériel génétique

La section 2.1.2.2 m’avait permis d’introduire le CCMV comme possédant un ARN simple brin de polarité positive. Pour autant, une infection effective du CCMV sera caractérisée par la présence de trois particules virales empaquetant quatre segments d’ARN différents les uns des autres (van de Waterbeemd et al. (2016)). Le CCMV est dit multipartite et on appelle ARN1, ARN2, ARN3 et ARN4, les quatre séquences de nucléotides qui permettent de décrire le génome de ce virus. La répartition du matériel génétique est la suivante : les ARN 1 et 2, longs de 3171 et 2774 nucléotides respectivement, sont présents dans deux capsides différentes alors que les ARN 3 et 4, longs de 2173 et 824 nucléotides, sont co-empaquetés dans une même particule (Figure 2.14.A). En somme, chaque capsid va contenir approximativement la même longueur de génome avec environ 3000 nucléotides ce qui correspond à une masse molaire de 1 MDa. Par ailleurs, la capacité d’encapsider plusieurs segments d’ARN dans une même particule a été observée *in vitro* par Cadena-Nava et al. (2012). En effet, ils ont montré que pour des séquences relativement courtes ( $< 2000$  nucléotides), une capsid peut contenir un à quatre segments distincts. De plus, si les séquences deviennent plus longues que l’ARN natif ( $> 4000$  nucléotides), des structures plus exotiques peuvent se former avec un segment d’ARN partagé par des doublets, triplets voire quadruplets de capsides (Figure 2.14.B).



**FIGURE 2.14** – **Diversité des longueurs d’ARN dans une capsid de CCMV.** A) Illustration du génome du CCMV, figure adaptée de van de Waterbeemd et al. (2016). B) Caractérisation de la forme des capsides de CCMV selon la longueur d’ARN à empaqueter : (1) Évolution du diamètre des capsides en fonction de la longueur des séquences d’ARN simple brin qu’elles contiennent (zone I = plusieurs segments d’ARN par capsid, II = un seul segment d’ARN par capsid, III = plusieurs capsides peuvent se former autour d’un même segment d’ARN), (2) Images de microscopie électronique à coloration négative des différents assemblages observés, figures adaptées de Cadena-Nava et al. (2012).

## 2.3 Quantifier l’assemblage de virus ...

Dans la première partie de ce chapitre, j’ai évoqué les travaux de plusieurs équipes qui ont permis d’en apprendre davantage sur les virus, et notamment leurs mécanismes d’assemblage. Ces recherches ont été permises grâce à l’utilisation de différents outils de mesure. Toutes ces techniques ont des spécificités propres mais chacune peut donner des informations, à différents degrés de précision et plus ou moins facilement, sur l’état d’assemblage de virus. Il est possible de classer ces techniques de différentes façons, mais ici j’ai choisi de les différencier selon leur capacité à quantifier et qualifier l’assemblage d’un point de vue du virus unique. Je reviendrai également sur les avantages et les limitations de ces systèmes de mesure. Il est important de préciser que les techniques présentées ci-dessous ne constituent en rien une liste exhaustive. Je me suis principalement concentré sur les outils de mesure utilisés par les équipes dont j’ai présenté les travaux lors de la section 2.1.4.

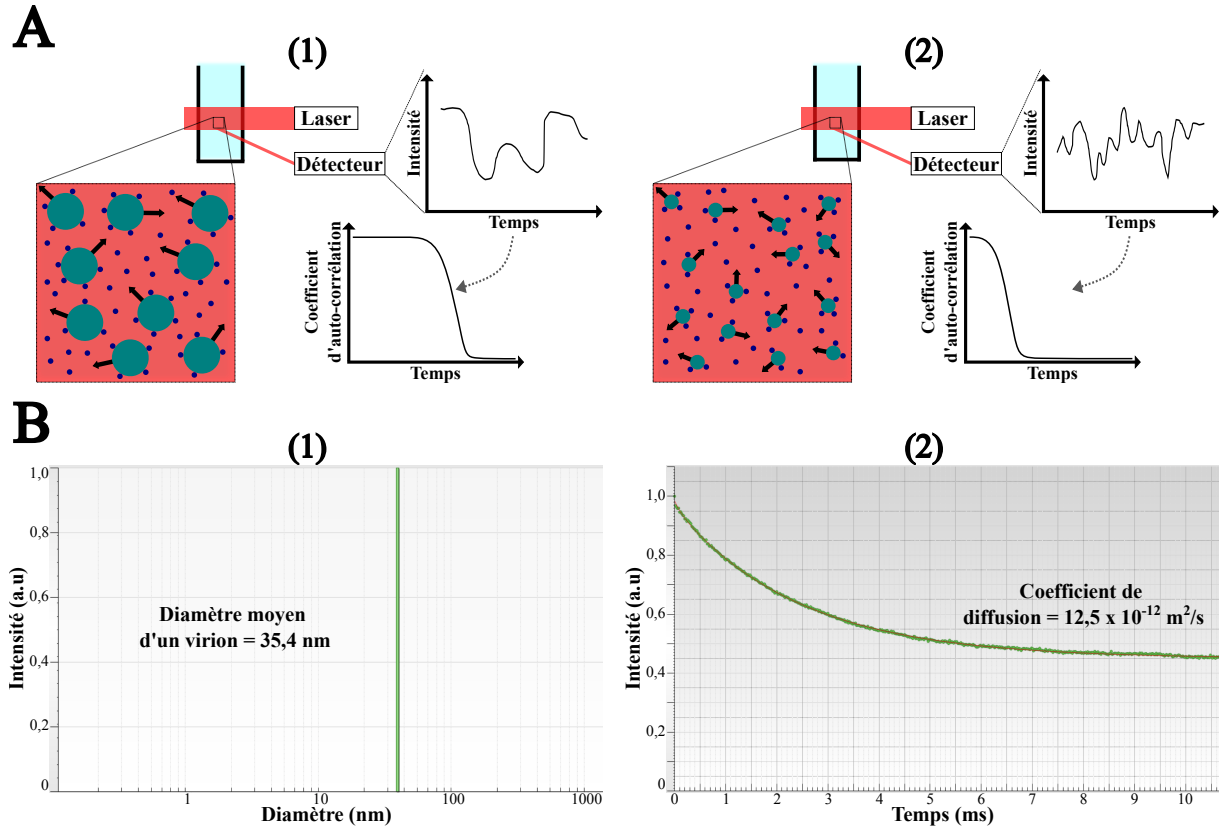
### 2.3.1 ... à l’aide de mesures d’ensemble

Pour commencer, je vais présenter certaines des techniques qui se basent sur des mesures d’ensemble afin de décrire le comportement global de l’échantillon étudié.

#### 2.3.1.1 Diffusion dynamique de la lumière

Pour vérifier que des capsides se sont bien assemblées, une méthode assez simple est d’effectuer des mesures de diffusion dynamique de la lumière ou DLS (*Dynamic Light*

*Scattering*). Cette technique permet d'estimer le diamètre des particules présentes en solution en mesurant les fluctuations d'intensité de la lumière diffusée par ces dernières. Ces variations proviennent du mouvement des particules et sont directement liées à la taille de ces dernières : plus le diamètre des particules est grand, moins elles se déplacent rapidement et donc plus les fluctuations d'intensité évoluent lentement (Figure 2.15.A).



**FIGURE 2.15** – Principe d'une mesure par DLS et exemple expérimental. A) Illustration du principe de mesure faite en DLS avec l'enregistrement de la lumière diffusée par des particules et l'allure de la fonction d'auto-corrélation : Exemple avec de grandes (1) et petites (2) particules, les molécules d'eau sont représentées en bleu foncé et les flèches noires indiquent le mouvement des particules. B) Mesure expérimentale réalisée avec l'appareil *Vasco-Flex*, *Cordouan Technology* sur une solution de virions de CCMV stockée à pH = 4,8 : (1) Histogramme du diamètre des particules détectées, (2) Fonction d'auto-corrélation.

L'intensité collectée  $I$  est ensuite analysée à l'aide d'une fonction d'auto-corrélation qui permet de comparer le signal à l'instant  $t$  avec lui même à un instant  $t + \tau$  où  $\tau$  correspond à un léger décalage temporel :

$$G(\tau) = \frac{\langle I(t)I(t + \tau) \rangle}{\langle I(t) \rangle^2} \quad (2.1)$$

Plus  $\tau$  sera grand, plus les deux signaux seront différents et donc plus  $G$  sera faible. La diminution de  $G$  est d'autant plus rapide que les fluctuations de  $I$  sont importantes et donc que les particules sont petites (Figure 2.15.A). Différents modèles mathématiques

peuvent être utilisés pour caractériser la décroissance de la fonction d'auto-corrélation et ainsi obtenir le coefficient de diffusion  $D$  des particules en solution. Enfin, en utilisant la relation de Stokes-Einstein, le diamètre hydrodynamique  $d_H$  des particules peut être calculé :

$$d_H = \frac{kT}{3\pi\eta D} \quad (2.2)$$

où  $k$  est la constante de Boltzmann,  $T$  la température et  $\eta$  la viscosité du milieu. La valeur de  $d_H$  est généralement plus élevée que le diamètre effectif de la particule puisqu'il prend en compte les interactions avec les molécules du solvant. La Figure 2.15.B montre une mesure expérimentale de DLS (*Vasco-Flex, Cordouan Technology*) effectuée au laboratoire LuMIn sur des virions de CCMV. Les virions étudiés ont un diamètre hydrodynamique d'environ 35 nm, ce qui est, comme attendu, légèrement supérieur à la valeur de 28 nm que l'on attend pour des virions natifs de CCMV stockés à pH = 4,8.

Les principaux avantages de cette technique sont qu'il est possible d'obtenir en une à deux minutes une estimation de la taille des particules que l'on souhaite mesurer (de 0,3 nm à 10  $\mu$ m), le dispositif expérimental prend peu de place et la mesure en elle-même consomme assez peu d'échantillons. En revanche, comme je l'ai dit, une surestimation systématique est effectuée sur le diamètre de la particule, ce qui peut compliquer la comparaison d'échantillons de tailles proches. Par ailleurs, les mesures dépendent grandement des conditions expérimentales (température, viscosité de milieu) pouvant ainsi rendre difficile les comparaisons entre différentes expériences.

### 2.3.1.2 Diffusion des rayons X aux petits angles

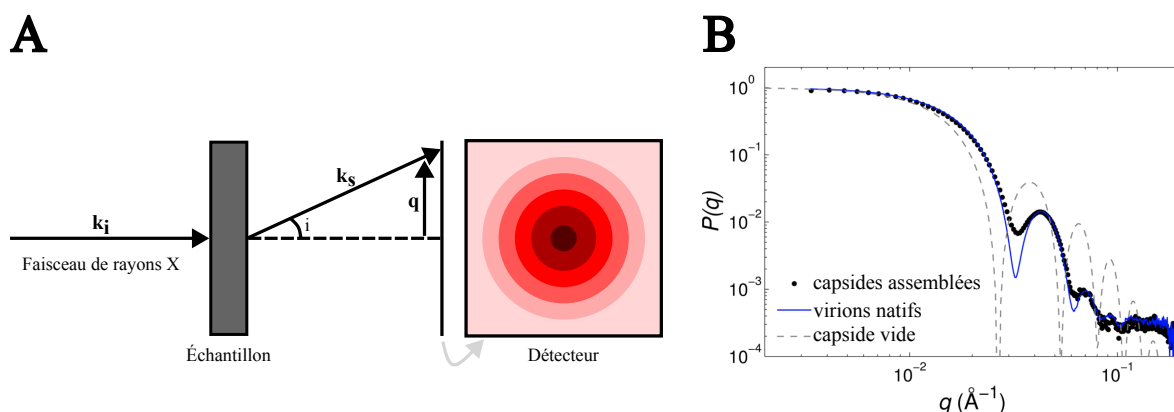
Toujours selon le principe de la diffusion de lumière, la technique de *Small-Angle X-ray Scattering* (SAXS) permet de bien quantifier la taille des particules que l'on souhaite analyser. Avec cette méthode, l'échantillon est éclairé par un faisceau incident de rayons X de vecteur de propagation  $\vec{k}_i$  et un détecteur permet de collecter les rayons diffusés par les particules de l'échantillon, pour lesquels le vecteur de propagation est noté  $\vec{k}_s$  (Figure 2.16.A). Ensuite, le but est de mesurer l'évolution de l'intensité diffusée  $I$  en fonction du vecteur d'onde  $\vec{q}$  qui caractérise la déviation par rapport au faisceau incident de rayons X. Des signaux typiques de virions natifs de CCMV et de capsides pleines ré-assemblées sont illustrés à la Figure 2.16.B. Les positions des arches aux faibles valeurs de  $q$  vont caractériser la taille globale des capsides alors que les plus grandes valeurs de  $q$  donnent des informations relatives à l'arrangement des sous-unités de capsides. Les résultats expérimentaux illustrés par la Figure 2.16.B montrent que la structuration des capsides pleines ré-assemblées n'est pas tout à fait celle des virions natifs.

Par ailleurs, il est possible de quantifier la cinétique d'assemblage de virus grâce à des mesures de *Time Resolved-SAXS* (TR-SAXS) en enregistrant l'évolution de l'intensité  $I$



au cours du temps. Des outils d'analyse, détaillés dans la thèse de Maelenn Chevreuil (Chevreuil (2020)), permettent d'évaluer avec précision la croissance des capsides (voir Figures 2.9 et 2.11).

Ainsi, les mesures par SAXS nous donnent accès à la taille des particules étudiées avec plus de précision et de détails que celles réalisées par DLS. En revanche, de manière générale, il est difficile de travailler avec des échantillons de diamètre supérieur à 100 nm, ce qui peut devenir limitant pour des virus plus grands que le CCMV. Dans certains cas, il est possible d'observer des particules de plusieurs centaines de nanomètres de diamètre, mais cela dépend en grande partie de la longueur de la ligne de lumière utilisée. De plus, les outils mathématiques utilisés pour l'analyse sont quand même plus complexes à prendre en main que ceux de la DLS. Enfin, que ce soit avec les dispositifs expérimentaux de laboratoire ou les lignes de lumière des synchrotrons, les expériences sont moins simples et accessibles que celles de la DLS.



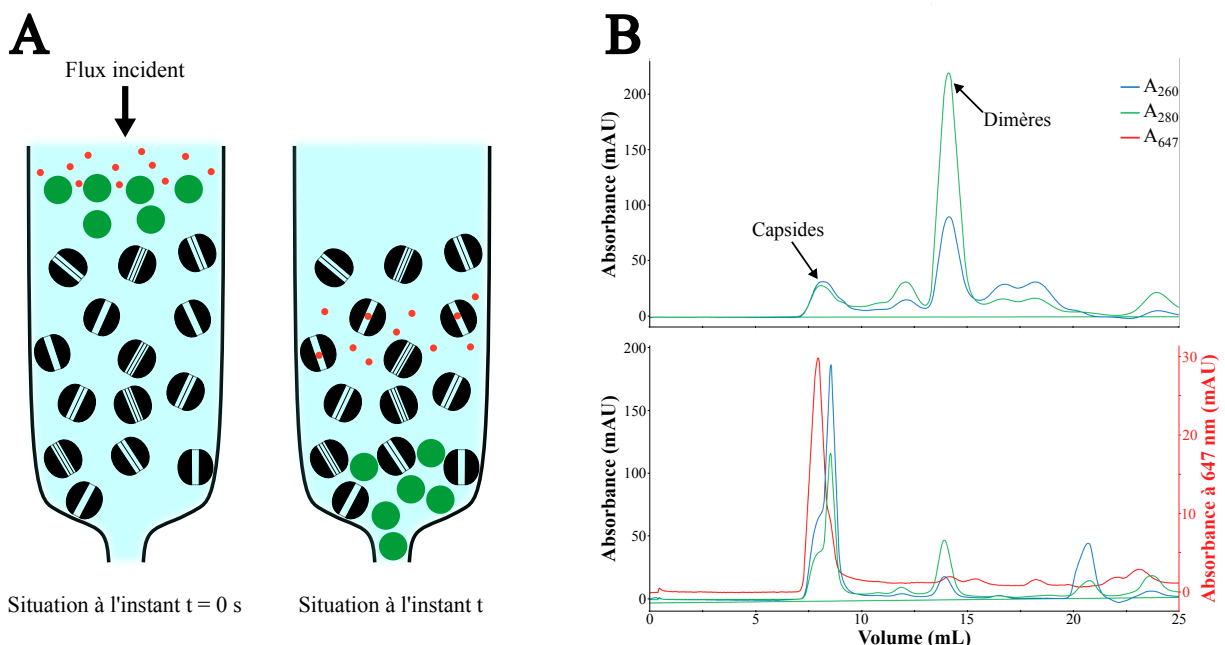
**FIGURE 2.16** – Principe d'une mesure par SAXS et exemple expérimental. A) Illustration du principe de mesure faite en SAXS. B) Comparaison des signaux de diffusion pour des capsides vides de CCMV (pointillés gris, données simulées), des virions de CCMV natifs (courbe bleue, mesure expérimentale) et des capsides pleines de CCMV ré-assemblées (points noirs, mesure expérimentale), figure adaptée de Chevreuil et al. (2018).

### 2.3.1.3 Chromatographie d'exclusion de taille

Cette dernière mesure d'ensemble est un peu à part parmi toutes les méthodes expérimentales que je présente dans cette section puisqu'elle ne permet pas de mesurer la taille des particules étudiées. En effet, la *Size Exclusion Chromatography* (SEC) se base sur l'utilisation d'une colonne remplie de billes pour séparer des particules d'un même mélange selon leur taille. Les billes ont la particularité d'être plus ou moins creuses formant ainsi un maillage qui a pour but de retenir les macro-molécules avec une efficacité qui dépend de leur taille. On distingue alors la phase fixe ou stationnaire, qui correspond aux billes creuses, de la phase mobile qui permet le déplacement des molécules d'intérêt au sein de ce réseau. Ainsi, les objets plus gros que le maillage ne seront pas retenus et sortiront en premier de la colonne alors que les macro-molécules de taille intermédiaire

seront plus ou moins freinées et sortiront plus tard (Figure 2.17.A). Il est important de préciser que les billes sont supposées n'avoir aucune affinité avec les macro-molécules, ces dernières ne sont donc sélectionnées que selon leur taille.

Sur la Figure 2.17.B deux chromatogrammes comparent des solutions de capsides de HBV marquées et non marquées ayant été dialysées dans un tampon de désassemblage. Les trois absorbances étudiées ici illustrent les contributions du matériel génétique à 260 nm, des protéines à 280 nm et du marqueur fluorescent Atto647N à 647 nm. L'équivalence entre les volumes auxquels les pics d'absorbance ont lieu et les masses molaires des objets élués nécessite une calibration sur laquelle je ne reviendrai pas mais qui est détaillée dans la thèse de Maeleann Chevreuil pour le CCMV et qui suit le même principe ici (Chevreuil (2020)). Dans notre cas, les pics d'absorbance à 8 - 8,5 mL caractérisent les capsides et ceux à 14 mL les dimères : les autres pics correspondent à des oligomères ou des molécules plus petites que des dimères. Ici, on remarque que l'absorbance à 280 nm des dimères est environ 8 fois supérieure à celle des capsides pour la solution non marquée (218 contre 28 mAU) alors qu'elle est environ 2,5 fois inférieure pour la solution marquée (47 contre 116 mAU). De plus, l'absorbance à 647 nm est la plus élevée au niveau du pic de capsides (30 mAU) alors qu'elle est très faible pour celui des dimères (2 mAU). En résumé, le marquage des capsides de HBV rend le désassemblage moins efficace et les dimères obtenus ne semblent pas être marqués.



**FIGURE 2.17** – Principe d'une mesure par SEC et exemple expérimental. A) Illustration du principe d'éluion de particules de différentes tailles. B) Chromatogrammes de solutions de capsides de HBV non marquées (graphique du haut) et marquées (graphique du bas) après dialyse dans un tampon de désassemblage. Les traits bleus, verts et rouges indiquent respectivement les absorbances à 260 nm, 280 nm et 647 nm.

Ainsi, la chromatographie d'exclusion de taille permet d'identifier les différentes populations présentes en solution et, comme ici, de caractériser l'efficacité d'un désassemblage. Une fois cette méthode utilisée, il est possible de récupérer les volumes d'élution des pics d'intérêt afin de quantifier chaque population (Lutomski et al. (2018)). En revanche, seule, elle ne permet pas d'obtenir des informations quantitatives sur la taille ou la structure des capsides.

### 2.3.2 ... à l'échelle du virus unique

L'inconvénient qu'ont les trois techniques présentées précédemment est qu'elles vont décrire un comportement global des particules présentes en solution. En moyenne, on peut savoir que les virus ont une certaine taille ou qu'ils se sont assemblés avec une certaine vitesse. Dès lors, il est difficile de faire état de l'hétérogénéité de l'échantillon étudié que ce soit en termes de structure finale ou de voie d'assemblage empruntée. Dans cette nouvelle section, je vais m'attarder sur des méthodes qui fonctionnent à l'échelle du virus unique.

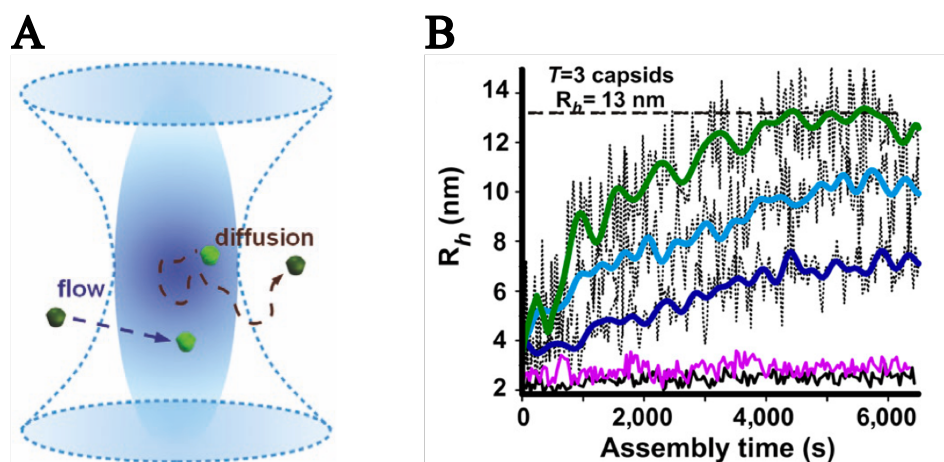
#### 2.3.2.1 Méthodes de mesures indirectes

Le terme de "mesures indirectes" indique que les deux techniques qui seront mises en avant ne permettent pas d'observer directement les particules d'intérêt, au contraire de celles présentées dans la section suivante.

##### - Spectroscopie de corrélation de fluorescence :

Borodavka et al. (2012) ont mené des expériences d'assemblage sur le virus du MS2 qui ont mis en lumière l'existence d'interactions spécifiques, *i.e. packaging signals*, entre les protéines de capsid et leur ARN grâce à des expériences de *Fluorescence Correlation Spectroscopy* (FCS). Par exemple, ils ont été capables de mesurer les rayons des capsides de MS2 pour des assemblages réalisés autour de différents ARN (Figure 2.18.B) : en vert, bleu foncé et bleu clair il s'agit d'ARN viraux du MS2 de différentes tailles alors qu'en violet et noir, des ARN non viraux ont été utilisés.

Le principe de cette méthode ressemble fortement à celui de la DLS vu que l'on s'intéresse aux fluctuations temporelles d'intensité. Cependant, au lieu d'enregistrer la lumière diffusée par les particules d'intérêt, seule la fluorescence des particules marquées est collectée. Ensuite, l'analyse des fluctuations se base une nouvelle fois sur le calcul de la fonction d'auto-corrélation  $G$  (équation 2.1) qui permet de retrouver deux propriétés importantes des échantillons marqués : le nombre moyen de fluorophores ainsi que leur temps de résidence moyen par unité de volume. En parallèle, il est possible d'avoir accès au coefficient de diffusion des particules marquées et d'en déduire leur diamètre hydrodynamique (équation 2.2). Sur la Figure 2.18.B, les mesures sont réalisées en continu afin de pouvoir suivre l'évolution de la taille des capsides de MS2 durant leur assemblage.



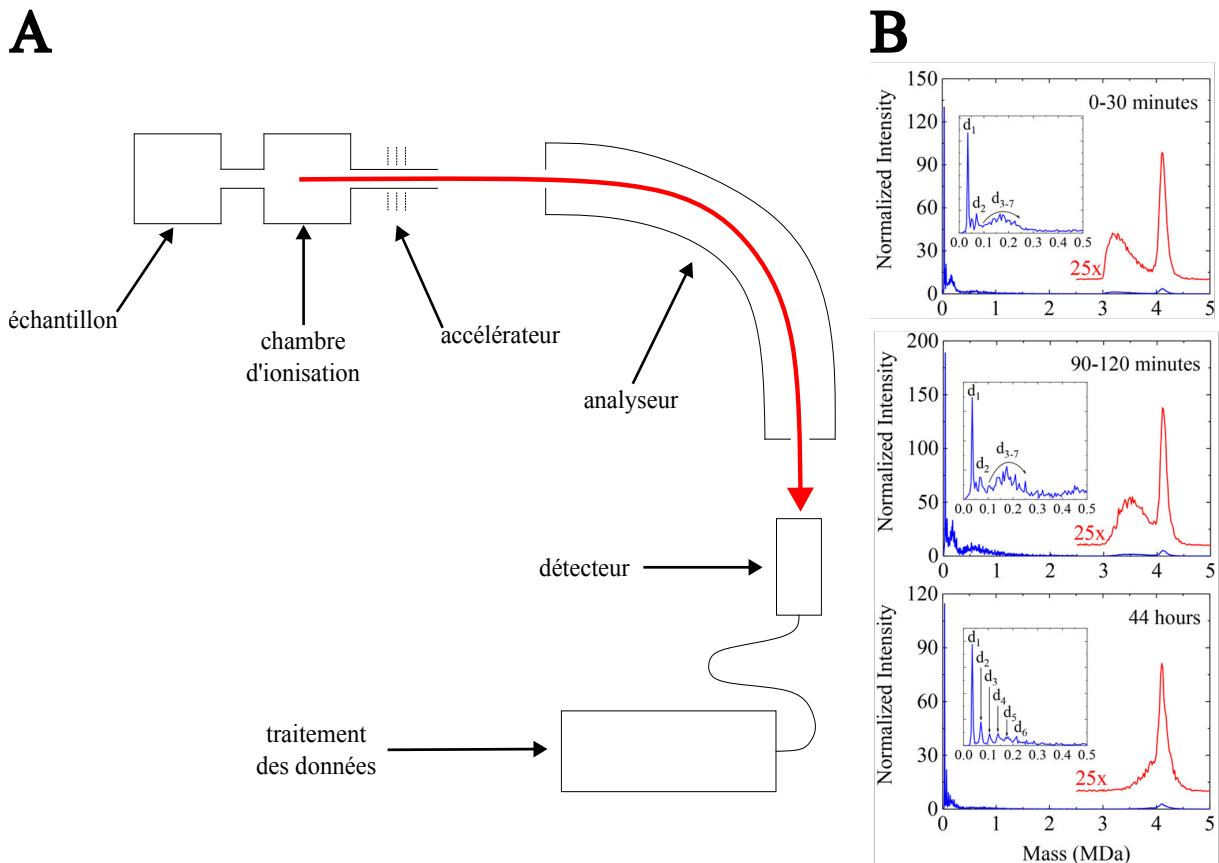
**FIGURE 2.18** – Schéma de l'illumination en FCS et exemple expérimental. A) Volume d'excitation d'un microscope confocal et diffusion de particules fluorescentes. B) Évolution du rayon hydrodynamique moyen des capsides de MS2 durant l'assemblage des protéines de capside autour de différents ARN. Les traits pleins correspondent au lissage des données brutes (pointillés), figure adaptée de Borodavka et al. (2012).

En se basant sur son mode de fonctionnement, la FCS peut être considérée comme appartenant à la catégorie des mesures d'ensemble, au même titre que la DLS. Néanmoins, cette technique est généralement utilisée à l'aide d'un microscope confocal, rendant ainsi le volume d'observation très petit (Figure 2.18.A). Cette propriété, combinée aux faibles concentrations choisies en molécules d'intérêt (entre le picomolaire et le nanomolaire), fait que la FCS est souvent considérée comme une technique de molécule unique. En effet, des événements uniques d'entrée, de sortie ou de changement de propriété photophysique des fluorophores dans le volume d'intérêt peuvent être enregistrés. En contrepartie, le fait de travailler à de très faibles concentrations peut altérer les mécanismes biologiques d'intérêt en les rendant notamment moins favorables. Par ailleurs, une autre limitation de cette technique peut être associée à la spécificité des instruments nécessaires à l'élaboration de ce montage optique : microscope confocal, détecteur de photon unique ...

#### - Spectrométrie de masse à détection de charge :

La spectrométrie de masse est une technique inventée au début du  $XX^e$  siècle et qui permet de déterminer la masse molaire de particules d'intérêt. Ces dernières sont d'abord transformées en ions plus ou moins chargés à l'aide d'une source ionisante. Ensuite, les ions créés sont accélérés avant de passer par un analyseur où ils sont triés en fonction de leur rapport masse sur charge. La dernière étape consiste à transformer les ions séparés en signal électrique via un détecteur avant de traiter des données acquises (Figure 2.19.A). Il existe plusieurs variantes de spectromètre de masse en fonction des méthodes d'analyse et de détection utilisées mais globalement le résultat en sortie reste le même : séparer les ions en fonction de leur rapport masse sur charge. Néanmoins, des difficultés expérimentales concernant la transformation de macro-molécules en macro-ions mais aussi la séparation

précise des charges contenues par ces derniers ont rendu l'utilisation de cette technique peu précise pour des particules plus grosses que 100 kDa. C'est grâce aux travaux de John Bennett Fenn puis de Stephen D. Fuerstenau et W. Henry Benner que ces difficultés expérimentales ont pu être surmontées, rendant la caractérisation de macro-molécules possible (Yamashita and Fenn (1984) et Fuerstenau and Benner (1995)). On parle alors de *Charge Detection Mass Spectrometer* (CDMS).



**FIGURE 2.19** – Schéma d'un spectromètre de masse et exemple expérimental. A) Montage général d'un spectromètre de masse. B) Suivi de l'évolution des distributions des différentes particules présentes en solution lors de l'assemblage de capsides de HBV, les encarts sont des agrandissements des tracés allant jusqu'à 500 kDa et les signaux tracés en rouge ont été multipliés par 25, figure adaptée de Lutomski et al. (2018).

Cette méthode expérimentale, dite à ion unique, permet d'avoir accès directement à la masse individuelle de chaque objet à l'aide d'un montage plus précis dans lequel la vitesse de chaque ion est mesurée conjointement à la charge induite par ce dernier sur un cylindre métallique. C'est en utilisant ce montage que Lutomski et al. (2018) peuvent avoir accès à toute la population des particules présentes en solution : des dimères aux capsides complètes de HBV en passant par les oligomères piégés cinétiquement. Par ailleurs, la Figure 2.19.B montre que cette méthode peut être utilisée afin de suivre la répartition de ces trois populations de particules dans le temps : on parle alors de *Time Resolved-CDMS*.

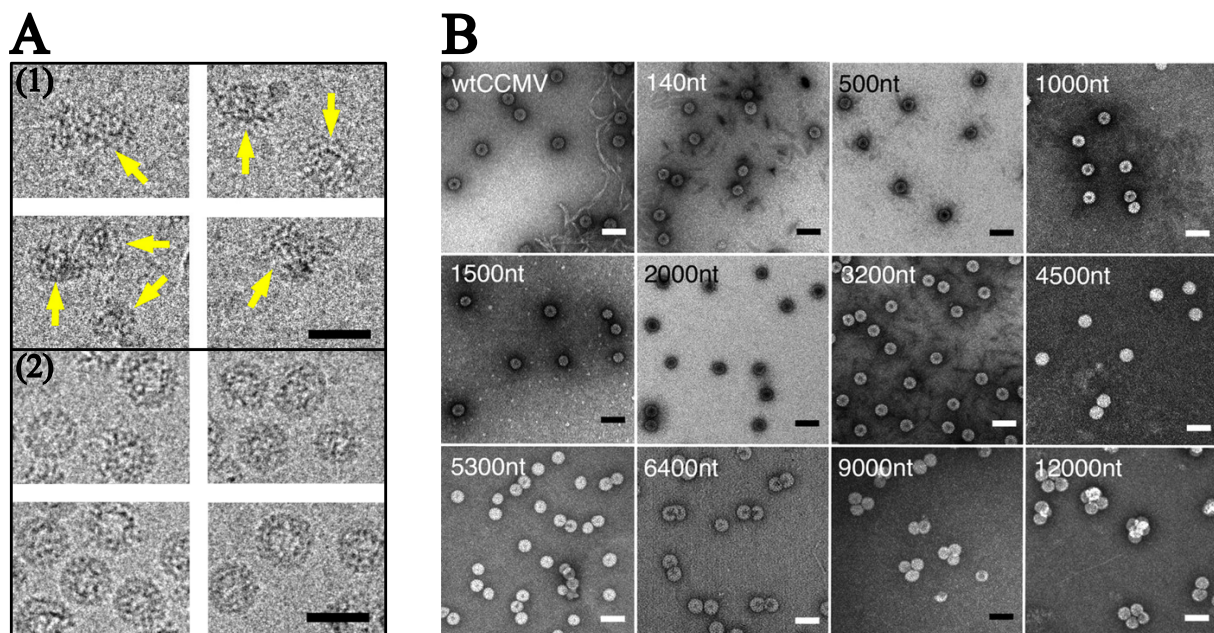
Les améliorations expérimentales réalisées sur le montage initial d'un spectromètre de

masse ont donc permis de mesurer individuellement les charges de macro-ions et remonter à la masse moléculaire des macro-molécules initiales. Néanmoins, à la différence des techniques précédentes, les échantillons ne sont plus utilisables à la suite d'une mesure et le dispositif expérimental est assez encombrant et donc peu pratique à utiliser.

### 2.3.2.2 Méthodes d'observations directes

#### - Microscopie électronique :

À plusieurs reprises dans la section 2.1, j'ai pu m'appuyer sur des images obtenues en microscopie électronique à transmission (*Transmission Electron Microscopy* ou TEM) pour illustrer mon point de vue (Figures 2.1, 2.3, 2.6, 2.7, 2.12, 2.14). Ces microscopes utilisent un faisceau d'électrons qui ont une longueur d'onde bien plus courte que celle d'un photon visible. Cette propriété entraîne une résolution proche de la centaine de picomètres contre environ 250 nm pour des microscopes optiques classiques. De plus, le facteur de grossissement peut aller jusqu'à  $2.10^6$  contre environ  $2.10^3$  pour les meilleurs microscopes optiques. Ces avantages font qu'il est possible d'observer individuellement de petits virus tels que le CCMV et le HBV (30 nm), de comparer leur forme à celle de virions natifs et ainsi mettre en lumière des défauts durant l'assemblage comme des capsides de CCMV incomplètes ou qui s'associent de façon particulière (respectivement Figure 2.20.A et B).



**FIGURE 2.20** – Observations de capsides de CCMV en TEM. A) Comparaison de capsides de CCMV incomplètes illustrées par des flèches jaunes (1) et complètes (2) observées en cryo-microscopie électronique, barres d'échelle de 30 nm, figures adaptées de Chevreuil et al. (2018). B) Résultats de l'assemblage de capsides de CCMV autour d'ARN de différentes tailles observées en TEM à coloration négative, barres d'échelle de 50 nm, figure adaptée de Cadena-Nava et al. (2012).

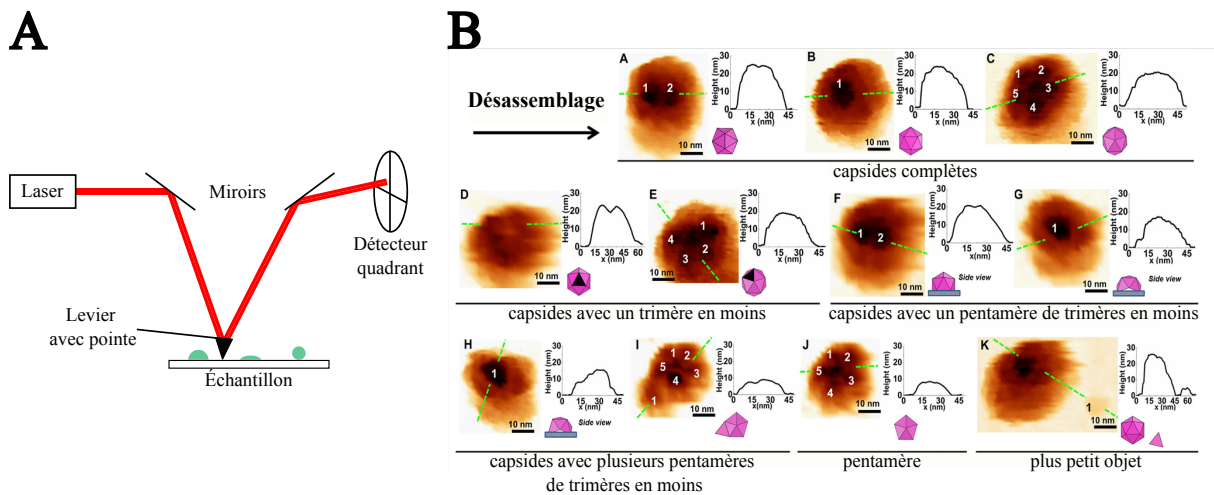
Néanmoins, un microscope électronique est à la fois très onéreux et encombrant, la

préparation des échantillons peut s'avérer délicate (utilisation de grilles spécifiques, coloration, cryo-microscopie) et comme pour le spectromètre de masse, les échantillons ne peuvent plus être récupérés après avoir été imagés.

- *Microscopie à force atomique* :

Les travaux de Medrano et al. (2016) ont permis de mettre en lumière des intermédiaires d'assemblage du virus minute de la souris (MVM), un virus  $T = 1$ , de 25 nm de diamètre. Ces résultats ont été rendus possibles grâce à la combinaison de données obtenues en microscopie électronique et en microscopie à force atomique (*Atomic Force Microscopy* ou AFM).

Cette technique utilise une pointe possédant un rayon de courbure de quelques nanomètres afin de cartographier une surface et y observer des échantillons d'intérêt. La topographie de la surface est possible grâce à l'enregistrement du déplacement d'un faisceau laser réfléchi par la face arrière de la pointe (Figure 2.21.A). Avec cette technique, la résolution (X,Y) est directement liée à la taille de la pointe et peut atteindre quelques nanomètres en latéral contre quelques angströms pour la résolution verticale. C'est grâce à ces propriétés que Medrano et al. (2016) ont pu identifier des états intermédiaires de la capsid de MVM et en observer certains formés de quelques oligomères (Figure 2.21.B). Par ailleurs, Buzon et al. (2021) ont pu utiliser un AFM à haute vitesse (*High-Speed AFM*) afin de suivre en temps réel la dynamique d'assemblage d'oligomères de HBV.



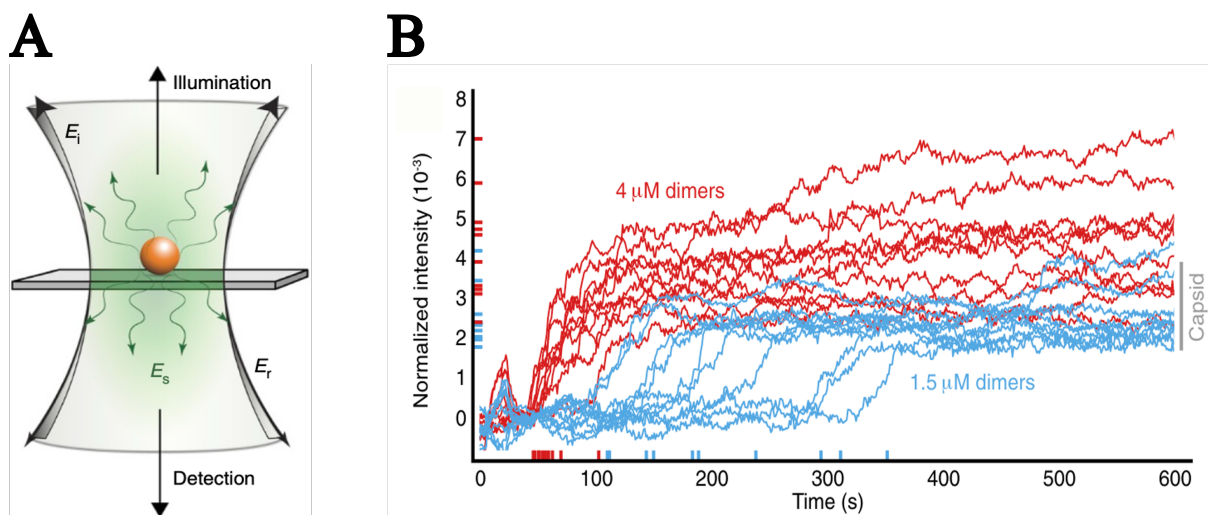
**FIGURE 2.21** – Principe d'une mesure par AFM et exemple expérimental. A) Illustration du principe de mesure par AFM. B) Observation par AFM de différents intermédiaires lors du désassemblage de capsides de MVM, les profils de hauteur sont donnés selon la direction des pointillés verts, barres d'échelle de 10 nm, figure adaptée de Medrano et al. (2016).

Cependant, cette méthode se base sur des mesures réalisées en surface où les objets étudiés interagissent avec cette dernière. À ce stade, nous ne pouvons pas observer l'assemblage d'une capsid en trois dimensions autour de son matériel génétique. Par ailleurs,

la fragilité et l'usure des pointes sont à souligner même si ces deux points peuvent être améliorés en choisissant judicieusement le mode de fonctionnement du microscope. En effet le mode *contact* use plus facilement la pointe car cette dernière est toujours au contact de l'échantillon alors qu'avec le mode *tapping*, elle ne fait que le tapoter par intermittence.

- Microscopie à diffusion interférométrique :

Pour finir, je vais rapidement présenter la technique de microscopie à diffusion interférométrique (*interferometric Scattering* ou iSCAT) notamment utilisée dans les travaux de Garmann et al. (2019). Le principe optique se base sur la création d'interférences entre un faisceau de référence  $E_r$ , généralement réfléchi par une lamelle de microscope, et la lumière diffusée  $E_s$  par la particule étudiée (Figure 2.22.A). Le contraste des interférences obtenues est directement proportionnel au volume  $V$  de la particule (et non à  $V^2$  comme l'est l'intensité diffusée) ainsi qu'à son indice de réfraction, ce qui permet d'observer des virus adsorbés sur une lamelle de verre. Garmann et al. (2019) utilisent cette propriété afin d'observer l'auto-assemblage des capsides de MS2 autour de leur matériel génétique. Ici, plus la capside grandit, plus son diamètre augmente et donc plus le contraste des interférences est important (Figure 2.1.C). Un des avantages de cette technique est qu'il est possible de suivre la totalité de l'assemblage de plusieurs capsides à la fois (Figure 2.22.B). À partir de ces données expérimentales, on peut obtenir la distribution des durées de croissance des capsides et en déduire une valeur moyenne.



**FIGURE 2.22** – **Schéma de l'illumination en iSCAT et exemple expérimental.** A) Schéma d'une particule d'intérêt imagée lors d'une expérience d'iSCAT, avec le champ incident  $E_i$ , réfléchi  $E_r$  et diffusé  $E_s$ , figure adaptée de Ortega Arroyo et al. (2016). B) Suivi en temps réel de l'assemblage de capsides de MS2 autour de leur ARN viral en fonction de la concentration en dimères. La barre grise indique le signal typique d'un virion natif, figure adaptée de Garmann et al. (2019)

En revanche, cette méthode est très sensible aux vibrations et les images acquises



peuvent être affectées par les interférences parasites provenant des réflexions sur les différentes optiques utilisées. Par ailleurs, comme je l'ai dit, le contraste des interférences d'intérêt varie en fonction du volume de la particule, ce qui limite la taille des objets que l'on peut observer mais également la résolution. En effet, la Figure 2.22.B montre que nous ne pouvons pas détecter des événements individuels d'accroche de sous-unités de capsid.

Je reviens plus en détails sur cette technique d'imagerie au chapitre 6 du manuscrit. Par ailleurs, les résultats de photométrie de masse que je présente à la section 5.1 sont obtenus à partir du même principe optique que l'iSCAT.

### 2.3.3 Conclusion

Ainsi, j'ai pu revenir plus en détails sur les outils expérimentaux qu'il est possible d'utiliser afin de rendre compte d'un point de vue qualitatif ou quantitatif de l'assemblage de capsides virales. Les mesures dites d'ensemble peuvent donner des informations à la fois importantes et précises sur l'état des capsides obtenues mais elles renvoient des grandeurs caractéristiques moyennées sur l'ensemble de la population sans avoir accès à toute la distribution. C'est donc dans une optique de montrer l'hétérogénéité des échantillons obtenus que les méthodes dites de molécule unique (ou de virus unique ici) sont utilisées. De toutes les méthodes étudiées, la microscopie électronique est celle qui est la plus utilisée dans les travaux que j'ai présentés puisqu'en plus elle permet d'ajouter un côté visuel. Par ailleurs, il est important de retenir que même si le fait d'accéder à toute la distribution des populations de capsides est intéressant, le résultat final est toujours donné sous la forme d'une moyenne : la taille moyenne (Cadena-Nava et al. (2012) - EM), la masse moyenne (Lutomski et al. (2018) - CDMS) ou encore la durée moyenne de croissance (Garmann et al. (2019) - iSCAT) des capsides.

En revanche, parmi toutes les techniques expérimentales décrites, aucune ne permet de travailler à l'échelle du dimère unique ou de l'accroche unique. Il est possible de suivre la cinétique d'assemblage (TR-SAXS, FCS, TR-CDMS, iSCAT) mais pas de voir des sous-unités de capsides individuelles s'accrocher au matériel génétique. De plus, Chevreuil et al. (2018) montrent que dans certaines conditions expérimentales des capsides de CCMV avec 75 dimères sur 90 sont obtenues. À ce stade, les capsides sont toujours en contact des sous-unités libres en solution et il est intéressant de se demander si, dans cet état d'équilibre, des échanges de dimères ont lieu. Il serait intéressant de caractériser ces phénomènes d'arrivée et de départ de sous-unités en ayant accès, par exemple, au temps de résidence moyen d'un dimère sur le matériel génétique. Avant d'aller plus loin sur cet aspect, je vais désormais détailler le montage expérimental que j'ai utilisé durant ma thèse et expliquer en quoi il peut me permettre d'observer des événements individuels d'arrivée et de départ de sous-unités de capsid.

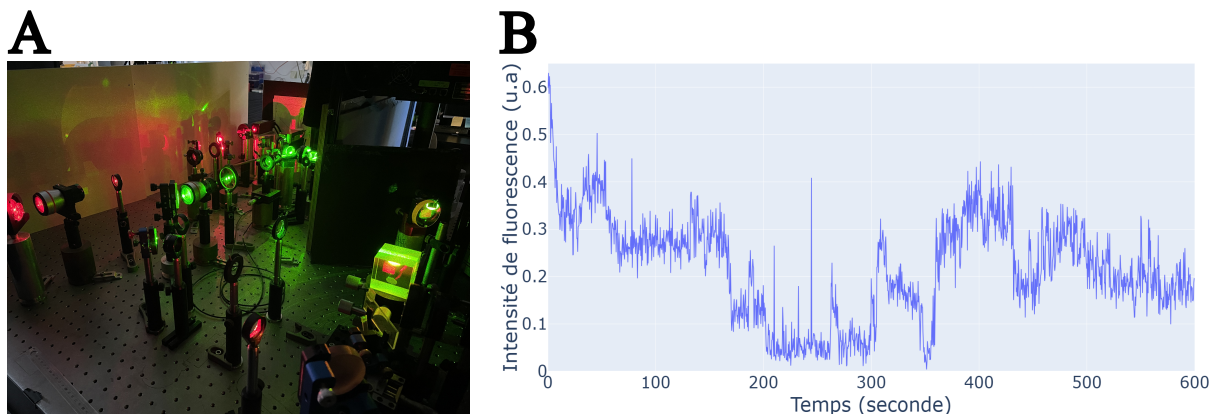
## Chapitre - 3

---

### Imagerie d'un virus à l'échelle de la molécule unique

---

Ce chapitre est entièrement consacré à la présentation des différentes composantes essentielles au bon déroulement des expériences de molécule unique réalisées pendant ma thèse. Je commence en décrivant le montage optique du TIRF (Figure 3.1.A), puis je reviens en détails sur l'étape de fonctionnalisation des lamelles de microscope. Dans la troisième partie, je rappelle brièvement quelques notions concernant le principe de fluorescence afin de décrire certaines propriétés photophysiques du fluorophore utilisé pour marquer les protéines, l'Atto647N. Enfin, ce chapitre se termine par la présentation du protocole de traitement des données appliqué après chaque expérience afin d'observer l'évolution de l'intensité de fluorescence au cours du temps comme illustré en Figure 3.1.B.



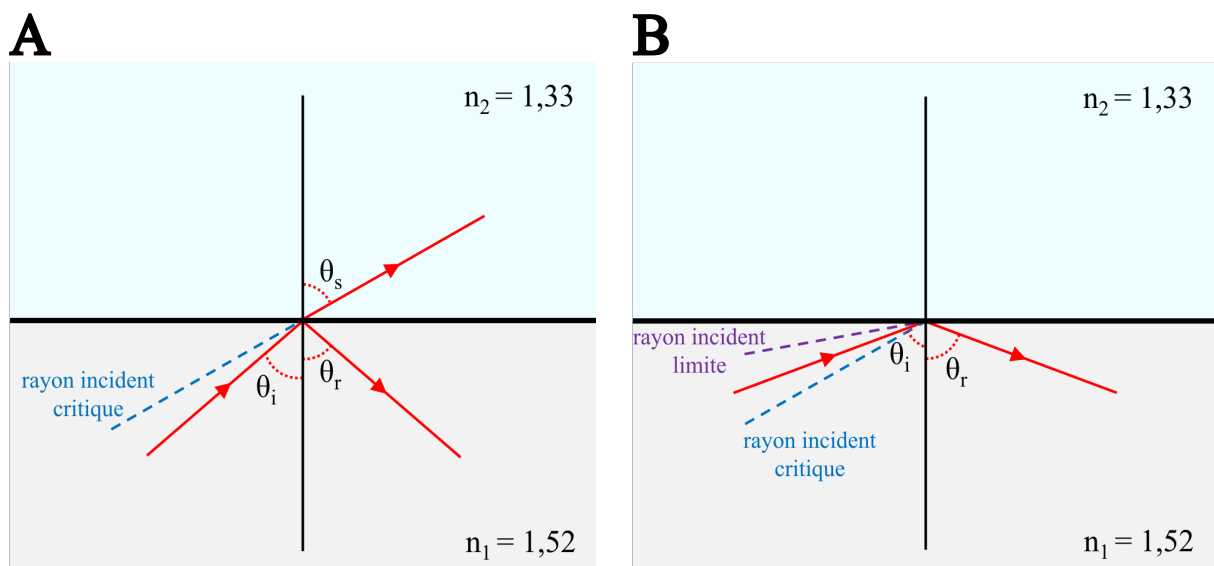
**FIGURE 3.1 – Vers l'imagerie d'un virus à l'échelle de la molécule unique.** A) Photo du montage optique qui permet l'illumination de l'échantillon en réflexion totale interne. B) Suivi de l'intensité de fluorescence émise par l'Atto647N au cours du temps, traduisant la dynamique d'assemblage de sous-unités marquées autour d'un ARN simple brin : temps d'exposition de 100 ms et temps de cycle de 375 ms.

### 3.1 Le TIRFM

La microscopie de fluorescence par réflexion totale interne (*Total Internal Reflection Fluorescence Microscopy* ou TIRFM) est une technique expérimentale très répandue en fluorescence puisqu'elle permet notamment de réduire le bruit parasite issu des zones de l'échantillon situées en dehors du plan de focalisation. Dans la première section, je reviens sur l'origine de cette propriété, en décrivant les paramètres physiques qui influencent la qualité de l'illumination utilisée. Dans un second temps, je détaille les différentes composantes du montage expérimental avec lequel j'ai travaillé pendant trois ans au laboratoire LuMIn. Pour finir, la dernière section est consacrée à la présentation de la cellule microfluidique dans laquelle les expériences de fluorescence sont réalisées.

#### 3.1.1 Principe de la technique

Durant une expérience de TIRFM, la réduction du bruit de fluorescence est obtenue grâce à la formation d'une onde lumineuse qui va éclairer l'échantillon sur une très faible épaisseur  $d$ . Cette onde électromagnétique, communément appelée onde évanescente, est obtenue dans des conditions d'illumination particulières qui nécessitent de rappeler le comportement d'un rayon lumineux au cours de son passage à travers une interface de deux milieux d'indice différent. Sur la Figure 3.2.A, on peut suivre le parcours d'un faisceau lumineux qui arrive sur une interface verre-eau, d'indices de réfraction respectifs  $n_1 = 1,52$  et  $n_2 = 1,33$ , avec un angle d'incidence  $\theta_i$ .



**FIGURE 3.2** – Propagation de rayons lumineux à travers une interface verre-eau. A) Cas de figure pour un angle incident inférieur à l'angle critique. B) Cas de figure pour un angle incident supérieur à l'angle critique mais inférieur à l'angle limite défini par l'ouverture numérique d'un objectif. Dans les deux cas :  $\theta_i$  = angle incident,  $\theta_r$  = angle réfléchi,  $\theta_s$  = angle réfracté.

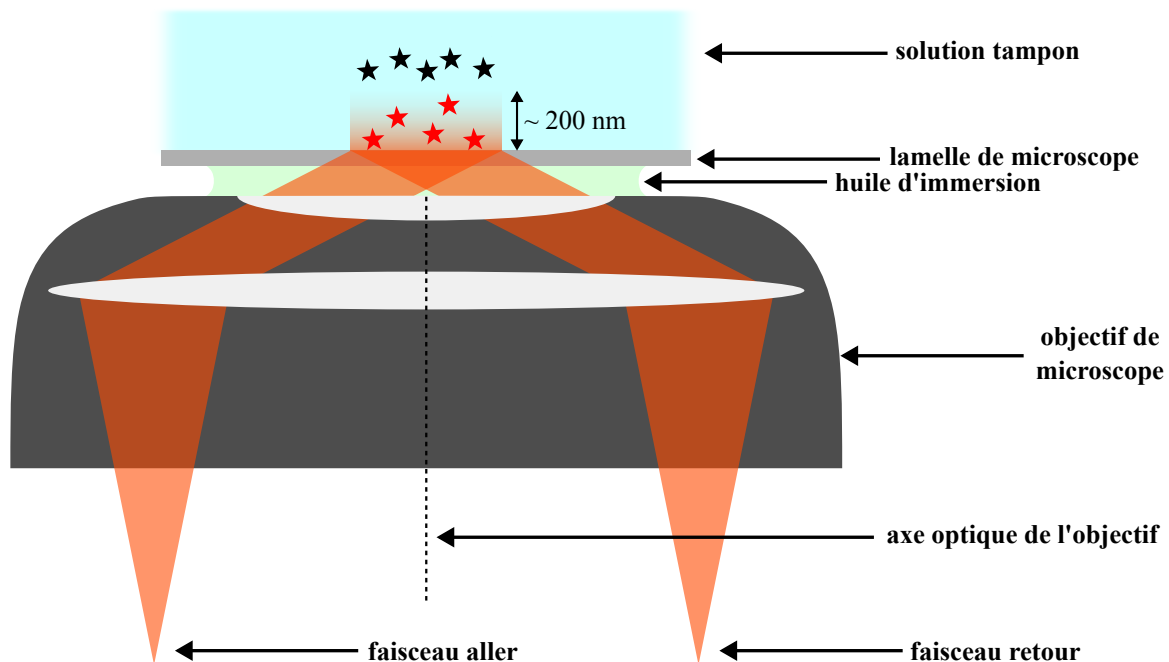
Cette interface décrit parfaitement les conditions expérimentales que nous utilisons puisque les échantillons d'intérêt présents en solution aqueuse sont déposés sur une lamelle de microscope. Dans ce cas, les lois de Snell-Descartes nous disent que le faisceau réfléchi fait un angle  $\theta_r$ , par rapport à la normale à l'interface, égal à  $\theta_i$  et que s'il y a un faisceau réfracté, son angle par rapport à la normale est donné par la relation suivante :

$$n_1 \sin(\theta_i) = n_2 \sin(\theta_s) \quad (3.1)$$

où  $\theta_s$  correspond à l'angle que fait le faisceau réfracté par rapport à la normale à l'interface. Pour une propagation allant du verre à l'eau, on remarque que l'angle  $\theta_s$  est plus grand que  $\theta_i$ . Ainsi, si  $\theta_i$  augmente, le rayon réfracté va s'éloigner encore plus de la normale, jusqu'à atteindre le cas limite où ce dernier sera confondu avec l'interface. Cette configuration est obtenue pour un angle incident critique  $\theta_{i,crit}$  que l'on peut calculer en posant  $\theta_s = 90^\circ$ . Dans le cas d'une interface verre-eau, et en réarrangeant légèrement l'équation 3.1, on obtient :

$$\theta_{i,crit} = \arcsin\left(\frac{n_2}{n_1}\right) \approx 61^\circ \quad (3.2)$$

Cela signifie que pour tout angle d'incidence  $\theta_i > \theta_{i,crit}$ , il n'y aura pas de rayon réfracté (Figure 3.2.B), c'est ce que l'on appelle le phénomène de réflexion totale interne. C'est dans ces conditions d'illumination particulières qu'une onde évanescente est créée dans le second milieu de propagation, comme cela est illustré par la Figure 3.3.



**FIGURE 3.3** – Excitation de marqueurs fluorescents en illumination TIRF. Les fluorophores excités par l'onde évanescente sont représentés par des étoiles rouges, alors que ceux qui sont trop éloignés de l'interface apparaissent en noir et n'émettent pas de lumière.

Quand on mentionne la microscopie TIRF, il ressort souvent que l'épaisseur  $d$  de l'onde évanescente correspond à une centaine de nanomètres. Plus généralement, il est possible de quantifier assez précisément cette grandeur puisque la relation suivante montre qu'elle dépend de plusieurs paramètres associés à l'interface étudiée :

$$d = \frac{\lambda_{exc}}{2\pi n_2} \frac{1}{\sqrt{\left(\frac{n_1}{n_2} \sin(\theta_i)\right)^2 - 1}} \quad (3.3)$$

où  $\lambda_{exc}$  correspond à la longueur d'onde d'excitation de la source lumineuse. Cette équation montre clairement que plus l'angle d'incidence augmente et tend vers  $90^\circ$ , plus l'épaisseur de l'onde évanescente diminue. Par exemple, en considérant les mêmes valeurs d'indice optique que précédemment,  $\lambda_{exc} = 638$  nm et  $\theta_i = 65^\circ$  ou  $85^\circ$ , on trouve respectivement :  $d = 282$  nm ou  $140$  nm. Ainsi, en modifiant l'indice de réfraction  $n_1$  ou l'angle d'incidence  $\theta_i$ , il est possible de confiner plus ou moins l'onde évanescente. Un confinement plus important permet de limiter la contribution des marqueurs fluorescents situés plus en profondeur et donc d'augmenter le rapport signal sur bruit.

Néanmoins, le choix de l'angle  $\theta_i$  est également contraint par l'ouverture numérique  $ON$  de l'objectif de microscope utilisé, puisque cette dernière lui impose une limite supérieure. Dans notre cas, nous utilisons un objectif à immersion dans l'huile (*Olympus*, UPlanApo 60X) d'ouverture numérique  $ON = 1,50$ . L'huile utilisée (*Zeiss*, Immersol 518F), possède un indice de réfraction  $n_h = 1,518$ , qui est très proche de l'indice du verre. De cette façon, les rayons émergents de l'objectif ne sont pas déviés lors du passage de l'huile au verre et c'est ce qui permet de travailler avec une grande ouverture numérique. Dans ces conditions, on peut conserver  $n_1 = 1,52$  et l'angle d'incidence limite peut se calculer grâce à la relation  $ON = n_1 \sin(\theta_i)$ , ce qui donne :  $\theta_{i,lim} \approx 81^\circ$ . Ainsi, en considérant les données expérimentales de notre montage, l'angle d'incidence est compris entre  $61^\circ$  et  $81^\circ$  (Figure 3.2.B), ce qui nous permet d'obtenir un confinement maximal de l'onde évanescente d'environ  $146$  nm. Dans sa thèse, Nathalie Barbier décrit précisément l'influence de l'ouverture numérique de l'objectif sur les performances d'une illumination TIRF (Barbier (2017)).

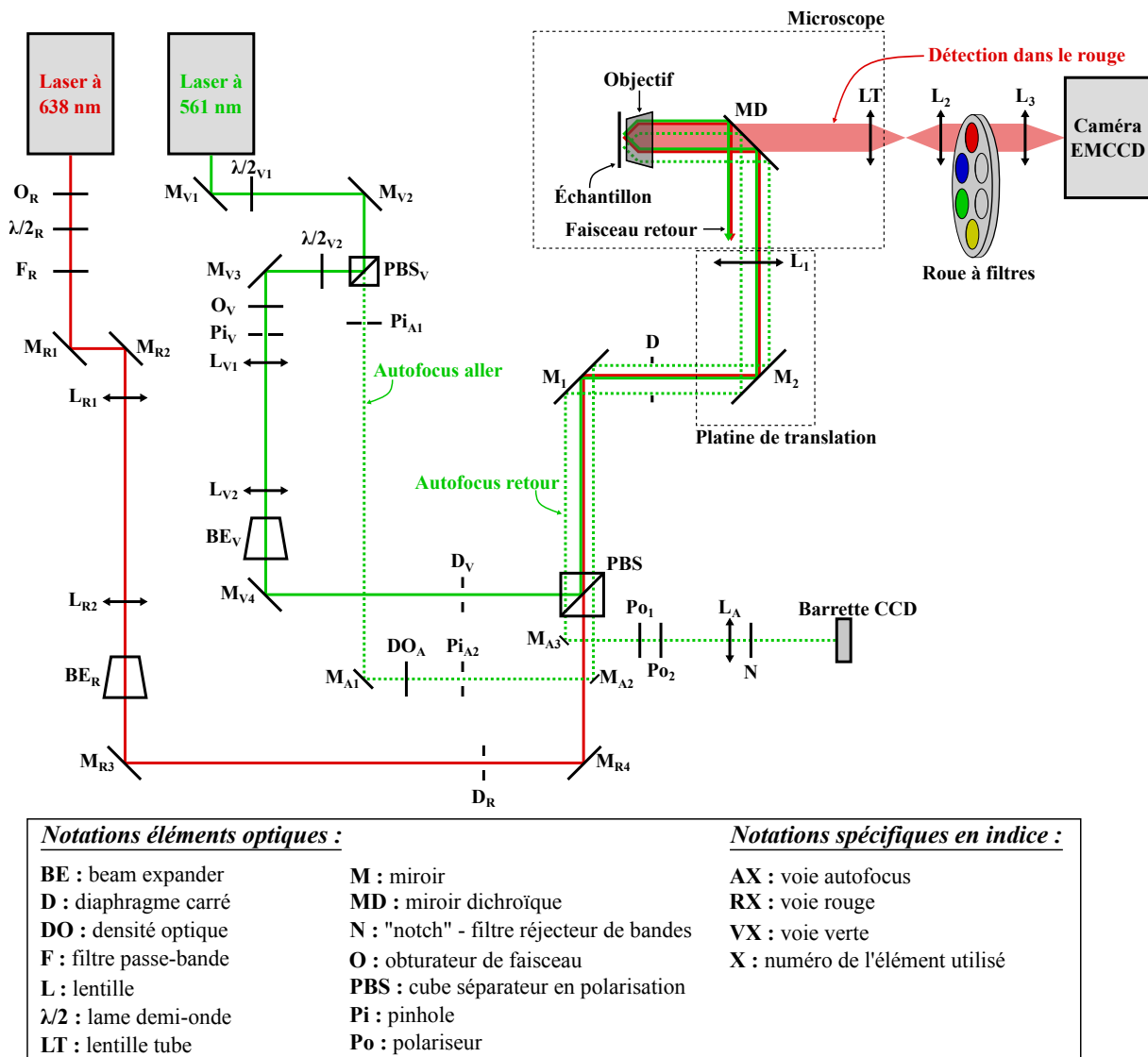
### 3.1.2 Description du montage

Dans la suite, je présente en détails le montage expérimental utilisé afin d'éclairer les échantillons en illumination TIRF. On distingue les voies d'excitation, de détection ainsi qu'une troisième voie qui nous permet de conserver la mise au point au cours du temps.

#### 3.1.2.1 Voies d'illumination et de détection

Le montage optique illustré par la Figure 3.4 permet la mise en forme et la propagation jusqu'à l'échantillon de deux faisceaux lasers : un laser rouge (*Cobolt*, 06-MLD), de

longueur d'onde d'émission 638 nm et dont la puissance nominale peut monter jusqu'à 250 mW et un laser vert (*Oxxius*, LCX 561), de longueur d'onde d'émission 561 nm et délivrant une puissance fixe de 75 mW. Le laser rouge permet d'observer les protéines marquées avec le fluorophore Atto647N alors que le laser vert est utilisé pour observer l'ARN marqué avec un Atto565 (voir section 3.2.3).



**FIGURE 3.4** – Schéma du montage optique permettant une illumination TIRF. Après le cube séparateur en polarisation, les voies rouges et vertes sont représentées l'une à côté de l'autre mais en réalité, elles sont confondues et surtout utilisées l'une après l'autre.

En début de montage, chaque faisceau laser mesure environ 700  $\mu\text{m}$  de diamètre et possède une polarisation linéaire *s*. Or, nous utilisons un objectif de microscope de grandissement  $\times 60$  et une caméra EMCCD (*Electron Multiplying Coupled Charge Device*) avec un détecteur de 512x512 pixels mesurant chacun 16  $\mu\text{m}$  de côté. Ainsi, chaque pixel voit une partie du plan de l'échantillon de 267x267  $\text{nm}^2$ , ce qui correspond à un champ total de 137x137  $\mu\text{m}^2$  environ. Dans la suite, je présente au fur et à mesure les différents

éléments qui nous permettent d'obtenir un champ d'excitation de cette taille.

Sur la voie rouge, juste après le laser, on retrouve une lame à retard de phase dite demi-onde, notée  $\lambda/2_R$ , qui va nous permettre de régler la puissance du faisceau d'excitation. Par la suite, le diamètre du laser rouge est agrandi à deux reprises, d'abord en passant à travers un télescope composé des lentilles  $L_{R1}$  (focale = 60 mm) et  $L_{R2}$  (focale = 200 mm) puis grâce à un *beam expander*, noté  $BE_R$  qui correspond ici à un télescope commercial de grandissement  $\times 10$ . En sortie de ces différents composants, le faisceau a été agrandi d'un facteur  $\times 33$  et mesure environ 28 mm de diamètre. Dans le cas du laser vert, le début du trajet optique est légèrement différent puisqu'une première lame demi-onde, notée  $\lambda/2_{V1}$ , est suivie d'un cube séparateur en polarisation ( $PBS_V$ ). L'association lame à retard - cube permet de décomposer la lumière incidente selon deux polarisations linéaires orthogonales entre elles,  $s$  et  $p$ , respectivement perpendiculaire et parallèle au plan d'incidence. Pour la voie d'excitation, on se concentre sur la composante réfléchie par le cube  $PBS_V$ . La polarisation transmise, décrite dans la prochaine section, est utilisée afin de maintenir la focalisation dans le plan de l'échantillon. Mise à part cette différence, la suite du trajet du laser vert est semblable à la voie rouge avec une seconde lame demi-onde, notée  $\lambda/2_{V1}$ , pour contrôler la puissance du faisceau d'excitation vert, un télescope composé des lentilles  $L_{V1}$  (focale = 30 mm) et  $L_{V2}$  (focale = 150 mm) puis finalement un *beam expander*, noté  $BE_V$ , de grandissement  $\times 10$ . Ici, le faisceau vert a été agrandi d'un facteur  $\times 50$  et mesure environ 34 mm de diamètre.

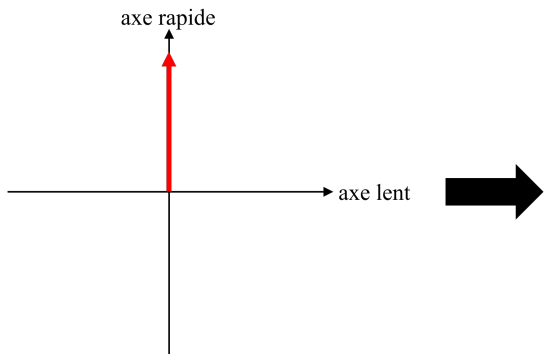
Ensuite, le cube séparateur noté  $PBS$  a pour but de recombinaison les deux voies d'excitation en réfléchissant le faisceau vert (polarisation  $s$ ) et en transmettant le faisceau rouge (polarisation  $p$ ). C'est donc à ce niveau que les lames à retard  $\lambda/2_R$  et  $\lambda/2_{V2}$  vont jouer leur rôle : l'orientation de leurs axes neutres respectifs va faire que la polarisation en sortie de lame sera plus ou moins proche de la composante  $p$  ou  $s$ , ce qui modifiera l'intensité transmise par le cube pour la voie rouge et réfléchie pour la voie verte. Je rappelle que pour une polarisation linéaire incidente, une lame demi-onde renvoie une polarisation linéaire symétrique par rapport à son axe neutre. De cette façon, une polarisation linéaire parallèle à un axe neutre de cette lame ne sera pas modifiée. La Figure 3.5 détaille le fonctionnement d'une lame demi-onde.

Avant d'arriver dans le microscope, les deux voies d'excitation passent par un diaphragme carré  $D$  afin d'obtenir des faisceaux de 15 mm de côté. Développée durant la thèse de Véronique Vienne (Vienne (2023)), cette étape permet de simuler une illumination dite *Flat-top*. En effet, le diaphragme  $D$  sélectionne la partie centrale de chacun des deux faisceaux gaussiens, c'est-à-dire là où l'intensité évolue très peu spatialement. De cette façon, l'échantillon est éclairé avec une intensité quasiment uniforme sur la totalité du champ de vue. En théorie, pour un faisceau de diamètre 34 mm centré sur  $D$ , on estime que l'intensité collectée en sortie du diaphragme est environ 30 % plus faible sur les bords

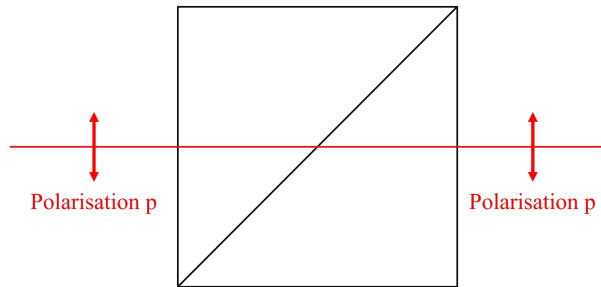
par rapport au centre.

**A**

Polarisation  $p$  non modifiée par la  $\lambda/2$

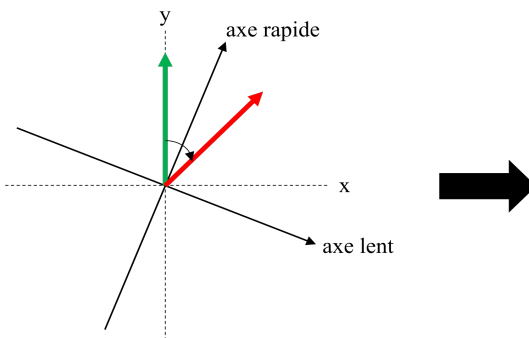


Polarisation  $p$  entièrement transmise par le cube PBS

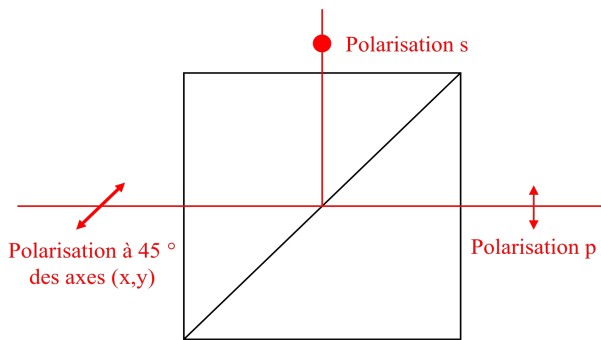


**B**

Polarisation  $p$  modifiée de  $45^\circ$  par la  $\lambda/2$



Polarisation incidente à moitié transmise en polarisation  $p$  et réfléchie en polarisation  $s$



**FIGURE 3.5** – Représentation de l'état de polarisation après passage par une lame demi-onde et un PBS. A) Cas de figure où la polarisation  $p$  est conservée jusqu'en sortie du cube séparateur en polarisation. B) Cas où la polarisation  $p$  est modifiée de  $45^\circ$  après la lame demi-onde puis est séparée en polarisation  $s$  et  $p$  de même intensité après le PBS.

Enfin, les faisceaux rouge et vert sont focalisés dans le plan focal arrière de l'objectif de microscope par la lentille  $L_1$  (focale = 300 mm). Ensuite, ils passent par un miroir dichroïque à bandes (*Chroma*, ZT405/488/561/640rpcv2-UF1), noté MD qui a pour but de séparer les photons d'excitation des photons émis par les fluorophores. Il est important de préciser que l'axe optique de  $L_1$  est volontairement décalé par rapport à celui de l'objectif. En effet, sur la Figure 3.3, nous voyons que les rayons qui arrivent sur l'objectif ne sont pas focalisés sur l'axe optique de ce dernier. Grâce à cette configuration, les rayons émergents de l'objectif sont collimatés et arrivent avec un angle  $\theta_i$  supérieur à l'angle critique sur l'échantillon, ce qui rend possible l'illumination par réflexion totale interne. Le décalage entre les axes optiques de  $L_1$  et l'objectif est contrôlé par la platine représentée en pointillés sur la Figure 3.4. Cette dernière nous permet de régler l'angle d'incidence et donc le confinement de l'onde évanescente. Au niveau de l'échantillon, les



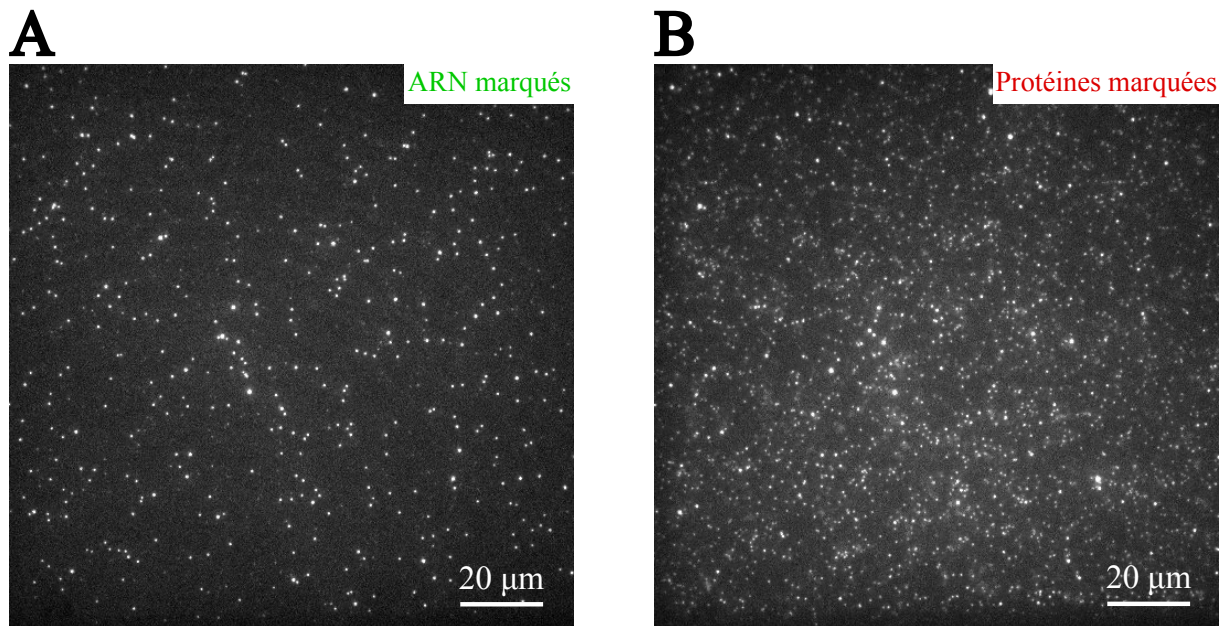
diamètres des faisceaux d'excitation sont divisés par un facteur  $100^1$  par rapport à leur taille en entrée de microscope, ce qui donne  $150\ \mu\text{m}$  : le diamètre du champ est donc légèrement supérieur aux  $137\ \mu\text{m}$  nécessaires pour éclairer la totalité du capteur de la caméra. L'échantillon est positionné sur un porte-échantillon qui repose lui même sur une platine de nano-positionnement contrôlable en X, Y et Z.

Une fois l'illumination par l'un des deux faisceaux lasers effectuée, le but est de faire une image de l'échantillon d'intérêt en dirigeant les photons émis par les marqueurs fluorescents vers la caméra : on parle désormais de la voie de détection. Dans un premier temps, ces photons vont être collectés par l'objectif de microscope *Olympus* si ceux-ci ont un angle d'incidence inférieur à l'angle limite caractérisé par l'ouverture numérique de 1,50. Le faisceau d'émission va être transmis par le même miroir dichroïque MD vers la lentille tube LT (focale 180 mm). En sortie de LT, le faisceau d'émission est focalisé dans le plan focal objet de la lentille  $L_2$  (focale = 100 mm) qui forme un télescope de grandissement  $\times 1$  avec la lentille  $L_3$  (focale = 100 mm). Ainsi, le grandissement final est égal à celui de l'objectif. Entre ces deux lentilles, le faisceau est collimaté et passe par une roue à filtres qui a pour but d'éliminer les traces résiduelles du faisceau d'excitation qui auraient été mal filtrées par le miroir MD. Chaque couleur possède son propre filtre d'émission (*Chroma*, HQ610/75m pour le vert et HQ665lp pour le rouge) et le contrôle de la roue à filtres est géré par un programme LabVIEW. Pour finir, le faisceau d'excitation est focalisé sur la caméra EMCCD (*Andor*, iXon Ultra 897). L'avantage de ce type de caméra est qu'elle possède un gain réglable, permettant d'augmenter l'intensité collectée et donc d'avoir accès à des signaux incidents très faibles. Les images faites sur la caméra sont directement observables en temps réel sur son logiciel associé.

C'est grâce au montage optique décrit ci-dessus que nous avons obtenu des images d'échantillons marqués dans le rouge ou dans le vert, comme nous pouvons le voir sur la Figure 3.6.

---

1. Le grandissement de l'objectif est de 60 pour une lentille tube de 180 mm, ce qui fait que la distance focale de l'objectif est de 3 mm. Comme la focale de  $L_1$  est de 300 mm, le diamètre des faisceaux est bien divisé par 100.

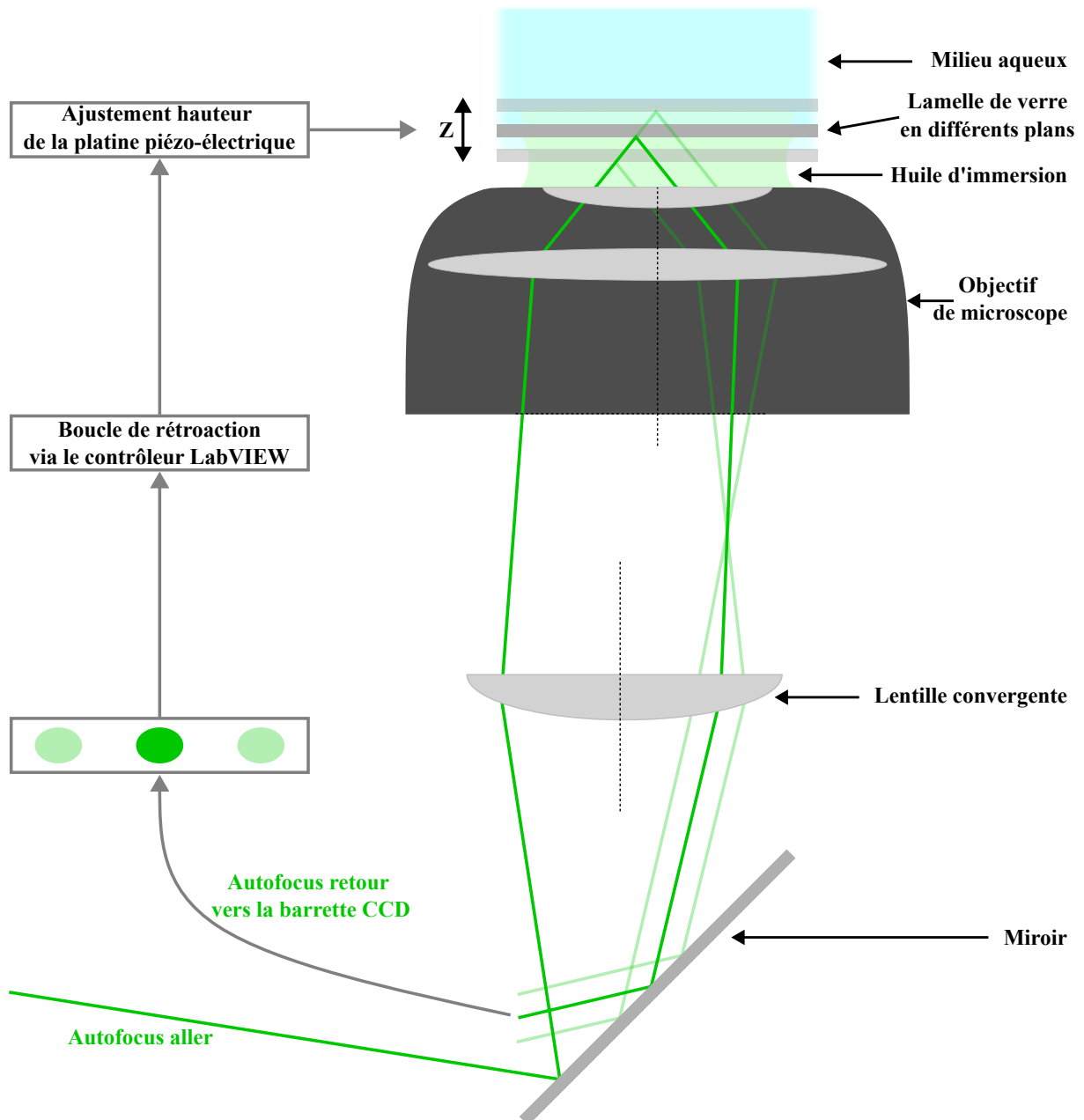


**FIGURE 3.6** – Exemples d’images expérimentales acquises en microscopie TIRF. A) Observation dans la voie d’excitation verte d’ARN simples brins marqués par un Atto565. B) Observation dans la voie d’excitation rouge de dimères marqués par un Atto647N, cette image a été prise durant une expérience d’assemblage où 50 nM en protéines avec un taux de marquage de 10 % ont été injectées.

### 3.1.2.2 Maintien de la focalisation dans le temps

Si l’on souhaite acquérir des images comme celles de la Figure 3.6 sur des durées de l’ordre de plusieurs minutes, il est important d’avoir un contrôle sur la netteté de l’image au cours du temps. En effet, la profondeur de champ étant inversement proportionnelle au carré de l’ouverture numérique, elle vaut seulement quelques centaines de nanomètres dans notre cas, ce qui rend la mise au point très sensible aux différentes perturbations extérieures. Afin de maintenir la mise au point dans le temps, on utilise la portion du faisceau laser vert qui a été extraite au début de son trajet optique par le cube  $\text{PBS}_V$  : on appelle ce faisceau l’autofocus. La polarisation de ce dernier étant orthogonale à celle du faisceau d’excitation, la propagation de l’autofocus est dirigée par plusieurs miroirs afin qu’il soit transmis par le cube PBS comme le faisceau rouge. Ensuite, jusqu’à l’échantillon, le trajet optique de l’autofocus est le même que pour les deux voies d’excitation. Le passage par le diaphragme D s’effectue par une fente située juste en dessous du carré qui permet la propagation des faisceaux principaux. Une fois que l’autofocus a été réfléchi par la lamelle de microscope, ce dernier emprunte le même trajet qu’à l’aller mais en étant légèrement décalé par rapport au faisceau incident. Ce décalage est directement lié à la distance entre l’objectif et la lamelle. Le faisceau retour de l’autofocus passe lui aussi par le cube PBS avant d’être dévié par un petit miroir vers une barrette CCD. Juste avant, l’autofocus passe par deux polariseurs consécutifs :  $\text{Po}_1$  afin de garder seulement la polarisation de l’autofocus et  $\text{Po}_2$  pour contrôler la puissance incidente. Ensuite, une

lentille convergente  $L_A$  focalise le faisceau sur la barrette et un filtre réjecteur de bandes  $N$  (*Semrock*, NF-633U-25) élimine les potentielles traces résiduelles du laser rouge qui a la même polarisation que l'autofocus. Ainsi, une modification de la distance objectif - lamelle entraînera le déplacement de l'autofocus le long de la barrette comme illustré en Figure 3.7.

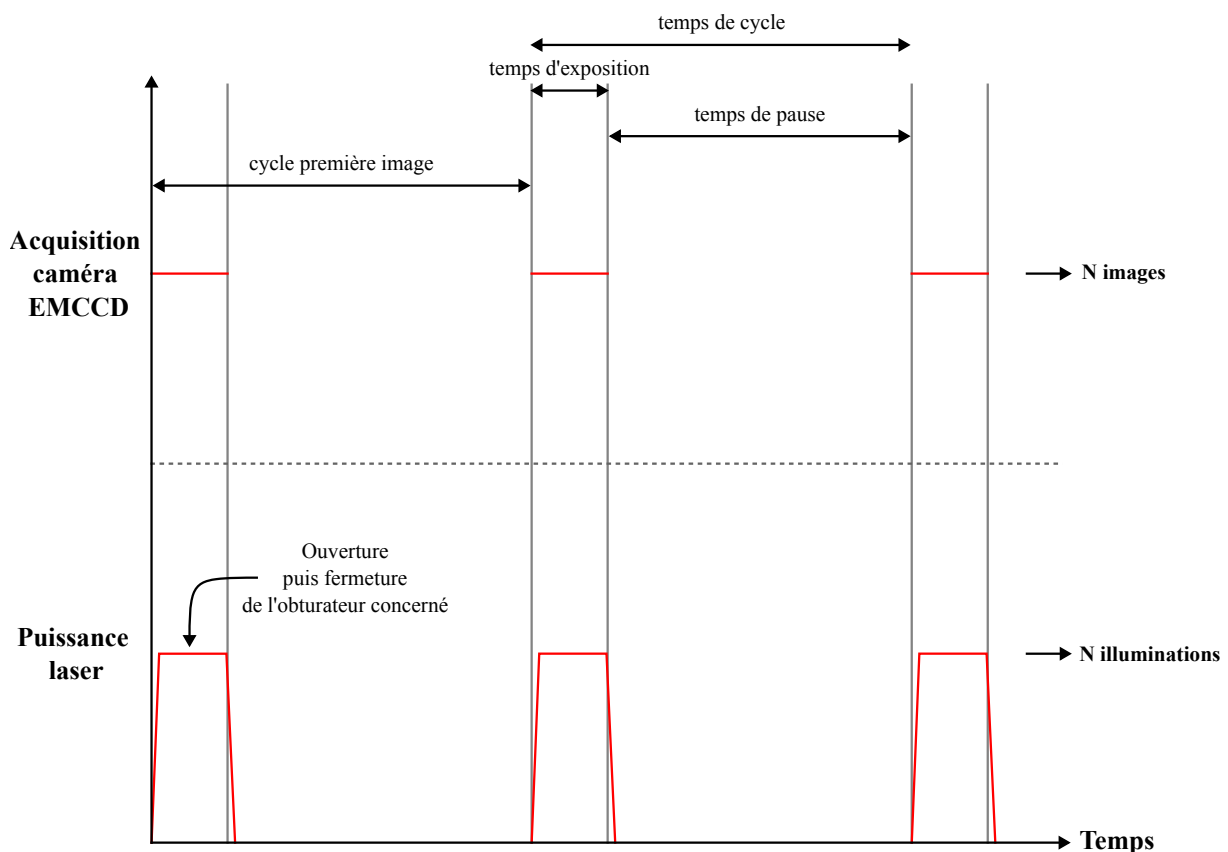


**FIGURE 3.7** – Schéma explicatif du maintien de la focalisation au cours du temps. Les deux traits noirs en pointillés représentent les axes optiques de la lentille  $L_1$  présente en amont du microscope et de l'objectif.

Une boucle de rétroaction est appliquée à la platine de nano-positionnement afin de contrôler le déplacement du plan de l'échantillon selon l'axe  $Z$  au cours du temps.

À l'aide d'un programme LabVIEW, une rampe de tension de 3 volts est envoyée à la platine dans le but d'entraîner un déplacement de cette dernière de 3  $\mu\text{m}$  selon l'axe Z. Au même moment, l'autofocus va se déplacer sur la barrette CCD ce qui permet d'établir une relation avec le positionnement en Z de la platine. Cette relation doit être linéaire afin de permettre une bonne rétroaction au cours du temps. Une fois cette calibration réalisée, il est possible d'enregistrer une première valeur consigne de la tension à appliquer sur la platine lorsque l'image est nette dans l'une des deux couleurs. Or, sachant que les lentilles souffrent d'un léger chromatisme, le passage à la deuxième couleur va entraîner une perte du plan de focalisation. Pour passer d'une voie d'excitation à une autre facilement et sans perdre la mise au point, une deuxième valeur de consigne est enregistrée. Ainsi, il est possible d'alterner l'utilisation des deux lasers pendant une acquisition tout en maintenant la netteté des images prises dans chaque voie au cours du temps.

### 3.1.2.3 Déroulement d'une acquisition



**FIGURE 3.8** – Cycle de synchronisation entre la caméra et l'obturateur de faisceau durant une acquisition. Ici seule la synchronisation avec l'obturateur présent devant le laser rouge est représentée mais le même mécanisme s'opère avec le laser vert.

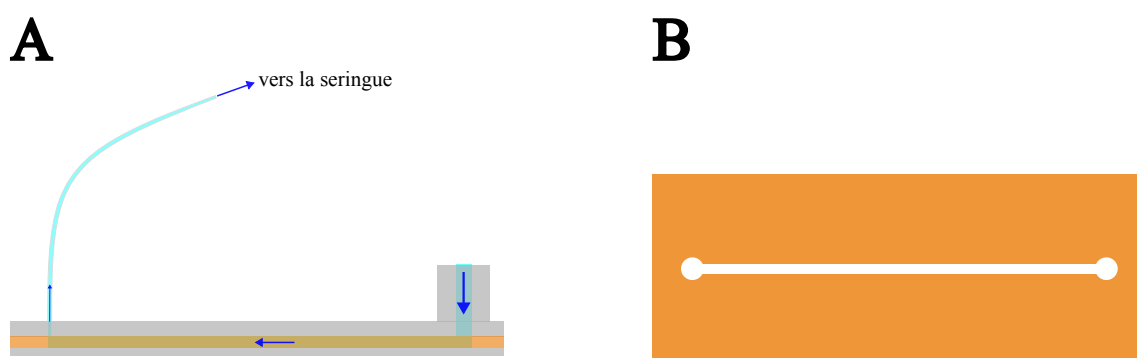
Les images (A) et (B) de la Figure 3.6 appartiennent à des séquences d'acquisition respectives de 5 et 400 images. Dans le cas de l'observation des protéines marquées, le

film a duré 10 minutes. Or, afin de limiter les effets du photoblanchiment (section 3.3.2.1), l'illumination des fluorophores ne se fait pas de manière continue et est contrôlée à l'aide d'obturateurs, notés O, qui sont positionnés au début de chaque voie d'excitation. De cette manière, les lasers sont en permanence allumés, mais ce sont les obturateurs qui vont être activés plus ou moins longtemps selon les durées d'illumination souhaitées. En parallèle, la caméra va acquérir une image à chaque fois que l'un des deux faisceaux se propage jusqu'à l'échantillon. Il est donc important de pouvoir synchroniser l'ouverture des obturateurs avec la prise d'images par la caméra. Pour y arriver, le contrôle d'une d'acquisition s'effectue également à l'aide d'un programme LabVIEW. La Figure 3.8 décrit le fonctionnement combiné de la caméra et des obturateurs dans le cas d'une séquence monocolor. Pour les expériences de microscopie de fluorescence du chapitre 5, les séquences utilisées sont explicitées dans chaque cas.

Le programme LabVIEW nous permet également de réaliser une acquisition bicolore en alternant entre le rouge et le vert voire en faisant deux images vertes entre chaque image rouge. Dans ce type de configuration, la roue à filtres est également synchronisée avec les obturateurs et la caméra.

### 3.1.3 La cellule micro-fluidique

En expliquant le principe de réflexion totale interne à la section 3.1.1, je précise que les échantillons biologiques sont déposés sur une lamelle de microscope. Cependant, pour pouvoir mener à bien une expérience d'assemblage, plusieurs étapes d'incubation et de rinçage doivent se succéder, rendant l'utilisation d'une simple lamelle de microscope très complexe. C'est pour cela que les expériences ont lieu dans des cellules micro-fluidiques faites maison, dont un schéma est représenté à la Figure 3.9.A.



**FIGURE 3.9** – Schéma d'une cellule micro-fluidique. A) Vue latérale des trois étages d'une cellule micro-fluidique où les flèches bleues indiquent la direction du flux. B) Vue de dessus du scotch double face utilisé.

Ces cellules possèdent trois niveaux distincts :

- Une lame de microscope percée à deux endroits afin d'insérer un tube de 1 mm de diamètre dans un premier trou, puis un réservoir dans le second. Les premières versions

de réservoirs ont été découpées à partir de pipettes Pasteur. Ils sont collés avec de la colle epoxy (*Loctite®*, *Double Bubble*) soit directement à la lame soit sur un petit cube de PDMS (Polydiméthylsiloxane) lui-même collé à la lame. Ces réservoirs sont remplacés au fur et à mesure par des cylindres creux en verre qui peuvent être collés par leur face inférieure, rendant le suivi du volume à injecter plus facile.

- Un scotch double face (*3M™*, 410M), découpé au FABLAB de CentraleSupélec, afin d'obtenir précisément un canal d'un millimètre de largeur reliant le réservoir au tube en plastique (Figure 3.9.B).

- Une lamelle de microscope (*Epredia™*) de 170  $\mu\text{m}$  d'épaisseur, 24 mm de largeur et 60 mm de longueur.

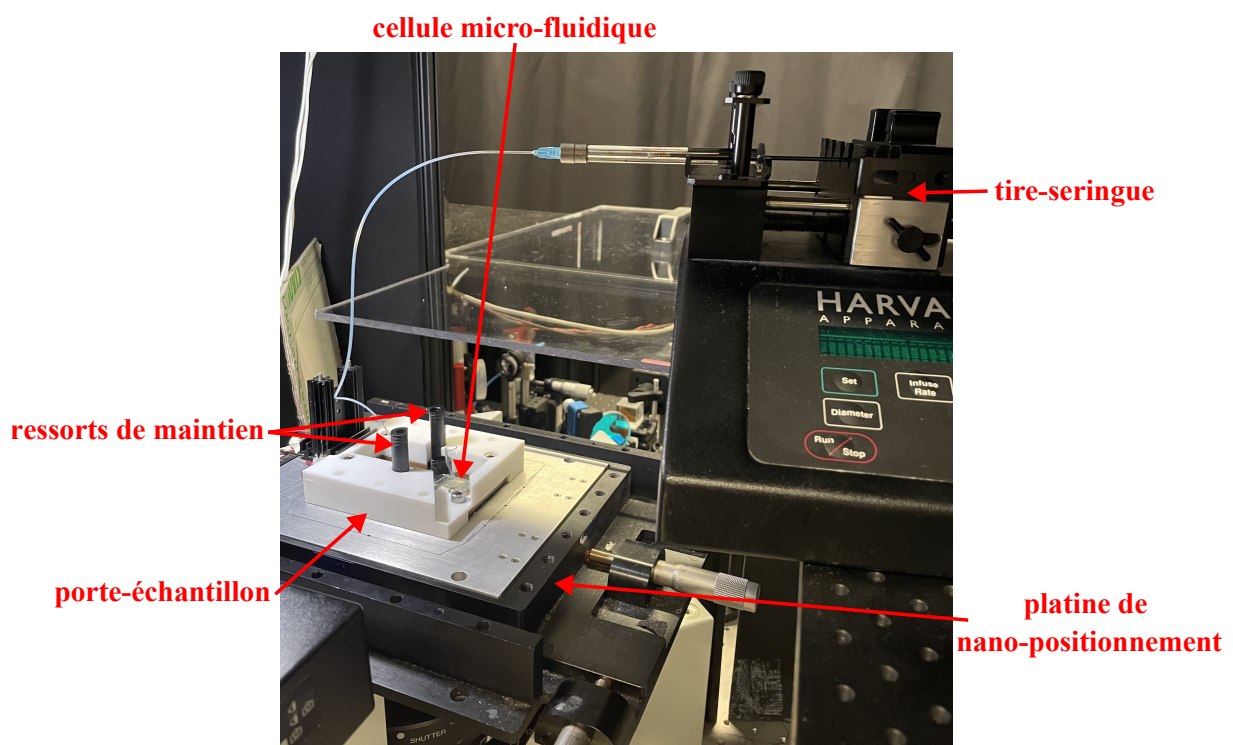


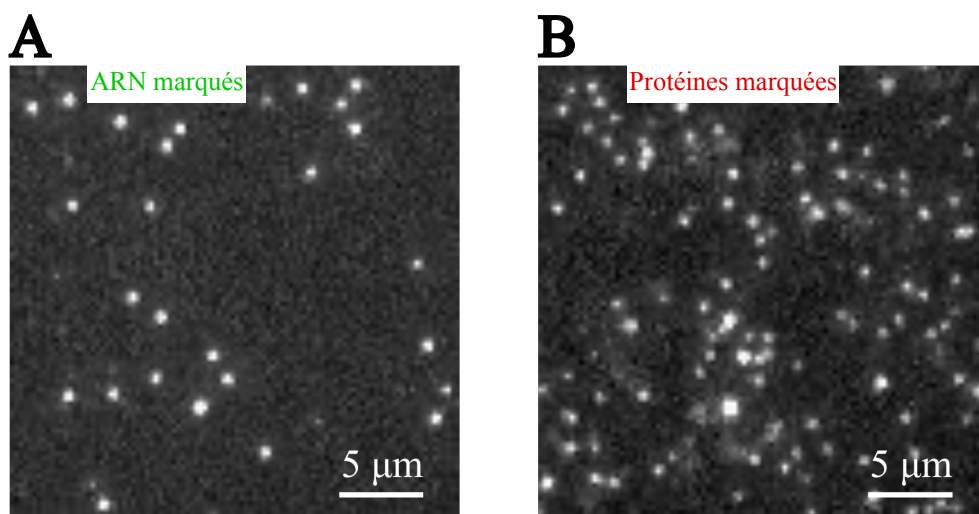
FIGURE 3.10 – Cellule micro-fluidique sur le microscope.

Une fois que les cellules sont montées, une seringue *Gas tight* (*Hamilton*, référence 074511) est insérée à l'autre extrémité du tube en plastique. La première solution aqueuse est ensuite injectée dans le réservoir puis déposée dans le canal grâce à la seringue. Généralement, les premières injections ont lieu sur la paillasse de préparation avant de se poursuivre sur le montage expérimental. Avant d'installer une cellule micro-fluidique au-dessus de l'objectif, il faut d'abord ajouter une goutte d'huile d'immersion. Ensuite, la cellule est déposée sur le porte-échantillon, fixée à l'aide de deux ressorts et la seringue est insérée dans le tire-seringue qui sera utilisé afin de déposer les dernières solutions aqueuses nécessaires aux expériences. Les différents éléments que je viens de mentionner pour l'installation des cellules micro-fluidiques sont représentés sur la Figure 3.10.

Les cellules sont montées généralement au début de chaque journée d'expérience puis démontées le soir même. Les lames de microscope et les seringues sont réutilisées d'une expérience à une autre à condition de les avoir soigneusement lavées en fin d'expérience avec des solutions d'eau ultra-pure et de RNase-Zap™ (*Invitrogen by Thermo Fisher Scientific*). La solution de RNase-Zap™ a pour but d'éliminer les potentielles ribonucléases qui sont connues pour dégrader les ARN. Par ailleurs, au plus tard à la veille d'une expérience, les lamelles de microscope sont passées au bain à ultrasons dans des solutions d'acétone pure puis d'eau ultra-pure afin qu'elles soient les plus propres possible. Cette étape est d'autant plus indispensable puisque nous réalisons des expériences à l'échelle de la molécule unique où chaque site d'intérêt doit être distingué de son plus proche voisin.

### 3.2 Fonctionnalisation des lamelles de microscope

La Figure 3.6 illustre parfaitement le principe d'une expérience en molécule unique qui est de pouvoir étudier plusieurs sites d'intérêt (ici l'ARN simple brin) individuellement. En effet, sur la Figure 3.6.A, chaque tache blanche correspond à un ARN pouvant constituer un site potentiel d'assemblage pour les sous-unités de capsid du CCMV. Dans la plupart des cas, on peut voir qu'il est clairement possible de distinguer deux points proches l'un de l'autre et donc deux sites d'ARN (3.11.A). En parallèle, la Figure 3.11.B montre cette fois-ci les sites d'accroche des dimères marqués du CCMV. Dans ce cas, on remarque que le champ de vue est clairement plus chargé en sous-unités de capsides qu'en ARN mais que là aussi, il est possible de différencier deux points proches l'un de l'autre. À ce stade, je précise que je ne m'intéresse pas à l'intensité de chaque point repéré, seulement à notre capacité à bien séparer deux sites d'accroche.



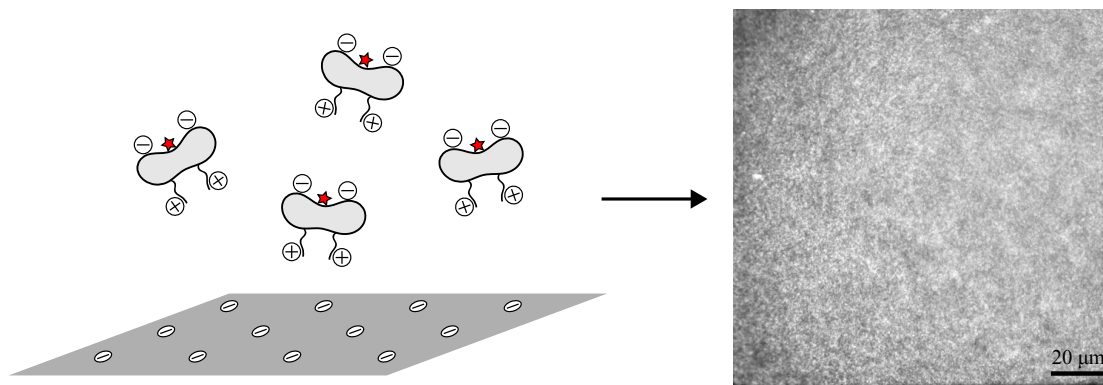
**FIGURE 3.11** – Mise en évidence de la distinction entre molécules uniques. A) Observation d'ARN simples brins marqués par un Atto565. B) Observation de dimères marqués par un Atto647N. Ces images proviennent d'une même zone agrandie de la Figure 3.6.

Afin d'obtenir ce type d'images, il est primordial de traiter chimiquement les lamelles de microscope. Ce traitement, auquel je fais référence sous le nom de "chimie de surface", a deux objectifs :

- **Fixer les ARN sur la surface de manière spécifique.**

- **Empêcher les accroches non spécifiques des dimères à la surface.** En effet, en solution, une lamelle de verre sera chargée négativement rendant ainsi les interactions électrostatiques possibles avec le bras N-terminal des sous-unités de capsid, plutôt chargé positivement, comme le montre la Figure 3.12. Dans ce cas, les interactions non spécifiques sont beaucoup trop nombreuses et rendent le champ de vue entièrement "blanc" : il est impossible de distinguer des accroches individuelles et donc de travailler à l'échelle de la molécule unique. Il est donc important de rendre la surface la plus neutre possible.

Pour répondre à ces deux problématiques, je commence par présenter en détails la chimie de surface à base de longues chaînes de poly-L-lysine avec laquelle nous travaillons, les tests effectués et les améliorations qui ont été envisagées afin de réduire le nombre d'accroches non spécifiques. Ensuite, je compare les performances de la chimie de surface à base de poly-L-lysine avec celles de trois autres méthodes que l'on peut trouver dans la littérature. Enfin, dans la dernière section je détaille le protocole d'hybridation de l'ARN simple brin qui nous permet de le marquer avec un Atto565 et de le fixer à la surface d'une lamelle de verre.



**FIGURE 3.12 – Observation des accroches non spécifiques des dimères marqués sur la lamelle de verre.** Des interactions électrostatiques ont lieu entre le bras N-terminal des protéines chargé positivement et la lamelle de verre chargée négativement en solution. L'image obtenue en illumination TIRF a été prise après avoir injecté 50 nM de protéines, dont 10 % sont marquées, mais en l'absence d'ARN.

Dans toute cette partie, je vais souvent mentionner deux solutions tampons dans lesquelles les protéines ont été diluées : le tampon B (100 mM MES, 10 mM NaCl à pH = 6,1) et le tampon B' (50 mM MES, 100 mM NaCl à pH = 6,1). De plus, l'analyse quantitative du nombre d'accroches non spécifiques a été faite avec le module *Trackmate* du logiciel de traitement d'images *ImageJ*.



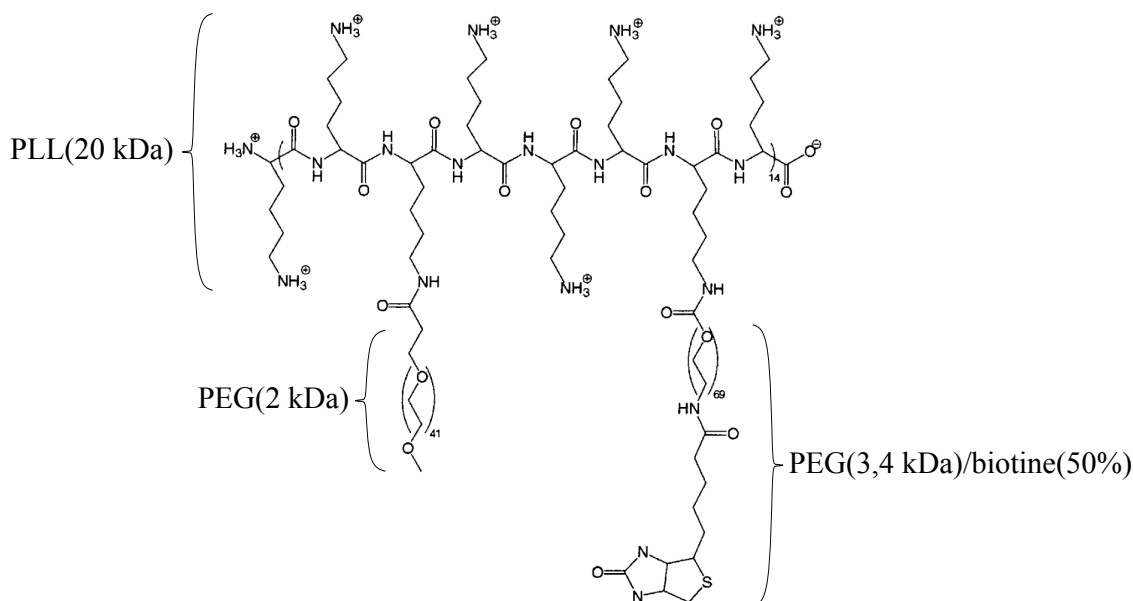
### 3.2.1 Focus sur la chimie de surface PLL-g-PEG

Pour les expériences d'assemblage réalisées au cours de ma thèse, j'ai utilisé une chimie de surface créée à partir de chaînes de poly-L-lysine et dont l'acronyme est le suivant : PLL-g-PEG pour *poly(L-lysine)-graft-poly(ethylene glycol)*. Commençons par regarder plus en détails la composition de cette méthode de fonctionnalisation. **Message au lecteur : les expériences décrites dans cette section sont assez techniques. La chimie de surface finalement utilisée est explicitée en toute fin de partie 3.2.1.4. Le protocole correspondant et un tableau résumé des différentes conditions expérimentales testées sont disponibles à l'annexe C.**

#### 3.2.1.1 Présentation de l'architecture

La chimie de surface de type PLL-g-PEG se base sur des interactions électrostatiques entre la longue chaîne de poly-L-lysine et la surface de verre. En effet, pour une solution dont le pH se situe aux alentours de la neutralité, le groupement amine de la chaîne latérale de chaque lysine sera chargé positivement alors que le verre sera lui chargé négativement (Perry et al. (2009)). De l'autre côté de la chaîne de PLL interagissant avec la lamelle de verre, du PEG, un composé électriquement neutre en solution, est greffé à certaines lysines. La fréquence d'apparition d'une chaîne de PEG est donnée par le paramètre  $g$  présent dans le nom de l'échantillon acheté. Par exemple, l'un des deux composants de base de la chimie de surface que nous utilisons est décrit selon la formule suivante : PLL(20)-g(3,5)-PEG(2). Ici il faut comprendre que la chaîne de PLL possède une masse molaire de 20 kDa et qu'en moyenne on retrouve un PEG de 2 kDa toutes les 3,5 lysines. Par ailleurs, cette première chaîne de PLL-g-PEG est mélangée avec une seconde qui présente une légère modification au niveau de certains PEG comme la formule suivante le décrit : PLL(20)-g(3,5)-PEG(2)/PEG(3,4)-biotine(20%), notée PLL-biotine. Pour ce second échantillon, la chaîne principale est conservée, mais un PEG de 2 kDa sur cinq est remplacé par un PEG plus long (3,4 kDa) et qui possède une modification de type biotine à son extrémité. Cette biotine est nécessaire pour fixer l'ARN simple brin spécifiquement à la lamelle comme je le présente dans la section 3.2.3. La Figure 3.13 montre la formule semi-développée d'une chaîne de PLL(20)-g(3,5)-PEG(2)/PEG(3,4)-biotine(50%).

Les deux échantillons de PLL-g-PEG et PLL-biotine sont achetés auprès de l'entreprise *SuSoS* qui les fournit en poudre. Les tests effectués durant la première partie de ma thèse ont été réalisés à partir des protocoles donnés par cette entreprise mais aussi présents dans la littérature. Je précise ici que pour les tests effectués de chimie de surface, les lamelles de microscope ont toujours été passées au préalable au *plasma cleaner* à oxygène (*electronic Diener, Plasma-Surface-Technology*). Ce premier traitement de 10 minutes nettoie une dernière fois les lamelles avant une expérience mais surtout il rend la surface plus réactive, favorisant les interactions avec la chimie de surface.



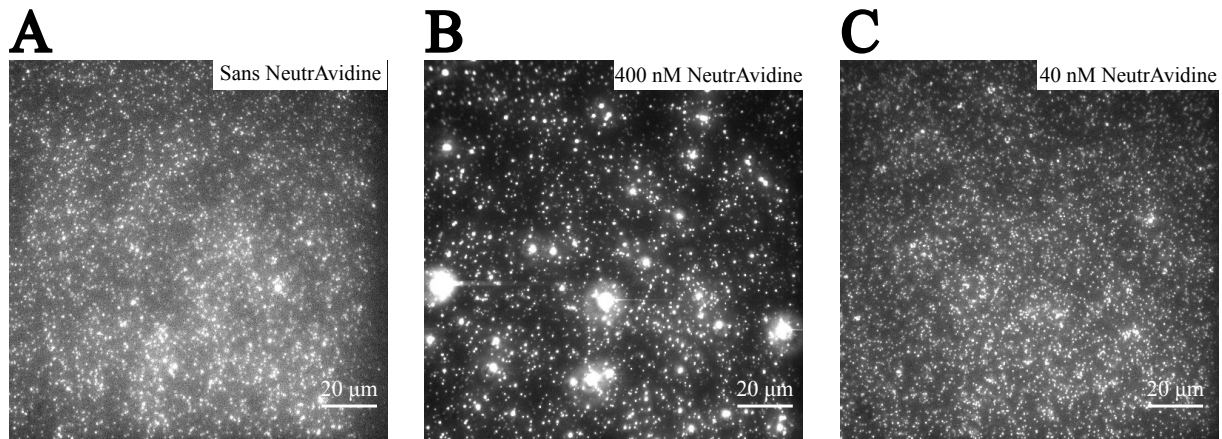
**FIGURE 3.13** – Formule semi-développée d'une chaîne de PLL-biotine. L'échantillon représenté est la PLL(20)-g(3,5)-PEG(2)/PEG(3,4)-biotine(50%) et la figure est adaptée de Huang et al. (2002)

Les différents tests présentés dans les sections suivantes ont pour but d'illustrer les problématiques qui ont été rencontrées tout au long de l'optimisation du protocole de chimie de surface. Les différentes méthodes employées pour améliorer les performances de la PLL-g-PEG ne sont pas détaillées chronologiquement parlant mais plutôt d'une façon à maintenir une certaine cohérence dans la réflexion. Je précise d'ores et déjà que ces tests ont été d'autant plus longs qu'il a été difficile d'être reproductible, ce qui peut principalement s'expliquer par une dépendance envers plusieurs paramètres. **Dans les trois sections suivantes, les champs de vue ont généralement été obtenus après une incubation de 2 minutes des protéines marquées, sauf indication contraire.**

### 3.2.1.2 Les premiers tests

En combinant les informations récoltées dans la littérature et données par *SuSoS*, deux premiers paramètres ont été mis en évidence : la concentration des échantillons et le tampon d'incubation. En effet, nous avons pu voir qu'il est conseillé de travailler avec des concentrations massiques comprises entre 0,1 et 1 g/L. En ce qui concerne les solutions tampons, il est possible de choisir entre le tampon appelé H1 (10 mM HEPES pour 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid à pH = 7,4) ou H2 (10 mM HEPES, 150 mM NaCl à pH = 7,4). Comme la plupart des articles trouvés décrivent plutôt des expériences avec des périodes d'incubation en absence de sel et que nos premiers tests n'ont pas montré de différences particulières entre les deux tampons, **nous avons commencé avec le tampon H1** (Huang et al. (2001), Huang et al. (2002), Pasche et al. (2003), Michel et al. (2005)). De plus, **pour tous les tests décrits dans la section 3.2.1,**

la PLL-g-PEG et PLL-biotine ont été incubées pendant 30 minutes puis le surplus est éliminé au cours d'une étape de rinçage réalisée dans le tampon d'incubation. Par ailleurs, chaque phase d'injection d'une espèce chimique a été forcément suivie d'une étape de rinçage avec le tampon adapté.

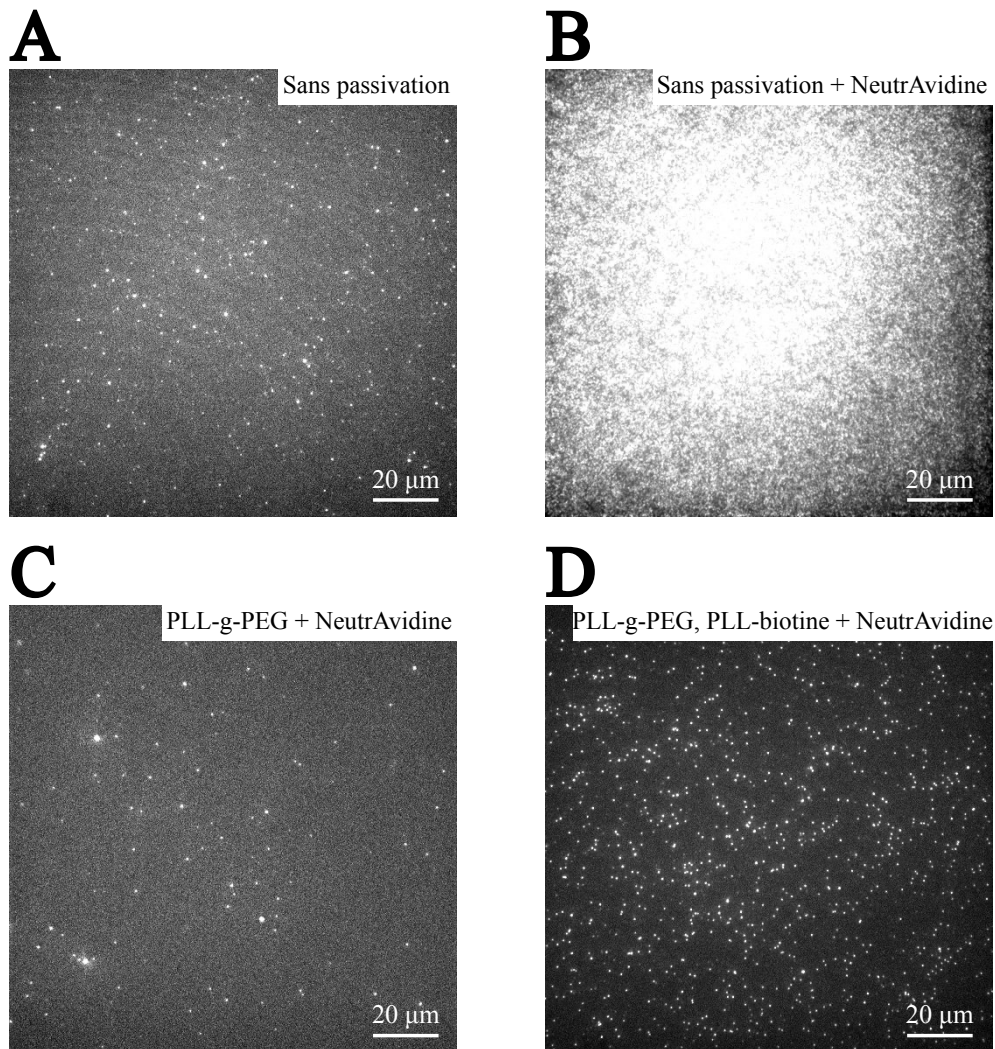


**FIGURE 3.14** – Influence du couple biotine-NeutrAvidine sur les accroches non spécifiques en protéines. A) Chimie de surface : 1 g/L en PLL-g-PEG. B-C) Chimie de surface : 1 g/L en PLL-g-PEG et PLL-biotine et respectivement 400 nM et 40 nM en NeutrAvidine. Les adsorptions non spécifiques proviennent d'une injection de 15 nM en protéines marquées et le tampon B a été utilisé pour les dilutions et rinçages.

Rapidement, nous avons pu voir que pour une même concentration en protéines marquées, ici 15 nM, une chimie de surface réalisée à partir de la PLL-g-PEG ou de la combinaison PLL-g-PEG / PLL-biotine ne donnait pas les mêmes résultats. En effet, sur la Figure 3.14.A, il est possible d'observer près de 2500 accroches sur une chimie de surface préparée seulement à partir de la PLL-g-PEG. En revanche, pour les Figures 3.14.B et C, on compte respectivement environ 2800 et 3500 points, avec notamment de très gros agrégats dans le cas de la Figure 3.14.B. Ces deux dernières couvertures ont été obtenues avec une chimie de surface comprenant à la fois 1 g/L en PLL-g-PEG et PLL-biotine. La seule différence entre les cas B et C repose sur la concentration en NeutrAvidine (*Thermo Scientific*, référence 31000) qui vaut respectivement 400 et 40 nM. La NeutrAvidine est une protéine d'environ 60 kDa qui possède une très grande affinité avec la biotine et qui a pour but de former une liaison entre celle fixée à la surface à partir de la PLL-g-PEG et celle que l'on retrouve au niveau de l'extrémité 5' de l'ARN simple brin<sup>2</sup> (voir section 3.2.3). En d'autres termes, le couple biotine-NeutrAvidine est indispensable afin de rendre immobile l'ARN du CCMV durant une expérience d'assemblage. Pour réduire le plus possible les interactions avec les protéines de capsid, le pH du tampon d'assemblage, qui est aussi celui utilisé pour diluer la NeutrAvidine, est égal au point isoélectrique de cette dernière, à savoir 6,1. De cette façon, la NeutrAvidine sera globalement neutre en solution et sera moins encline à interagir avec d'autres espèces chargées. Plus globalement,

2. Au maximum, quatre biotines peuvent se fixer à une même NeutrAvidine

les tests décrits dans les sections suivantes ont été réalisés avec **une concentration en NeutrAvidine de 400 nM et une durée d'incubation de 10 minutes dans le tampon B ou B', sauf indication contraire.**



**FIGURE 3.15** – Mise en lumière des accroches non spécifiques de la NeutrAvidine. A) Lamelle de verre non passivée. B) Lamelle de verre non passivée et NeutrAvidine. C) Chimie de surface : 1 g/L en PLL-g-PEG et NeutrAvidine. D) Chimie de surface : 1 g/L en PLL-g-PEG,  $10^{-4}$  g/L en PLL-biotine et NeutrAvidine. Les taches blanches correspondent à des Atto565-biotinylés, injectés à 1 nM.

Pour compléter l'observation faite sur l'influence du couple biotine-NeutrAvidine, j'ai souhaité comprendre un peu mieux le comportement de la NeutrAvidine. Sur la Figure 3.15.A, on peut compter environ 250 points d'accroches sur une lamelle de verre n'ayant pas été passivée. Ces points proviennent de l'injection à 1 nM du marqueur fluorescent Atto565 ayant une modification de type biotine. En revanche, lorsque l'on injecte au préalable de la NeutrAvidine, toujours sur une lamelle non passivée, les Atto565-biotinylés recouvrent entièrement le champ de vue comme l'illustre la Figure 3.15.B. Par ailleurs, si la surface de verre est traitée avec 1g/L en PLL-g-PEG avant d'injecter la NeutrAvidine,

le nombre d'Atto565-biotinylés fixés sur cette dernière chute à 75 points ( Figure 3.15.C). Pour finir, la Figure 3.15.D nous montre que l'utilisation d'une chimie de surface préparée à partir de PLL-g-PEG et PLL-biotine permet d'obtenir une couverture très satisfaisante en Atto565-biotinylé après que la NeutrAvidine a été injectée. Ainsi, ces différentes observations nous permettent d'arriver à la conclusion suivante : **la NeutrAvidine peut se fixer à une lamelle de verre nue mais pas à une surface correctement passivée et ne présentant pas de biotines. De cette manière, on peut penser que les potentiels défauts de la PLL-g-PEG qui permettent les accroches non spécifiques des protéines de CCMV sont plus petits que la NeutrAvidine, sauf dans le cas d'une très mauvaise passivation.** Dans les quatre conditions expérimentales de la Figure 3.15, le tampon B a été utilisé et les marqueurs fluorescents ont été incubés pendant 2 minutes avant rinçage.

Ainsi, les premiers tests réalisés avec la chimie de surface PLL-g-PEG et PLL-biotine nous ont permis de comprendre un peu mieux l'accroche non spécifique des protéines sur une lamelle de verre. En effet, avec une passivation simplement réalisée à partir de la PLL-g-PEG, les sous-unités de capsid peuvent se fixer à la surface, mettant en lumière de potentiels défauts de la chimie de surface. De plus, ces accroches semblent augmenter lorsque nous rajoutons le couple biotine-NeutrAvidine, et ce même si le choix du pH de la solution tampon a été fait en fonction du point isoélectrique de la NeutrAvidine.

Malgré tout, la Figure 3.14 ne montre pas une situation catastrophique en termes d'accroches de protéines. En effet, mise à part pour la condition B, où de gros agrégats sont présents, les autres cas montrent qu'il semble assez simple de détecter individuellement des points fluorescents. Nous sommes très loin des observations expérimentales faites en l'absence de passivation (Figure 3.12). En réalité, ces deux figures se différencient également par l'injection des protéines. Pour la Figure 3.14, je n'ai pas injecté directement 15 nM en protéines marquées, mais d'abord 150 pM puis 1,5 nM. En d'autres termes, le fait d'injecter au préalable une solution moins concentrée en protéines permettrait de passiver une seconde fois la surface. Je m'intéresse à cette possibilité dans la section suivante.

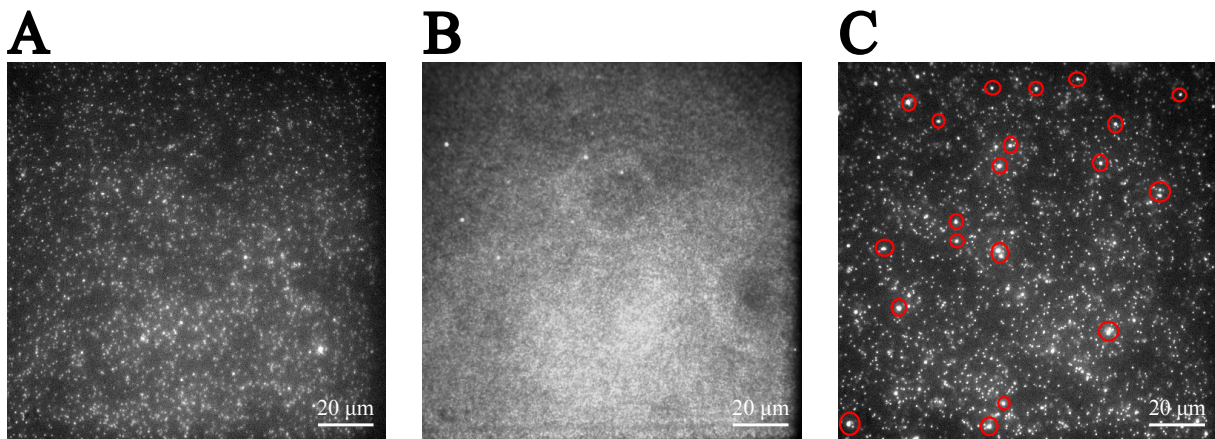
### 3.2.1.3 Est-il possible de combler les trous de la chimie de surface ?

Après quelques tests, il nous a semblé évident que la chimie de surface à partir de 1 g/L en PLL-g-PEG et PLL-biotine nécessitait quelques ajustements et notamment un agent passivant supplémentaire. Cette méthodologie, utilisée dans la plupart des protocoles de chimie de surface, a pour but de combler les éventuels défauts de passivation par une autre espèce chimique. Ici, je présente quatre solutions qui ont été testées pour améliorer les performances de la chimie de surface PLL-g-PEG / PLL-biotine. Parmi ces quatre pistes, les trois premières ont été réalisées en début de thèse, au moment où j'avais beaucoup de mal à trouver un protocole efficace. En revanche, la quatrième a été essayée plus tard, au

moment où la chimie de surface montrait de nouveau quelques signes de faiblesse.

*- Les sous-unités non marquées de CCMV :*

Une première idée a été d'utiliser des sous-unités de capsid non marquées. En effet, comme je l'ai mentionné à la fin de la section 3.2.1.2, les tests préliminaires nous ont permis de remarquer que le nombre d'accroches non spécifiques observées pour 15 nM en protéines marquées étaient plus faibles après avoir injecté au préalable deux solutions de protéines marquées concentrées à 150 pM puis 1,5 nM. Sur la Figure 3.16.A, près de 2300 points sont détectés, ce qui est très proche de ce que nous avons observé avec la Figure 3.14.A. En revanche, lorsque 15 nM en protéines marquées ont été injectées directement, nous pouvons voir que les accroches non spécifiques recouvrent la totalité du champ de vue (Figure 3.16.B). Ici, la chimie de surface a été préparée à partir de 1 g/L en PLL-g-PEG et sans NeutrAvidine.



**FIGURE 3.16 – Effet d'une passivation supplémentaire à partir de sous-unités.** A) Injections successives en protéines marquées : 150 pM, 1,5 nM puis 15 nM. B) Injection unique à 15 nM en protéines marquées. C) Injection de 12 nM en protéines non marquées puis 15 nM en protéines marquées : les cercles rouges mettent en évidence des agrégats. Les trois passivations ont été préparées à partir de 1 g/L en PLL-g-PEG sans NeutrAvidine.

Pour ces tests, le taux de marquage des protéines est d'environ 5 %. Cela signifie qu'en injectant 150 pM puis 1,5 nM en protéines marquées, les concentrations totales en protéines sont respectivement de 3 et 30 nM. Afin d'éliminer les accroches non spécifiques induites par les deux premières injections en protéines marquées, nous avons utilisé un lot de protéines non marquées à la place. Par exemple, pour la Figure 3.16.C, la chimie de surface a été préparée à partir de 1 g/L en PLL-g-PEG, puis 12 nM en protéines non marquées ont été incubées pendant 2 minutes avant de passer au lot marqué. Ici, pour 15 nM en protéines marquées, on retrouve environ 1800 points, ce qui est légèrement moins que pour les Figures 3.14.A et 3.16.A. Néanmoins, on observe de nombreux agrégats à différents endroits du champ de vue comme les différents cercles rouges le montrent. Par ailleurs, certains agrégats étaient déjà visibles sur les Figures 3.14.A et 3.16.A. Ces derniers

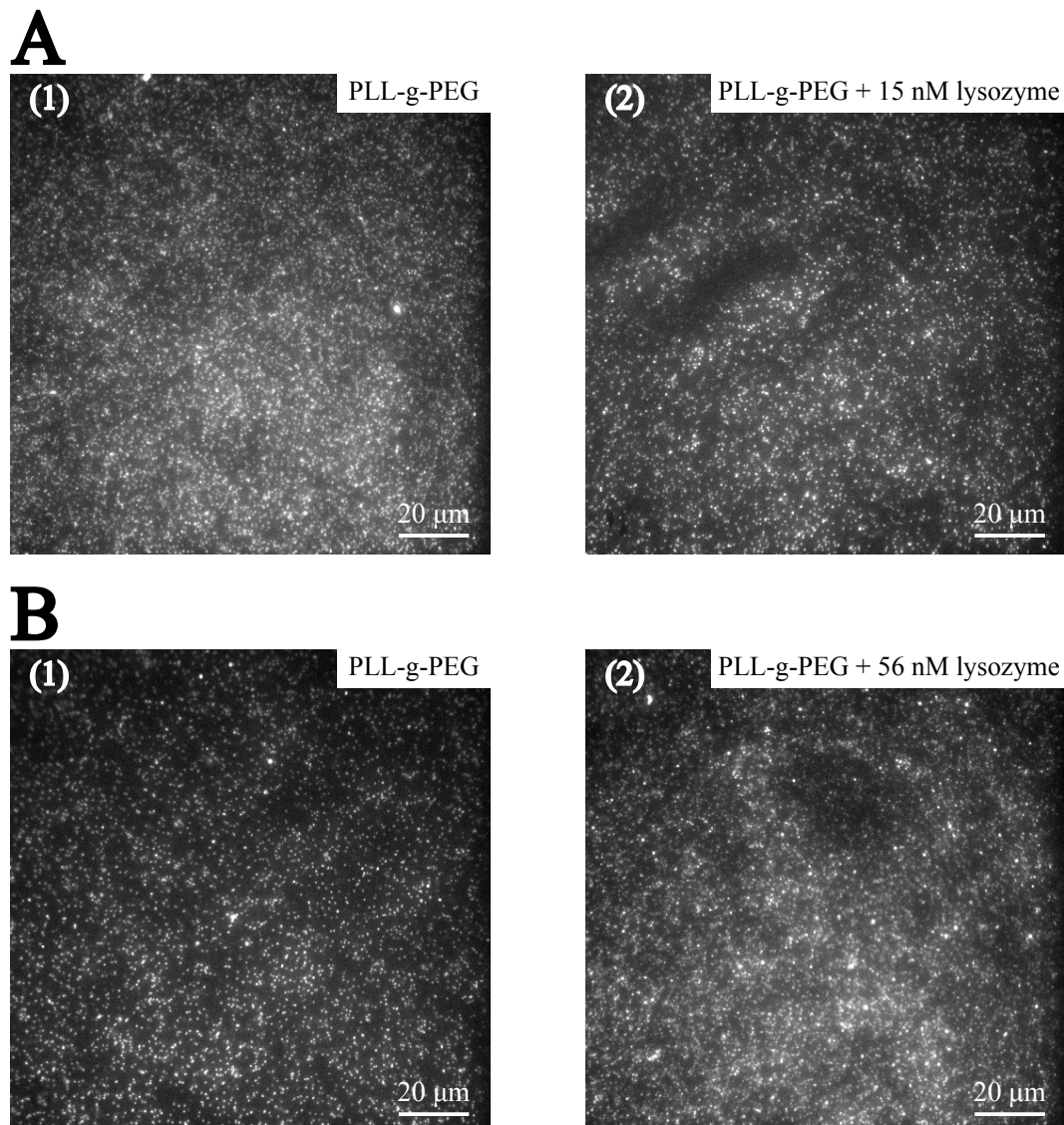
peuvent être le signe d'une interaction entre protéines marquées sur la surface, favorisée par les protéines non marquées déjà présentes. Or, **comme nous ne souhaitons pas altérer le comportement des protéines durant l'assemblage de capsides autour de leur ARN simple brin, cette piste n'a pas été exploitée plus en détails.**

Cependant, comme la concentration en sous-unités est un paramètre essentiel dans l'assemblage de capsides virales, nous ne pouvons pas continuer à injecter des concentrations croissantes en protéines marquées afin de réduire le nombre d'accroches non spécifiques. En effet, l'interprétation des résultats serait assez complexe. Ainsi, il est important de trouver une autre façon pour corriger les défauts de passivation de la PLL-g-PEG.

- Le Lysozyme :

En parallèle des essais prometteurs mais potentiellement risqués avec les protéines non marquées, nous avons utilisé une autre protéine, le lysozyme (masse molaire de 14 kDa). Cette protéine possède un point isoélectrique de 11, ce qui fait que globalement, elle possède une charge positive dans une solution tampon de pH 6,1. En théorie, le lysozyme peut tout à fait interagir avec le verre chargé négativement. Sur la Figure 3.17, je compare la couverture provenant d'une injection de 15 nM en protéines marquées entre deux passivations réalisées à partir de la PLL-g-PEG et deux autres sur lesquelles du lysozyme a été ajouté en plus.

Pour les deux configurations classiques (Figure 3.17.A.1 et B.1), on compte respectivement près de 4800 et 3500 points pour 0,75 g/L et 1 g/L en PLL-g-PEG. En comparaison, lorsque nous rajoutons du lysozyme en amont des protéines, respectivement 15 nM et 56 nM, le nombre d'accroches non spécifiques n'évolue que très peu : environ 3500 points pour 0,75 g/L en PLL-g-PEG contre 4400 points pour 1 g/L en PLL-g-PEG (Figures 3.17.A.2 et B.2). Même si nous sommes assez loin de la couverture totale en protéines marquées observées à la Figures 3.16.B, observer autant d'accroches sans biotine ni NeutrAvidine reste trop élevé. De plus, les différences de près de 1000 points observées, pour chacune des concentrations en PLL-g-PEG ne semblent pas provenir, à mon avis, d'un quelconque effet du lysozyme mais plutôt de la variabilité inhérente à la passivation à partir de la PLL-g-PEG. Ce manque de reproductibilité d'une journée à une autre est mis en évidence si l'on compare les Figures 3.14, 3.16 et 3.17 entre elles. En effet, pour 1 g/L en PLL-g-PEG et une injection en trois étapes des protéines marquées (concentrations de 150 pM, 1,5 nM puis 15 nM), nous passons de 2300 points (Figure 3.14.A) à 2500 points (Figure 3.16.A) puis 3500 points (Figure 3.17.B.1). Ces différentes observations semblent donc bien nous indiquer que **l'utilisation de lysozyme ne permet pas d'améliorer les défauts de passivation de la PLL-g-PEG.** Ici, le lysozyme a été incubé pendant 2 minutes puis rincé, avant que 15 nM en protéines marquées ne soient injectées. Toutes ces étapes ont eu lieu dans le tampon B.



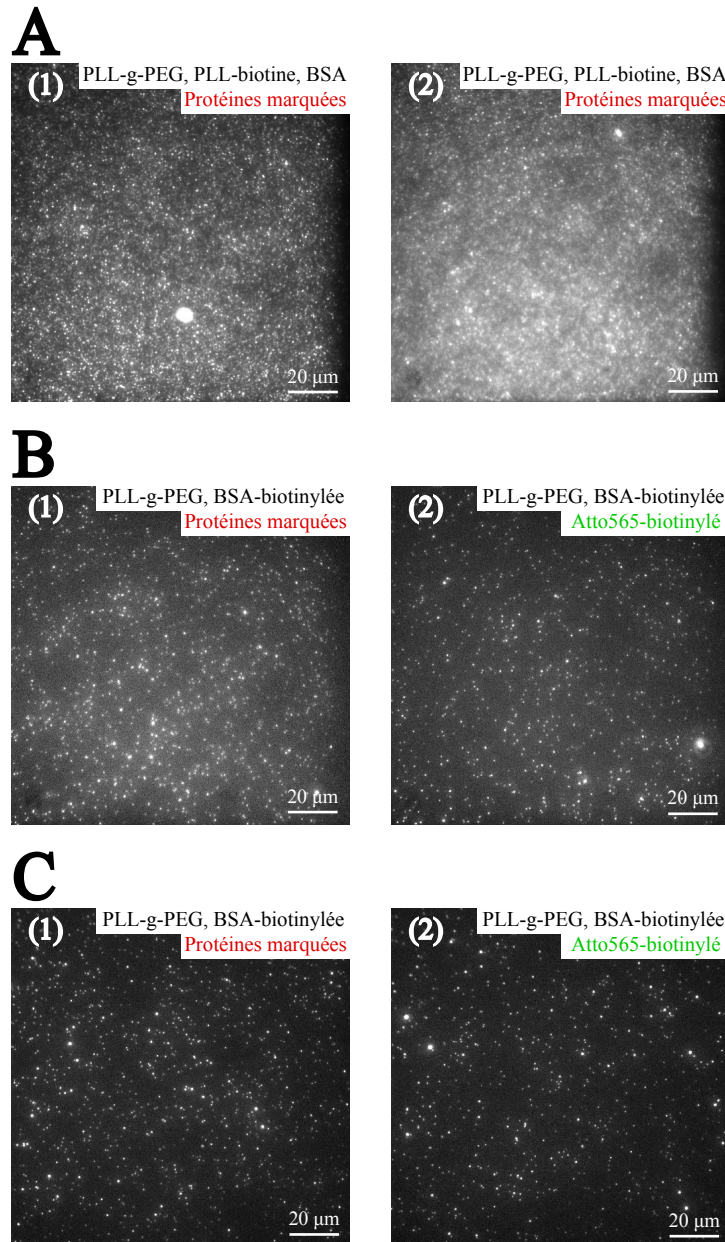
**FIGURE 3.17** – Effet d'une passivation supplémentaire à partir de lysozyme. A) Chimie de surface : 0,75 g/L en PLL-g-PEG (1) sans lysozyme et (2) avec 15 nM en lysozyme. B) Chimie de surface : 1 g/L en PLL-g-PEG (1) sans lysozyme et (2) avec 56 nM en lysozyme. Le lysozyme a été incubé pendant 2 minutes dans le tampon B, juste avant l'injection de protéines marquées. Les couvertures correspondent à une injection de 15 nM en protéines marquées dans les quatre cas.

Par ailleurs, je précise qu'au début de l'année 2023, j'ai utilisé le lysozyme une seconde fois pour réduire les adsorptions non spécifiques de 5 nM en protéines marquées sur une chimie de surface que l'on pensait être bien optimisée. Ces tests ont été réalisés dans le tampon B', avec des concentrations plus élevées en lysozyme (700 nM - 7 μM - 70 μM et 700 μM) et des durées d'incubation de 30 minutes. Au préalable, une chimie de surface de type PLL-g-PEG et PLL-biotine avait servi pour former une première couche de passivation de la lamelle. Toutefois, aucune différence notable n'avait été observée entre des conditions expérimentales avec et sans lysozyme.



- *La BSA (Bovine Serum Albumin)* :

La troisième solution est beaucoup plus classique en matière de passivation puisqu'il s'agit de la BSA, une protéine d'environ 66 kDa.

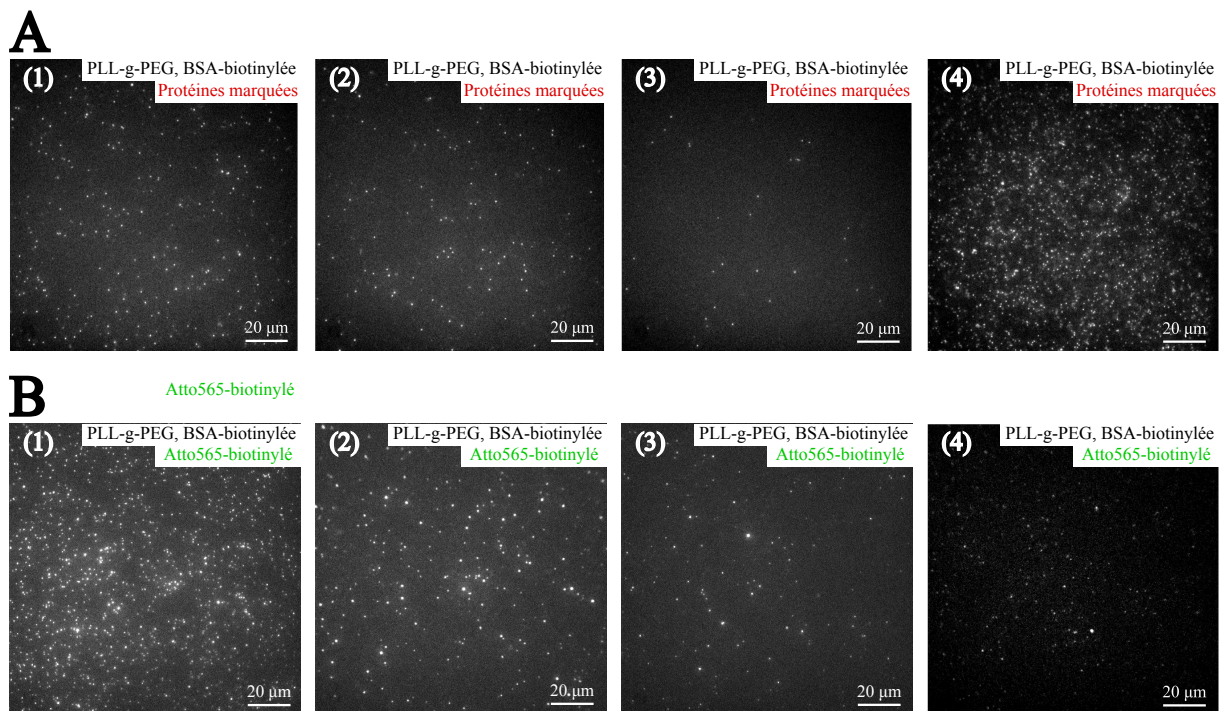


**FIGURE 3.18** – Effet d’une passivation supplémentaire à partir de BSA et BSA-biotinylée. A) Chimie de surface : 1 g/L en PLL-g-PEG, 0,5 g/L en PLL-biotine, NeutrAvidine et 10 nM en BSA ayant été incubée (1) 4 heures et (2) 6 heures. B-C) Chimie de surface : 1 g/L en PLL-g-PEG, 10 nM en BSA-biotinylée ayant été incubée 4 heures et NeutrAvidine. Figures A.1 et A.2 : accroches non spécifiques pour 1,5 nM en protéines marquées sur deux cellules. Figures B.1 et C.1 (respectivement B.2 et C.2) : accroches non spécifiques pour 1,5 nM en protéines marquées et interactions spécifiques entre Atto565-biotinylé et NeutrAvidine, sur un même champ de vue.

Dans un premier temps, j’ai évalué la capacité de 10 nM en BSA à réduire les adsorp-

tions non spécifiques en protéines marquées sur une chimie de surface préparée à partir de 1 g/L en PLL-g-PEG, 0,5 g/L en PLL-biotine et en présence de NeutrAvidine. Sur les Figures 3.18.A.1 et A.2, les accroches recouvrent quasiment la totalité des champs de vue pour une injection en deux étapes des protéines marquées (concentrations de 150 pM puis 1,5 nM) et ce malgré une incubation de 4 et 6 heures respectivement pour la BSA.

Dans un second temps, la PLL-biotine a été remplacée par de la BSA-biotinylée (10 nM) qui a été incubée pendant 4 heures et rajoutée 1 heure après l'injection de 1 g/L en PLL-g-PEG. La NeutrAvidine est toujours présente et la concentration en protéines marquées est de 1,5 nM (injectées après une première étape à 150 pM). Les Figures 3.18.B.1 et C.1 montrent deux champs de vue ayant été observés sur deux cellules micro-fluidiques mais pour les mêmes conditions expérimentales. Près de 1000 et 800 points ont été détectés, ce qui constitue une belle amélioration par rapport à la Figure 3.18.A. De plus, les Figures 3.18.B.2 et C.2 nous indiquent qu'en injectant 100 pM en Atto565-biotinylé près de 1000 et 700 points sont visibles respectivement. La statistique est convenable mais il est intéressant de voir s'il est possible de l'augmenter légèrement pour espérer avoir un plus grand nombre d'accroches en ARN durant une expérience d'assemblage.



**FIGURE 3.19 – Évaluation des performances d'une passivation à partir de PLL-g-PEG et BSA-biotinylée.** A) (1-2-3) Accroches non spécifiques issues de 1,5 nM en protéines marquées, (4) Couverture pour 5 nM en protéines marquées durant une expérience d'assemblage. B) Interactions spécifiques entre (1-2-3) les Atto565-biotinylés et la NeutrAvidine, (4) les ARN marqués et la NeutrAvidine : dans chaque cas, les champs de vue sont les mêmes que ceux de A. Pour les quatre cellules, la passivation utilisée est la suivante : 1 g/L en PLL-g-PEG, 100 nM en BSA-biotinylée incubée pendant 6 heures et avec NeutrAvidine.

Après ces premiers essais, j'ai donc voulu voir si augmenter la concentration en BSA-biotinylée de 10 à 100 nM permettait d'observer un plus grand nombre d'Atto565-biotinylés tout en conservant les bonnes performances de passivation observées aux Figures 3.18.B et C. De plus, l'incubation de la BSA-biotinylée est passée de 4 à 6 heures et la concentration en Atto565-biotinylé est désormais de 1 nM afin de favoriser son accroche à la surface. Les Figures 3.19.A.1-2-3 montrent que les adsorptions non spécifiques provenant de 1,5 nM en protéines marquées (injectées après une première étape à 150 pM) sont toujours très faibles et semblent même avoir diminué puisqu'on compte respectivement 286 - 142 et 35 points. Ces résultats nous indiquent que la stratégie de passivation employée est très efficace vis-à-vis des accroches non spécifiques en protéines. En revanche, il semblerait que l'augmentation de la concentration en BSA-biotinylée n'ait pas eu les effets escomptés sur le nombre d'accroches en Atto565. En effet, sur les Figures 3.19.B.1-2-3, on trouve respectivement 960 - 263 et 114 points. L'adsorption spécifique des marqueurs fluorescents biotinylés n'a pas trop évolué et surtout, la reproductibilité est difficile à obtenir.

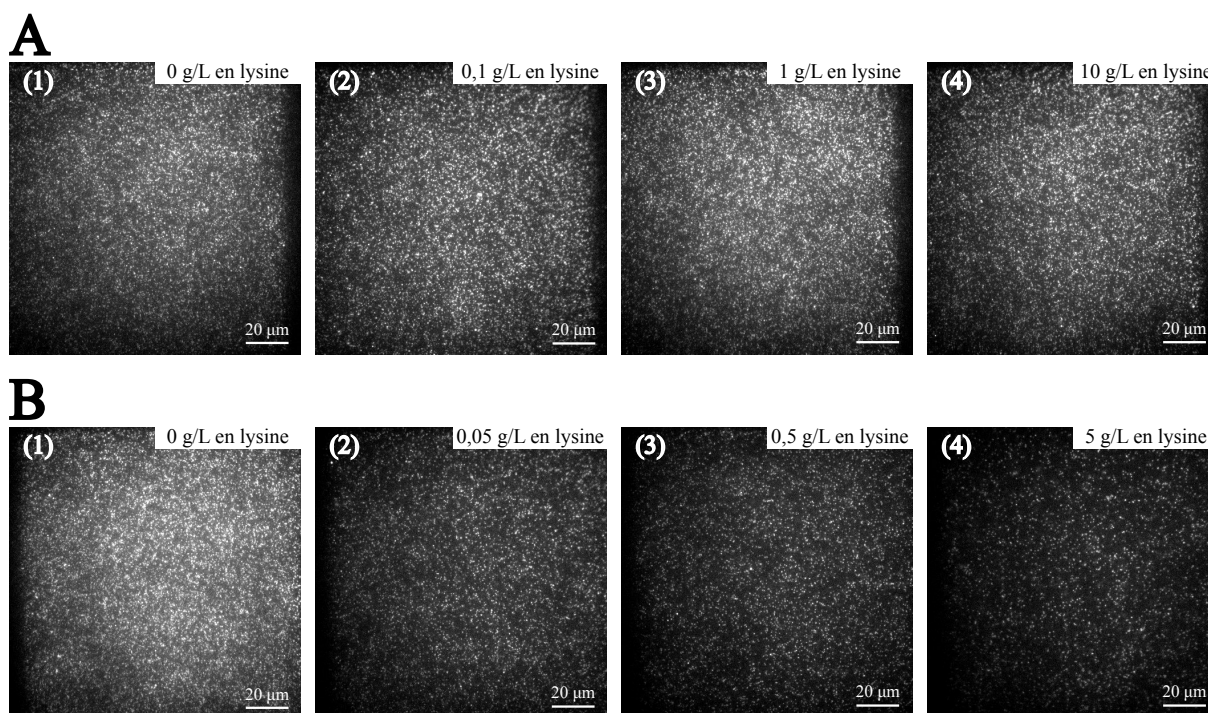
Pour résumer, la passage à la BSA-biotinylée semble montrer que l'un des principaux problèmes est bien la concentration en biotine, et *a fortiori*, le nombre d'accroches en NeutrAvidine. **Cependant, la BSA-biotinylée ne peut pas être considérée comme une solution viable, puisque son interaction avec la surface semble très aléatoire.** Par ailleurs, un test a été effectué en condition d'assemblage avec une chimie de surface préparée à partir de 1 g/L en PLL-g-PEG et 100 nM en BSA-biotinylées. Sur la Figure 3.19.A.4, on compte près de 1200 points pour une injection directe à 5 nM en protéines marquées, ce qui est vraiment très bien. En revanche, la Figure 3.19.B.4 montre que seulement 100 ARN simples brins marqués avec un Atto565 se sont fixés à la surface. Le nombre d'ARN doit être bien plus important compte tenu du champ de vue de près de 140  $\mu\text{m}$  de côté que nous avons à notre disposition. Je précise que dans toute cette section, les différentes dilutions ont été effectuées dans le tampon B.

#### - La lysine :

L'utilisation de la lysine est un peu différente par rapport aux essais présentés jusqu'à présent. En effet, cette piste a été envisagée au début de l'année 2023 quand le protocole de passivation, que l'on pensait plutôt fiable, a dû être repensé. Néanmoins, vu qu'il s'agit d'une volonté de réduire les adsorptions non spécifiques en protéines marquées, les tests effectués méritent d'être détaillés.

Cette piste se base sur le même principe que l'utilisation de protéines non marquées et de lysozyme puisque la lysine, à pH proche de la neutralité, est plutôt chargée positivement. En réalité, son mécanisme d'interaction avec la surface chargée négativement est similaire à celui de la poly-L-lysine, sauf qu'en étant beaucoup plus petite (environ

180 Da), elle peut facilement passer les potentiels trous de la chimie de surface. Plus précisément, deux protocoles ont été mis en oeuvre : un premier où la lysine est injectée 10 minutes après la chimie de surface PLL-g-PEG et PLL-biotine et un autre où elle est présente dans la solution de protéines marquées au moment de l'injection. Pour le premier protocole, la durée d'incubation de la lysine est de 30 minutes, le tampon d'incubation est le même que celui de la PLL-g-PEG et PLL-biotine à savoir le tampon H2 ici (10 mM HEPES, 150 mM NaCl à pH = 7,4) et les concentrations en lysine sont les suivantes : 0 - 0,1 - 1 et 10 g/L. Pour le second protocole, la lysine a été préparée dans le tampon B' (50 mM MES, 100 mM NaCl à pH = 6,1) avec 5 nM en protéines marquées et pour les concentrations suivantes : 0 - 0,05 - 0,5 et 5 g/L.



**FIGURE 3.20 – Effet de la lysine sur le nombre d'accroches non spécifiques en protéines marquées.** A) Accroches non spécifiques pour 5 nM en protéines marquées sur une chimie de surface préparée à partir de 1 g/L en PLL-g-PEG,  $10^{-4}$  g/L en PLL-biotine et (1-2-3-4) 0 - 0,1 - 1 et 10 g/L en lysine. B) Accroches non spécifiques pour 5 nM en protéines marquées contenant (1-2-3-4) 0 - 0,05 - 0,5 et 5 g/L en lysine, sur une chimie de surface préparée à partir de 1 g/L en PLL-g-PEG,  $10^{-4}$  g/L en PLL-biotine. Pour les huit conditions, la NeutrAvidine est présente avant l'injection des protéines.

Les Figures 3.20.A et B caractérisent respectivement les premier et second protocole, alors que la concentration en lysine croît de la gauche vers la droite. Globalement, nous pouvons voir que le fait d'injecter la lysine juste après la chimie de surface n'a aucune influence par rapport à la situation où la lysine est absente puisqu'on compte respectivement 4000 - 4200 - 4300 et 4000 points. En revanche, lorsque l'injection de 5 nM en protéines marquées se fait en même temps que la lysine, on peut voir un net effet lorsque la concentration de cette dernière vaut 5 g/L soit environ 27 mM. En effet, le nombre

d'accroches non spécifiques vaut environ 4500 points sans lysine contre 2300 pour 27 mM en lysine. Les autres concentrations en lysine, à savoir 0,27 mM et 2,7 mM, ne permettent pas d'améliorer significativement les performances de la chimie de surface (environ 3500 points dans les deux cas). Ces observations faites avec la lysine concentrée à 27 mM s'est reproduite à plusieurs reprises, ce qui était une bonne nouvelle compte tenu des nouvelles difficultés de passivation observées à l'époque. **Néanmoins, comme cet acide aminé serait chargé positivement dans un tampon d'assemblage, nous avons émis l'hypothèse, qu'à cette concentration, les interactions entre les protéines et l'ARN simple brin seraient altérées.** Ainsi, malgré des résultats très encourageants, nous avons finalement décidé de ne pas réaliser d'expériences d'assemblage dans ces conditions. Dans cette section consacrée, je précise que les figures proviennent de conditions expérimentales différentes de celles décrites jusqu'à présent. En effet, les protéines marquées ont été directement injectées à 5 nM et le tampon B' a été utilisé.

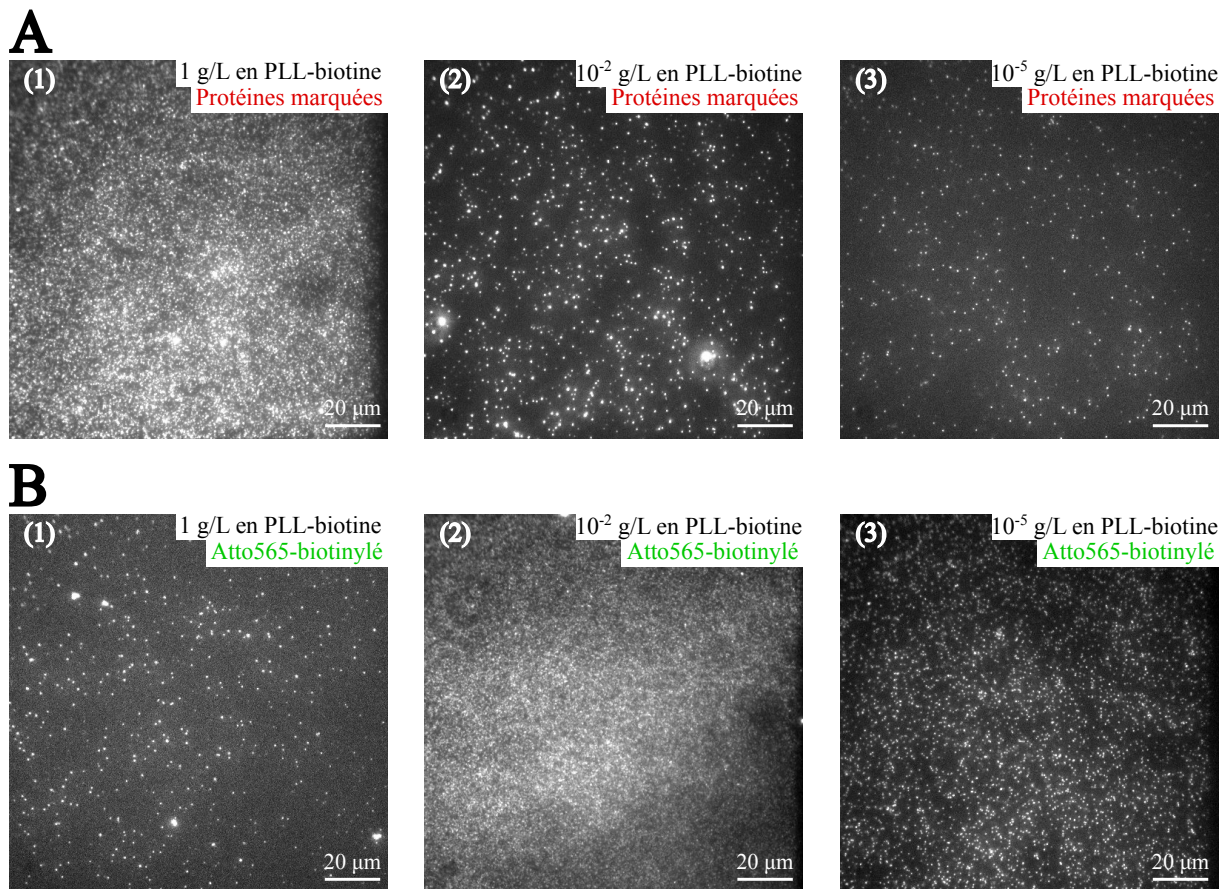
**Pour résumer, parmi les quatre pistes envisagées pour améliorer les performances de la chimie de surface PLL-g-PEG et PLL-biotine, trois d'entre elles ont montré des résultats intéressants.** En effet, utiliser une passivation complémentaire à partir de protéines non marquées élimine quelque peu les accroches non spécifiques en protéines marquées mais en même temps, elle tend à favoriser l'apparition d'agrégats. En ce qui concerne la BSA-biotinylée, son utilisation à la place de la PLL-biotine réduit le nombre d'accroches non spécifiques des sous-unités de capsides. Cela nous a fait comprendre que la concentration en biotine et donc par extension, la forte adsorption en NeutrAvidine, constitue un problème majeur pour notre méthode de passivation. Cependant, la BSA-biotinylée n'a pas été retenue du fait de la faible accroche de cette dernière : nous n'aurions eu que très peu d'ARN par champ de vue. Enfin, injecter de la lysine en même temps que les protéines marquées réduit également le nombre d'accroches non spécifiques. Néanmoins, la forte concentration qu'il est nécessaire d'utiliser pour observer une amélioration significative de la situation nous fait craindre une interaction avec les protéines de CCMV et donc une altération du mécanisme d'auto-assemblage des capsides.

#### 3.2.1.4 La recette finale

Les modifications apportées à la chimie de surface PLL-g-PEG n'ont pas été couronnées de succès mais ont eu le mérite de rappeler un problème évoqué à la section 3.2.1.2 : le non spécifique dépend directement de la présence du couple biotine-NeutrAvidine.

Pour réduire les effets de ce dernier, nous avons décidé de diminuer la concentration en PLL-biotine afin de nous rapprocher des conditions expérimentales utilisées avec la BSA-biotinylée. **Dans toute cette section, la chimie de surface a été préparée dans le tampon H1 à partir de 1 g/L en PLL-g-PEG. Les protéines ont été**

diluées dans le tampon B puis incubées pendant 2 minutes avant d'être rincées dans la même solution.

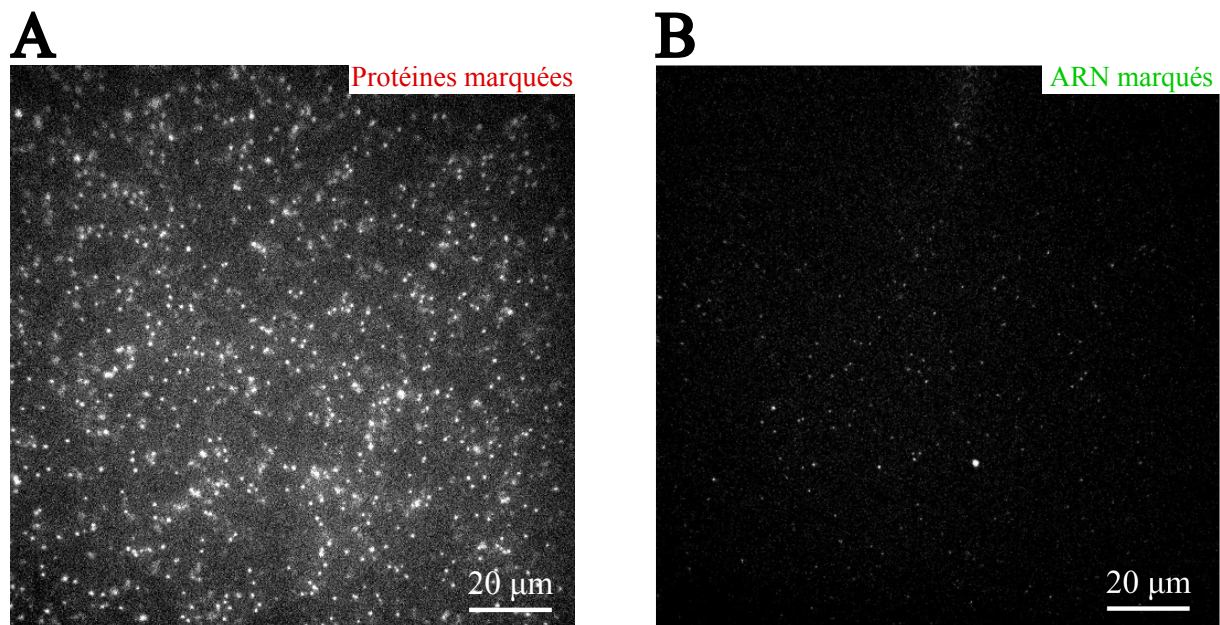


**FIGURE 3.21** – Effet de la concentration en PLL-biotine sur les accroches en protéines marquées et Atto565-biotinylé. A) Accroches non spécifiques issues de 1,5 nM en protéines marquées pour trois cellules différentes. B) Accroches spécifiques issues de 1 nM en Atto565-biotinylé sur les trois mêmes champs de vue présentés en A. La chimie de surface a été préparée à partir de 1 g/L en PLL-g-PEG, (1-2-3) 1 - 10<sup>-2</sup> - 10<sup>-5</sup> g/L en PLL-biotine et en présence de NeutrAvidine.

Sur la Figure 3.21.A, nous comparons les performances de la chimie de surface PLL-g-PEG pour trois concentrations en PLL-biotine : 1 - 10<sup>-2</sup> - 10<sup>-5</sup> g/L. Il apparaît assez clairement que le nombre de protéines marquées visibles (injectées à 1,5 nM après une première étape à 150 pM) à la surface diminue avec la concentration en PLL-biotine. En effet, on passe d'une couverture presque entière pour 1 g/L à environ 1200 points détectés pour 10<sup>-2</sup> g/L contre 500 points à 10<sup>-5</sup> g/L. En parallèle, sur la Figure 3.21.B, on observe les trois mêmes zones mais avec la voie d'excitation verte. De cette façon, nous observons environ 500 Atto565-biotinylés pour 1 g/L en PLL-biotine, un tapis de fluorophores pour 10<sup>-2</sup> g/L et près de 3000 points pour 10<sup>-5</sup> g/L. Les résultats détaillés ici sont très surprenants puisque nous nous attendions à observer une couverture totale en Atto565 pour 1 g/L en PLL-biotine. Le très faible nombre d'accroches obtenu ici peut s'expliquer soit par une mauvaise accroche de la PLL-biotine soit de la NeutrAvidine. Plus

généralement, cette observation est bien le signe que si l'un des paramètres utilisés pour la chimie de surface est défaillant, l'expérience finale est compromise.

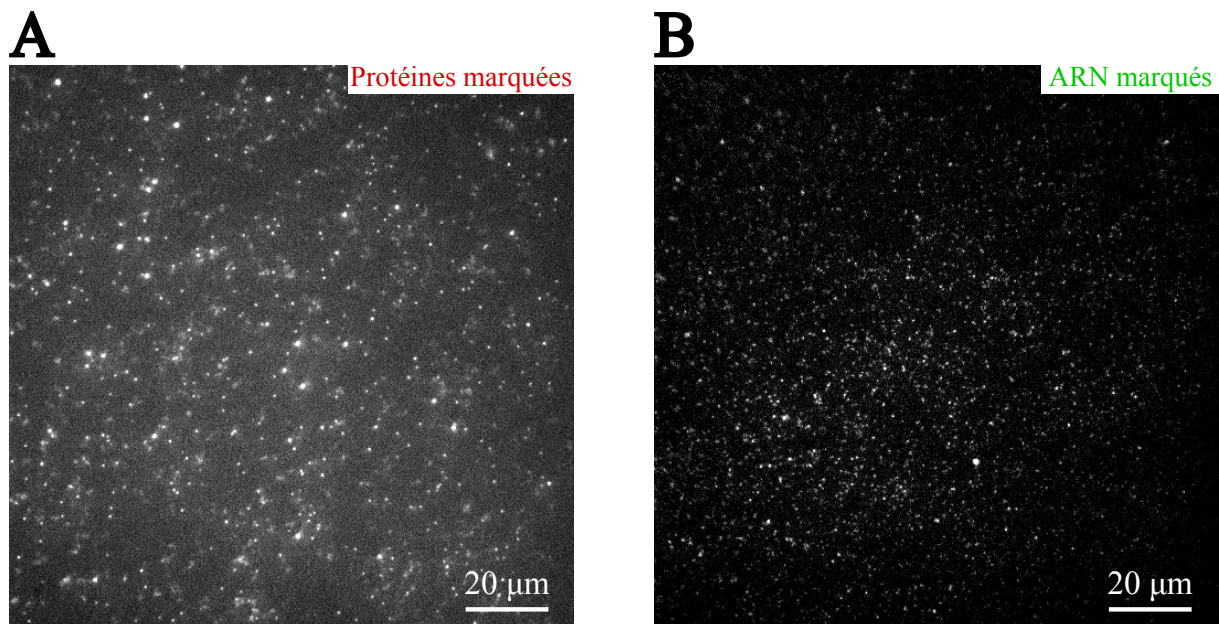
Ainsi, en partant du principe que  $10^{-5}$  g/L est la meilleure configuration, surtout en termes d'accroches non spécifiques en protéines marquées, nous avons voulu faire un test en condition d'assemblage. Après avoir injecté 1 nM en ARN simple brin marqué avec un Atto565, une acquisition a été réalisée afin de repérer leur position. La Figure 3.22.B représente la première image de cette acquisition et il est possible de compter près d'une centaine de points. Cette faible statistique nous rappelle le même résultat détaillé précédemment dans le cas de la BSA-biotinylée. À la suite de l'acquisition réalisée dans la voie d'excitation verte, 5 nM en protéines marquées ont été directement injectées au commencement d'un film rouge. Sur la Figure 3.22.A, nous observons près de 700 points, qui correspondent à des accroches de protéines marquées sur la lamelle ou sur les ARN. Le nombre de sous-unités repérées sur la surface est plus important que le nombre d'ARN, ce qui montre bien qu'il existe toujours des interactions non spécifiques entre les dimères de capsid et la surface, mais elles sont beaucoup moins nombreuses qu'au début.



**FIGURE 3.22 – Observation des accroches en ARN et protéines durant une expérience d'assemblage.** A) Accroches issues de 5 nM en protéines marquées. B) Accroches spécifiques issues de 1 nM en ARN marqués avec un Atto565. La chimie de surface a été préparée à partir de 1 g/L en PLL-g-PEG,  $10^{-5}$  g/L en PLL-biotine et en présence de NeutrAvidine.

Les Figures 3.21 et 3.22 illustrent bien l'amélioration des performances de la chimie de surface de type PLL-g-PEG lorsque l'on baisse la concentration en PLL-biotine et donc le nombre de NeutrAvidine fixées à la lamelle. En revanche, cette forte diminution du nombre de biotines entraîne également une perte importante au niveau des accroches en espèces biotinylées qui vont être observées sur un même champ de vue. Par ailleurs, au

cours d'une autre expérience d'assemblage, j'ai comparé le nombre d'accroches en ARN marqués pour deux concentrations différentes, mais toujours avec une chimie de surface contenant  $10^{-5}$  g/L en PLL-biotine. Pour 1 nM en ARN marqués, quelques dizaines de points se sont accrochés à la surface, contre près de 300 points pour 5 nM. Dans ce cas, les deux injections en ARN ont été réalisées successivement sur une même cellule microfluidique et un même champ de vue. Ce comportement rappelle celui observé avec la BSA-biotinylée, sauf que la faible statistique obtenue pouvait s'expliquer par une difficulté de la BSA à s'accrocher à la lamelle de verre. Avec la chimie de surface réalisée à partir de la poly-L-lysine, les interactions avec la lamelle de verre étant plus importantes qu'avec la BSA, nous avons la possibilité d'ajuster la concentration en PLL-biotine afin de contrôler le nombre d'accroches spécifiques en espèces biotinylées.



**FIGURE 3.23** – Amélioration de la statistique en ARN marqués au cours d'une expérience d'assemblage. A) Accroches issues de 5 nM en protéines marquées. B) Accroches spécifiques issues de 100 pM en ARN marqués avec un Atto565. La chimie de surface a été préparée à partir de 1 g/L en PLL-g-PEG,  $10^{-4}$  g/L en PLL-biotine et en présence de NeutrAvidine.

Pour justifier ce raisonnement, au cours d'une nouvelle expérience d'assemblage, nous avons préparé la même chimie de surface que précédemment mais en augmentant la concentration en PLL-biotine à  $10^{-4}$  g/L. La concentration en ARN simple brin marqué est passée à 100 pM et sur la Figure 3.23.B, nous pouvons observer près de 1400 points, ce qui est la meilleure statistique observée jusqu'alors. Après la détection des sites d'ARN, 5 nM en protéines marquées ont été directement injectées juste après le début d'un film rouge. Sur la Figure 3.23.A nous observons cette fois-ci près de 700 points. Dans les deux cas, on voit qu'il est assez simple de distinguer des molécules fluorescentes proches les unes des autres, ce qui est notre but.



Ainsi, il semblerait que préparer une chimie de surface avec 1 g/L en PLL-g-PEG et  $10^{-4}$  g/L en PLL-biotine dans le tampon H1 (10 mM HEPES à pH = 7,4) puis injecter 400 nM en NeutrAvidine nous permettent d'obtenir de bonnes conditions expérimentales pour travailler à l'échelle de la molécule unique. Sur la Figure 3.24, je propose une représentation schématique simplifiée de l'architecture de la chimie de surface PLL-g-PEG et PLL biotine.

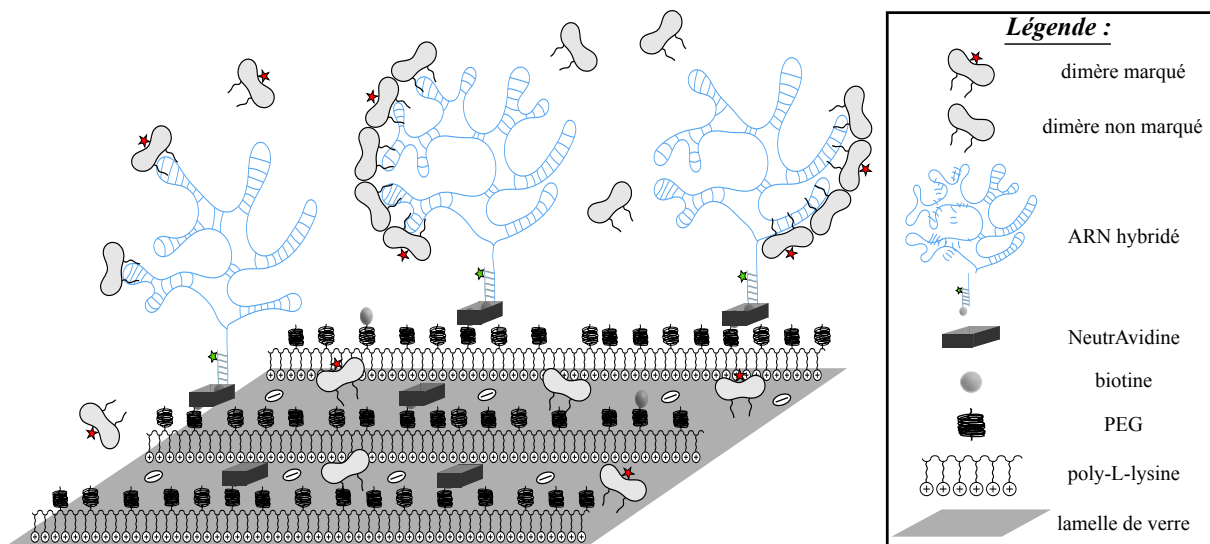


FIGURE 3.24 – Fonctionnalisation PLL-g-PEG et PLL-biotine utilisée en expérience.

Jusqu'à la fin de ma thèse j'ai utilisé cette méthode de passivation à base de PLL-g-PEG et PLL-biotine mais le protocole général a été adapté à plusieurs reprises. En effet, désormais les échantillons sont dissous dans le tampon H2 (10 mM HEPES, 150 mM NaCl à pH = 7,4) et la concentration en PLL-g-PEG est de l'ordre de 1,5 g/L. Le protocole actuel de chimie de surface est disponible à l'annexe C.

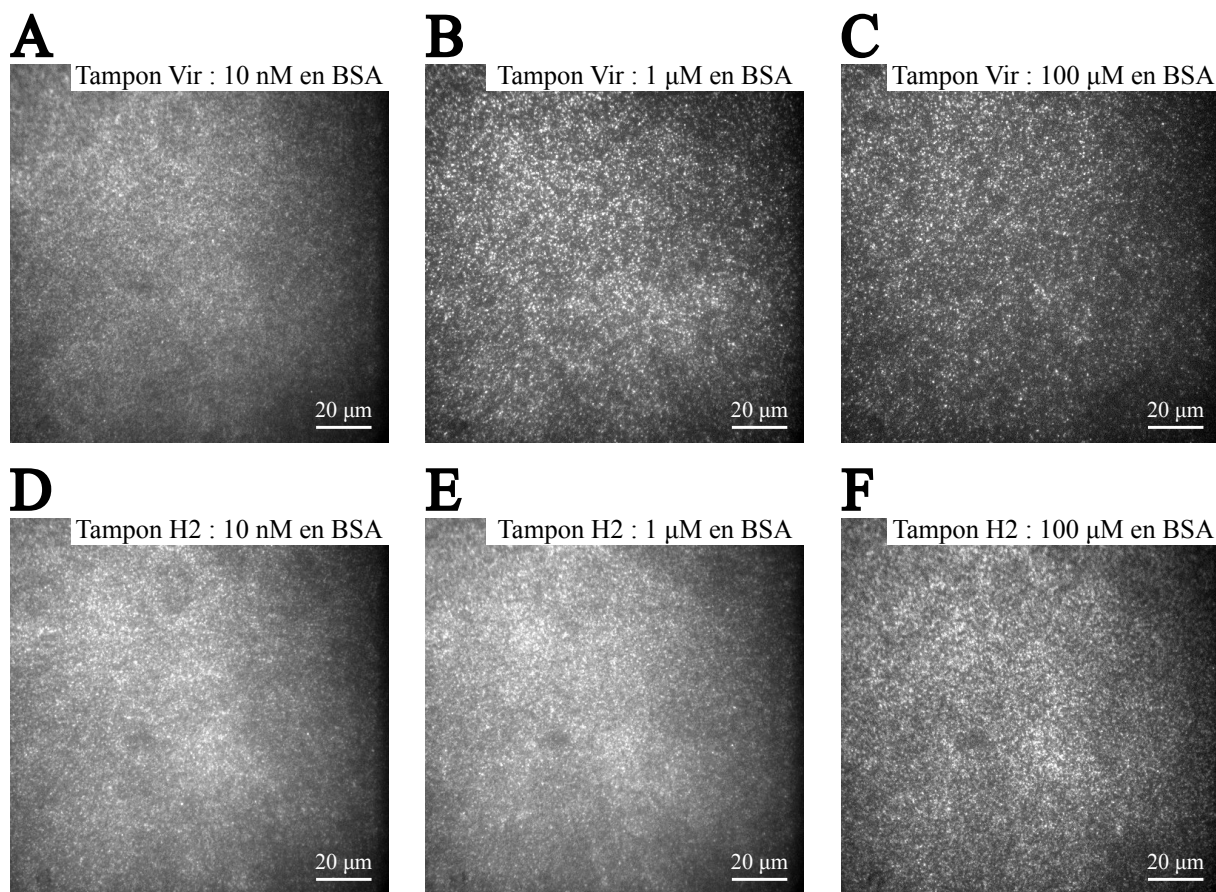
### 3.2.2 Comparaison avec d'autres types de chimie de surface

En parallèle des différents essais effectués avec la PLL-g-PEG, d'autres méthodologies de fonctionnalisation des lamelles de verre ont été testées. Dans cette section, je décris rapidement les résultats obtenus pour les chimies de surface suivantes : la BSA, le PEG-Silane et un exemple de chimie-clic. Par ailleurs, ici aussi les lamelles de microscope ont subi un traitement de 10 minutes au *plasma cleaner* à oxygène juste avant l'expérience. **De plus, pour ces trois architectures de chimie de surface, la concentration en NeutrAvidine sera de 400 nM, sauf indication contraire.**

#### 3.2.2.1 La chimie de surface à base de BSA

Comme je l'ai évoqué plus haut, une méthode très répandue pour passiver une surface consiste à y déposer de la BSA. En effet, malgré son faible niveau d'interaction avec cer-

taines surfaces telles que le verre (Kwok et al. (2007)), la BSA peut empêcher l'apparition d'accroches non spécifiques comme cela a été étudié par Steinitz (2000). Par ailleurs, Park et al. (2018) ont montré qu'en chauffant la BSA pendant plusieurs dizaines de minutes, ses propriétés sont modifiées la rendant ainsi plus adhérente aux surfaces de verre.



**FIGURE 3.25** – Adsorptions non spécifiques sur une chimie de surface de type BSA. A-D) 10 nM de BSA. B-E) 1 μM de BSA. C-F) 100 μM de BSA. Tampons de passivation : A-B-C) tampon Vir et D-E-F) tampon H2. Les accroches non spécifiques sont issues de 5 nM en protéines marquées.

Dans notre cas, la BSA était particulièrement intéressante car elle peut être obtenue avec une modification de type biotine, ce qui lui permettrait de répondre aux deux critères pour la chimie de surface que l'on souhaite utiliser. Néanmoins, comme je l'ai mentionné à la section 3.2.1.3, la BSA ne permet pas de réduire le non spécifique observé avec une passivation de type PLL-g-PEG et PLL-biotine mais remplacer cette dernière par de la BSA-biotinylée améliore grandement la situation. En revanche, que ce soit pour une question d'accessibilité à la lamelle ou d'interactions trop faibles avec cette dernière, la BSA-biotinylée ne permettait pas d'observer assez de points dans la voie verte (Atto565-biotinylés ou ARN marqués). Pour savoir si le problème vient réellement d'un manque de place pour se fixer à la surface, j'ai évalué les performances d'une passivation entièrement réalisée avec la BSA. Sur la Figure 3.25, nous pouvons voir des exemples de couverture

de protéines marquées après une passivation à partir de 10 nM, 1  $\mu$ M et 100  $\mu$ M en BSA (de gauche à droite) et pour deux solutions tampons différentes (de haut en bas) : le tampon H2 (10 mM HEPES, 150 mM NaCl à pH = 7,4) et le tampon Vir (50 mM NaAc à pH = 4,8). Dans les six cas, la BSA a été incubée pendant 30 minutes avant d'être rincée avec le tampon correspondant. Les accroches non spécifiques proviennent d'une incubation pendant 2 minutes de 5 nM en protéines marquées (tampon B' : 50 mM MES, 100 mM NaCl à pH = 6,1).

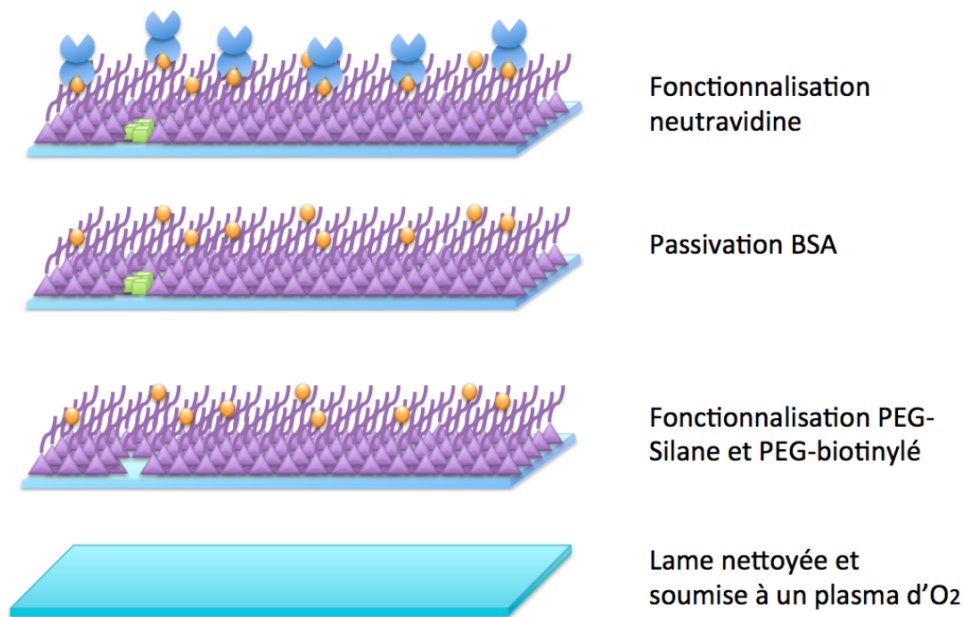
Nous pouvons voir que pour une incubation dans le tampon Vir (Figures 3.25.A-B-C), le nombre d'accroches non spécifiques en protéines marquées semble diminuer quand la concentration en BSA augmente. Néanmoins, même à 100  $\mu$ M en BSA, les adsorptions non spécifiques sont toujours très présentes, alors même que nous n'avons pas utilisé de NeutrAvidine. Dans le cas du tampon H2 (Figures 3.25.D-E-F), il semblerait que la passivation soit moins efficace pour chacune des concentrations en BSA, même si encore une fois, à 100  $\mu$ M, le non spécifique est moins présent qu'à 10 nM et 1  $\mu$ M. Le choix des solutions tampons a été fait comme suit : pour le tampon H2 il s'agit simplement de faire une comparaison avec le tampon utilisé pour la préparation de la chimie de surface de type PLL-g-PEG, alors qu'avec le tampon Vir, le pH est égal au point isoélectrique de la BSA ce qui devrait favoriser l'adsorption de cette dernière selon Yeung et al. (2009). Nos observations expérimentales semblent confirmer une meilleure passivation de la surface par la BSA pour une solution tampon dont le pH vaut 4,8. **Dans tous les cas, cette méthode de passivation ne semble pas suffisante pour contrer efficacement les adsorptions non spécifiques des protéines de CCMV.**

### 3.2.2.2 La chimie de surface PEG-Silane

La fonctionnalisation de surfaces avec le PEG-Silane était déjà utilisée avant mon arrivée au LuMIn par Karen Perronet et il est possible de retrouver des protocoles dans les thèses de Nathalie Barbier et Véronique Vienne (Barbier (2017) et Vienne (2023)). Avec cette méthode, de longues chaînes de PEG présentent chacune une fonction tri-éthoxy silane qui va interagir avec les groupements hydroxyyles d'une lamelle, activés par le *plasma cleaner* à oxygène. Ici, à la différence des lignes de PLL-g-PEG qui vont venir se déposer à la surface (voir Figure 3.24), chaque PEG va directement se fixer sur la lamelle. De plus, il est aussi possible d'utiliser des échantillons de PEG-Silane biotinylés (noté PEG-biotine) afin de permettre l'accroche spécifique des ARN simples brins. Un exemple de représentation de la chimie de surface de type PEG-Silane est donné par la Figure 3.26. Dans ce cas, la BSA était utilisée afin de combler les éventuels trous dans la couverture de PEG-Silane, mais ce n'est plus le cas actuellement.

Par ailleurs, ce type de passivation a déjà été utilisé pour observer le mécanisme d'assemblage des sous-unités de MS2 (Garmann et al. (2019)). Dans cette étude, les au-

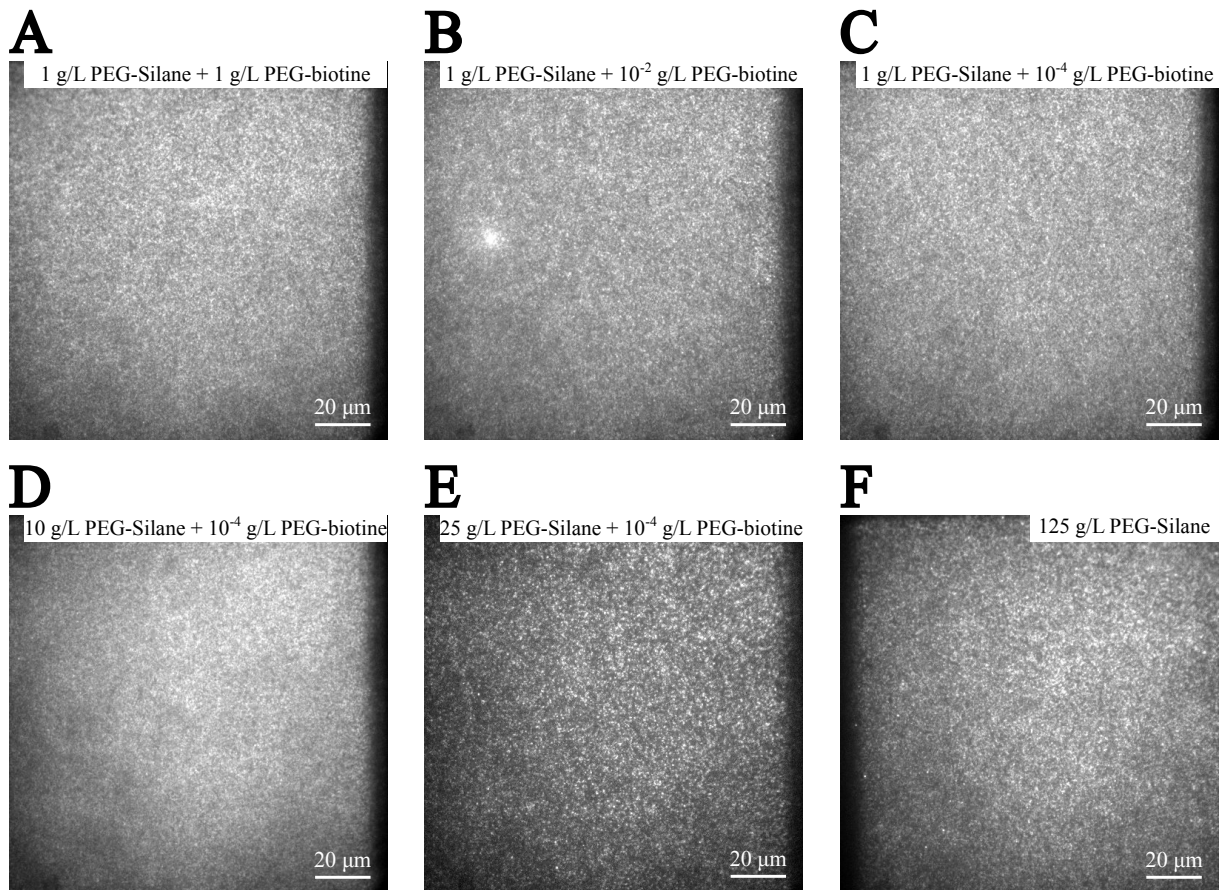
teurs n'utilisent pas exactement le même schéma réactionnel qu'avec le PEG-Silane/PEG-biotine mais l'idée est la même : passiver la surface de verre avec des chaînes de 5 kDa de PEG. Les sites d'accroche pour le matériel génétique ne sont pas obtenus à partir d'une association biotine / NeutrAvidine mais grâce à une réaction de type chimie clic. De plus, à la place de la BSA, un tensioactif appelé Tween-20, a été utilisé pour combler les défauts de la couverture en PEG. Ainsi, les bons résultats obtenus dans le cadre d'expériences d'assemblage semblables à celles que nous souhaitons mettre en place ainsi que l'utilisation en routine du PEG-Silane par Karen nous ont poussés à l'essayer.



**FIGURE 3.26** – Fonctionnalisation de surface PEG-Silane et PEG-biotine. La figure est adaptée de Barbier (2017).

Pour nos tests, nous avons utilisé des échantillons de PEG-Silane et PEG-biotine provenant de *Laysan Bio* et pour lesquels les chaînes de PEG font 5 kDa. Dans tous les cas, la chimie de surface a été préparée dans le tampon adéquat (Tampon borate : 100 mM  $H_3BO_3$  à pH = 8,8) avant d'être incubée pendant 30 minutes puis rincée dans le même tampon. Par ailleurs, la NeutrAvidine a été incubée pendant 10 minutes.

Sur la Figure 3.27, nous pouvons voir que peu importe la concentration en PEG-Silane (1 - 10 - 25 ou 125 g/L) ou celle en PEG-biotine (1 -  $10^{-2}$  -  $10^{-4}$  g/L), l'injection de 5 nM en protéines marquées (tampon B') entraîne des adsorptions non spécifiques sur la totalité du champ de vue. On remarque une légère amélioration pour 25 g/L en PEG-Silane et  $10^{-4}$  g/L en PEG-biotine mais le non spécifique reste trop important. Nous aurions pu remettre en cause la concentration en biotine puisqu'elle n'est pas la même pour  $10^{-4}$  g/L en PEG-biotine et  $10^{-4}$  g/L en PLL-biotine (les masses molaires sont différentes), mais la Figure 3.27.F montre que même sans biotine ni NeutrAvidine, 125 g/L en PEG-Silane ne permet pas de contrer les adsorptions non spécifiques des sous-unités de CCMV.



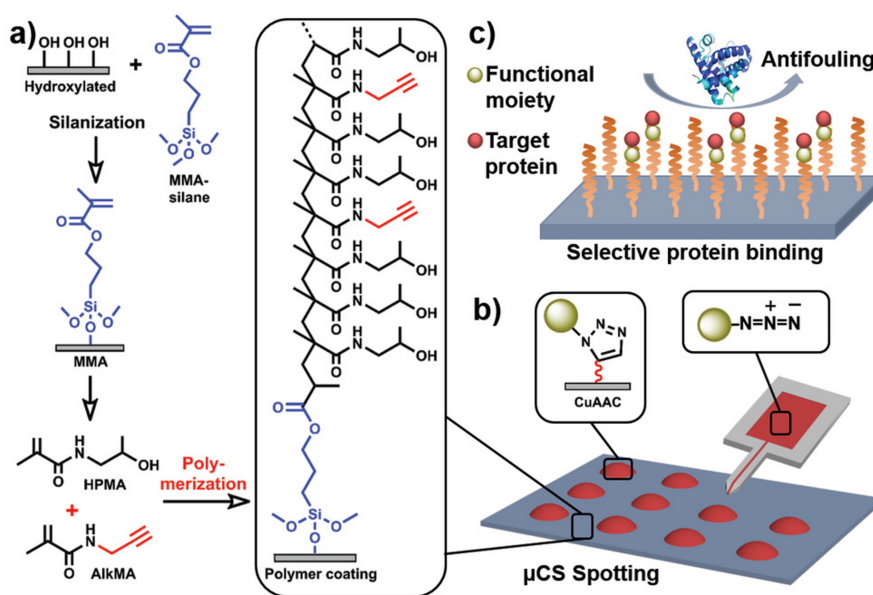
**FIGURE 3.27** – Adsorptions non spécifiques sur une chimie de surface de type PEG-Silane et PEG-biotine. A) 1 g/L en PEG-Silane et 1 g/L en PEG-biotine. B) 1 g/L en PEG-Silane et  $10^{-2}$  g/L en PEG-biotine. C) 1 g/L en PEG-Silane et  $10^{-4}$  g/L en PEG-biotine. D) 10 g/L en PEG-Silane et  $10^{-4}$  g/L en PEG-biotine. E) 25 g/L en PEG-Silane et  $10^{-4}$  g/L en PEG-biotine. F) 125 g/L en PEG-Silane. Les accroches non spécifiques sont issues de 5 nM en protéines marquées.

Ainsi, l'utilisation du PEG-Silane ne semble pas être appropriée aux expériences d'assemblage que l'on souhaite mettre en place. Les performances observées avec les protéines de CCMV peuvent surprendre quand on compare les bons résultats obtenus sur d'autres projets mais finalement les conditions expérimentales ne sont pas comparables. En effet, avec leur montage d'iSCAT Garmann et al. (2019) n'ont pas la même sensibilité que le TIRF d'un point de vue du dimère unique, ce qui rend les accroches non spécifiques beaucoup moins pénalisantes, voire inexistantes. En ce qui concerne les recherches menées par Nathalie Barbier et Véronique Vienne durant leur thèse respective, l'utilisation du PEG-Silane et PEG-biotine permettait d'accrocher des ARN marqués par deux fluorophores différents sur une lamelle. Les ARN étant chargés négativement, les adsorptions non spécifiques sont moins favorables que celles issues du bras N-terminal des dimères de CCMV. **Ce cas illustre bien le fait qu'une méthode de passivation peut fonctionner dans certaines conditions, sans constituer pour autant une solution universelle.** Pour terminer, j'ai essayé à plusieurs reprises de combiner le PEG-Silane

avec la PLL-g-PEG et PLL-biotine ou de remplacer cette dernière par du PEG-biotine mais les résultats étaient également très peu encourageants.

### 3.2.2.3 La chimie de surface de type chimie clic

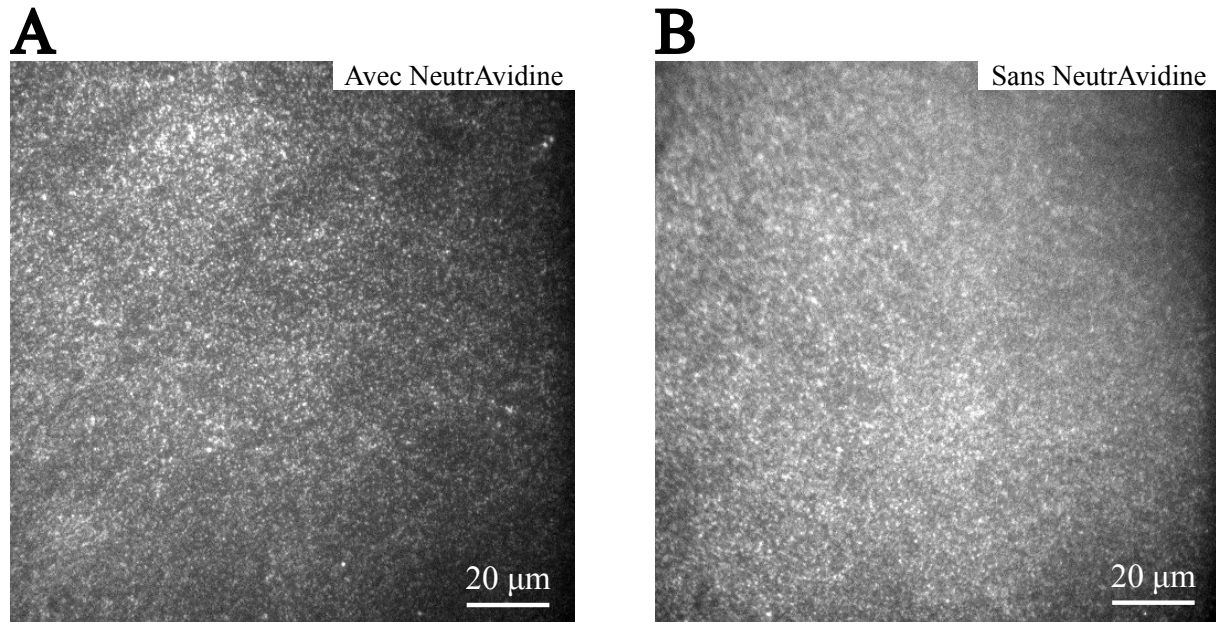
Cette troisième architecture de chimie de surface, qui se base sur une méthode de *click chemistry* (chimie clic), nous a été gentiment fournie par Petr Cigler, un collaborateur de François Treussart (Professeur à l'ENS Paris-Saclay et travaillant au LuMIn). Ici, les lamelles sont d'abord activées par *plasma cleaner* à oxygène avant d'être passivées par une fine couche de méthacrylamide selon la même réaction permettant l'adsorption du PEG-Silane. Ensuite, un copolymère est constitué à partir de deux segments distincts. Un des deux segments, appelé AlkMA (N-propargyl méthacrylamide), possède une triple liaison carbone - carbone (communément appelée alcyne) qui va être utilisée pour réagir avec des azotures (*azide* en anglais) selon une réaction dite CuAAC pour *copper-catalyzed alkyne-azide cycloaddition*. Dans ce mécanisme, les azotures sont en fait liés à un échantillon d'intérêt, comme une biotine, ce qui peut nous permettre d'injecter *a posteriori* de la NeutrAvidine. Cette fonctionnalisation développée par Kumar et al. (2022), a pour but d'être utilisée sur des surfaces de diamant. La Figure 3.28 donne une représentation schématisée des réactions qui viennent d'être décrites.



**FIGURE 3.28 – Fonctionnalisation de surface de type chimie clic.** Figure adaptée de Kumar et al. (2022) où a) montre le mécanisme réactionnel permettant de passiver une surface de verre, b) illustre la réaction CuAAC afin de fonctionnaliser la surface avec les sphères vertes (par exemple avec une biotine) puis c) montre la fixation de protéines d'intérêt en rouge.

Pour nos tests, Petr nous a envoyé plusieurs lamelles fonctionnalisées selon ce procédé et possédant directement une modification de type biotine. Durant la totalité du transport

et juste avant leur utilisation, elles ont été stockées dans un récipient contenant une solution de PBS (*phosphate-buffered saline*) à pH = 7,4. Pour chaque test, l'une des faces de la lamelle a été séchée avec du papier absorbant avant d'être déposée sur le porte-échantillon. Les injections ont donc eu lieu sur la face avec du PBS et si rinçage il devait y avoir, du papier absorbant a été utilisé.



**FIGURE 3.29** – **Adsorptions non spécifiques sur une chimie de surface de type chimie clic.** A) Chimie clic avec NeutrAvidine. B) Chimie clic sans NeutrAvidine. Les accroches non spécifiques sont issues de 5 nM en protéines marquées.

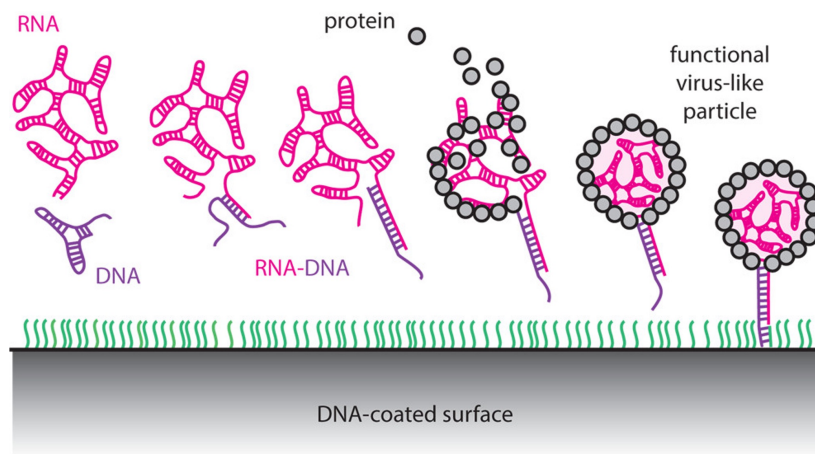
Sur la Figure 3.29, je compare deux conditions expérimentales :

- Injection de NeutrAvidine (400 nM dans le tampon B'), incubation 10 minutes, séchage puis injection de protéines marquées (5 nM dans le tampon B').
- Injection de protéines marquées (5 nM dans le tampon B').

Ici, les deux champs de vue montrent bien un nombre de protéines marquées beaucoup trop important, que ce soit avec ou sans NeutrAvidine. Cette méthode de passivation semble se comporter de la même façon que le PEG-Silane et PEG-biotine. En effet, pour les deux chimies de surface, ce ne sont pas de longues lignes qui viennent se déposer sur la lamelle, mais plutôt des sites unique. En revanche, dans le cas de la chimie clic, la concentration en biotine n'a pas pu être ajustée vu que la modification était déjà présente à la réception des lamelles. Or il est possible qu'une trop grande concentration en biotine vienne favoriser l'adsorption non spécifique des sous-unités de CCMV. Par ailleurs, comme les lamelles fournies par Petr ne sont pas adaptées aux cellules micro-fluidiques, nous avons décidé de ne pas réaliser davantage de tests avec cette fonctionnalisation. **Ainsi, nous avons conclu que dans ces conditions, la troisième architecture de chimie de surface ne répond pas aux critères attendus pour nos expériences.**

### 3.2.3 L'hybridation de l'ARN simple brin

La Figure 3.24 met en avant la façon dont l'ARN simple brin est accroché à la chimie de surface. Comme je l'ai dit plus haut, une biotine permet de faire la jonction avec la NeutrAvidine précédemment fixée à la PLL-biotine. De plus, un marqueur fluorescent, ici l'Atto565, est utilisé afin de repérer l'ARN au cours d'une illumination TIRF. La biotine et l'Atto565 ne sont pas directement présents sur l'ARN simple brin mais sur une séquence de nucléotides, appelée oligo-nucléotides, qui va être hybridée à l'ARN. Avant de revenir sur le protocole d'hybridation plus en détails, il est important de préciser que la Figure 3.24 montre que la partie de l'ARN proche de la chimie de surface, ici son extrémité 5', semble être assez rigide et surtout exclue de l'assemblage. Cette construction peut surprendre et mettre en doute la formation de capsides pleines mais Garmann et al. (2015) ont montré qu'elle est tout à fait possible. En effet, les auteurs ont observé des capsides de CCMV entièrement assemblées autour de son ARN hybridé à la surface. La partie rigide du matériel génétique semble passer par un petit trou présent dans la capside, sans que cela ne semble empêcher l'assemblage de cette dernière. La Figure 3.30 illustre l'assemblage de type *cherry bomb* observé par Garmann et al. (2015).



**FIGURE 3.30** – Illustration de l'assemblage sous forme de *cherry bomb*. Figure adaptée de Garmann et al. (2015).

Dans notre cas, nous avons rapidement changé la façon dont l'ARN simple brin était hybridé. En effet, pour les premières expériences d'assemblage, une simple hybridation était réalisée entre une séquence de nucléotides et l'extrémité 5' de l'ARN, comme le montre la Figure 3.31.A. Cette séquence longue de 40 nucléotides est composée : d'une modification en 3' de type biotine, de 20 adénines successives, d'un Atto565 puis de 20 nucléotides complémentaires de l'extrémité 5' de l'ARN du CCMV. Cependant, deux problèmes peuvent être mis en avant. En premier lieu, la présence des 20 adénines initiales peut altérer la rigidité de la *cherry bomb* vu qu'elles ne sont pas hybridées. De plus, le fait que la biotine et l'Atto565 soient présents sur la même séquence signifie que l'efficacité



de l'hybridation importe peu : dans tous les cas, l'oligo-nucléotide peut se fixer à la NeutrAvidine grâce à sa biotine. Un ARN correctement hybridé ou un oligo-nucléotide avec son Atto565 donneront un point fluorescent de même taille et même intensité.

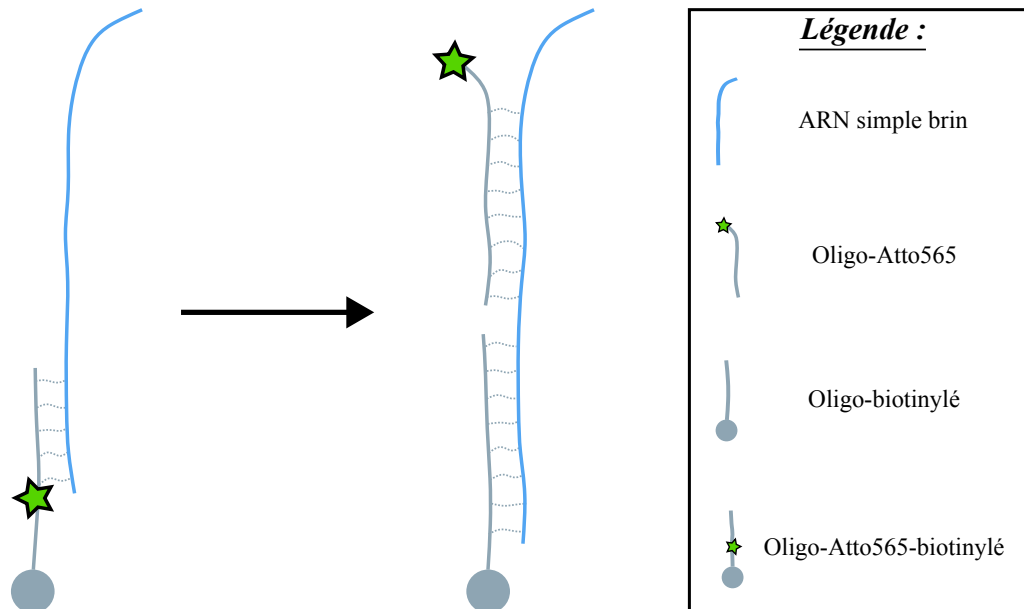
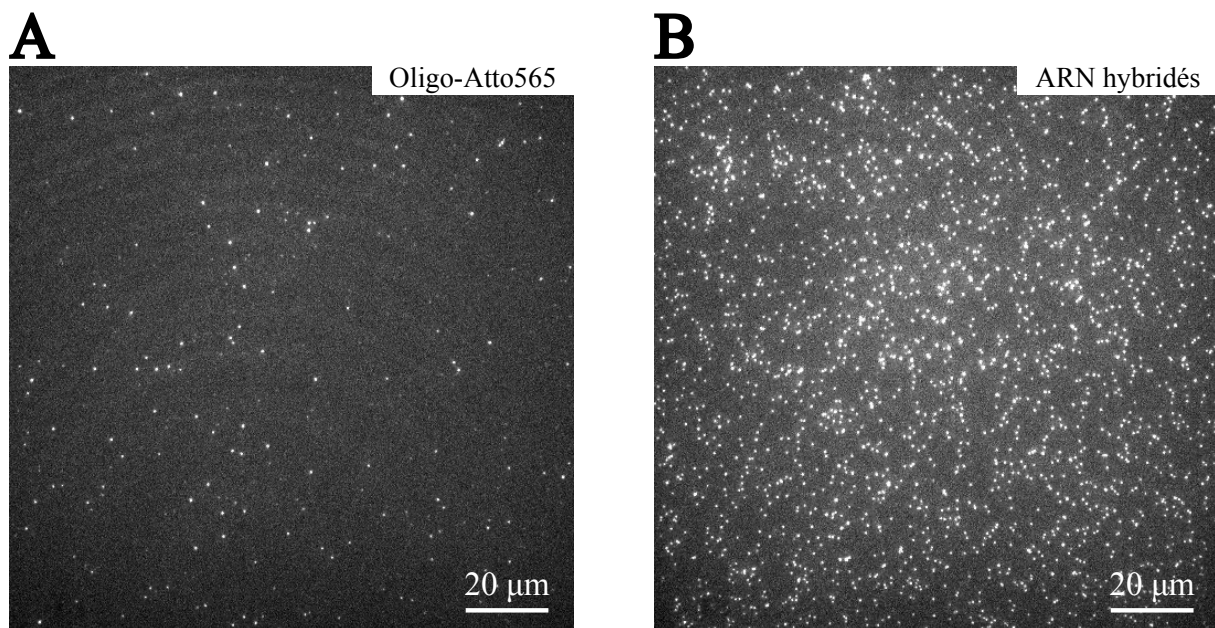


FIGURE 3.31 – Comparaison des constructions d'hybridation. A) Hybridation simple. B) Hybridation double.

Pour contrer ce problème, nous sommes passés à une construction avec deux oligo-nucléotides et donc à une double hybridation, comme le montre la Figure 3.31.B. La séquence décrite pour la simple hybridation est en fait découpée en deux séquences plus petites auxquelles je fais référence sous le nom OB pour oligo-biotinylé et OA pour oligo-Atto565. L'oligo-biotinylé possède une modification en 3' de type biotine et 53 nucléotides dont voici la séquence : ATG-GAG-ACC-CGT-GAG-TCG-ACA-AGG-GAT-TGA-ACC-TCG-CTC-TCG-TGG-ATT-ACA-AA-biotinTEG (dans le sens 5' > 3'). Les trois adénines proches de la biotine ne sont pas hybridées à l'ARN simple brin, elles servent à éloigner légèrement ce dernier de la chimie de surface. Pour l'oligo-Atto565, la modification est présente à son extrémité 5' et la séquence possède 51 nucléotides qui peut être décrite comme suit : Atto565-AAA-AAG-TAA-AGA-AGA-AAA-ATA-TCA-AGA-AGA-AAA-TGT-ATA-AAC-TGT-TTT-CAA (dans le sens 5' > 3'). Ici aussi, les trois adénines proches du marqueur fluorescent ne sont pas hybridées à l'ARN et permettent au fluorophore de ne pas être au contact de ce dernier. Au total, la séquence de 104 nucléotides hybridés à l'ARN a une longueur d'environ 30 nm, ce qui est bien inférieur à la longueur de persistance de 64 nm d'un ARN double brin estimée par Abels et al. (2005). L'extrémité hybridée de l'ARN peut donc bien être considérée comme étant rigide.

Cette double hybridation n'est en réalité bénéfique que si l'OA ne se fixe pas non spécifiquement sur la chimie de surface. Pour s'en assurer, 300 pM en OA ont été injec-

tés dans trois cellules micro-fluidiques fonctionnalisées selon le protocole le plus récent (1,5 g/L en PLL-g-PEG,  $10^{-4}$  g/L en PLL-biotine et 20 nM en NeutrAvidine). Dans les trois cas, l'injection a été suivie de deux minutes d'incubation puis d'un rinçage et nous avons détecté 114 - 227 - 425 points fluorescents, ce qui est vraiment très faible. À la suite de chacune de ces mesures, 100 pM en ARN hybridés ont été injectés et laissés incuber pendant deux minutes puis rincés, ce qui nous a permis d'observer 1426 - 2270 - 2738 points fluorescents. **Sans aucun doute, ces six tests nous montrent que les accroches non spécifiques provenant de l'oligo-Atto565 sont très faibles même s'il en existe quelques-unes.** La Figure 3.32 compare les accroches issues de 300 pM en OA (à gauche) contre 100 pM en ARN hybridés (à droite) sur une lamelle avec une chimie de surface de type PLL-g-PEG et PLL-biotine. Ici, les solutions d'OA et d'ARN hybridés ont été préparées dans le tampon B'. De plus, dans le protocole que je détaille à l'annexe D, je précise que le volume d'hybridation est trois fois plus concentré en oligo-Atto565 qu'en ARN ce qui explique pourquoi la concentration en ARN hybridés est de 100 pM.



**FIGURE 3.32** – Spécificité de la fonctionnalisation de surface pour l'ARN hybridé. A) Accroches non spécifiques pour 300 pM en Oligo-Atto565. B) Accroches spécifiques pour 100 pM en ARN hybridés.

Ce type de contrôle a été réalisé à plusieurs reprises et quelques fois, j'ai pu observer une forte augmentation des accroches non spécifiques des oligo-Atto565 pour une concentration de 300 pM. En effet, il m'est arrivé de détecter plusieurs centaines de points fluorescents. Ces observations sont surtout apparues depuis la réception du nouveau lot de chimie de surface début 2024, qui a eu lieu en même temps qu'un nouvel aliquotage en NeutrAvidine. La modification de ces deux échantillons essentiels pour obtenir une bonne fonctionnalisation des lamelles de verre peut expliquer l'apparition de nouvelles

sources d'accroches non spécifiques. Néanmoins, j'ai pu remarquer que ces adsorptions non souhaitées peuvent être réduites si le volume de rinçage est augmenté. À l'annexe D, je reviens plus en détails sur les accroches non spécifiques en OA en insistant notamment sur le fait que les protéines marquées sont peu enclines à s'y accrocher.

### 3.2.4 Conclusion

Dès lors que l'on souhaite contrôler les accroches spécifiques ou non spécifiques d'un échantillon d'intérêt sur une lamelle de verre, l'étape de fonctionnalisation apparaît rapidement comme étant primordiale. Afin de couvrir un large spectre des expériences qui nécessitent une chimie de surface, plusieurs protocoles assez généraux ont été développés, impliquant notamment le PEG-Silane ou la BSA. Par ailleurs, il est assez commun désormais d'utiliser des protéines, des séquences de nucléotides ou des marqueurs fluorescents ayant une modification de type biotine afin de fixer spécifiquement l'échantillon que l'on souhaite étudier.

Néanmoins, comme la caractérisation d'une chimie de surface en termes d'interactions non spécifiques est généralement faite avec un nombre limité de protéines ou autres, il est possible de rencontrer certaines difficultés avec un échantillon en particulier. Dans notre cas, les accroches non spécifiques des protéines de CCMV ont longtemps été un problème à cause, selon toute vraisemblance, de leur bras N-terminal chargé positivement. Cette difficulté est d'autant plus prépondérante dans nos expériences que chaque marqueur fluorescent, *i.e.* chaque sous-unité marquée, est détecté par la caméra EMCCD. Pour réduire la proportion en accroches non souhaitées, beaucoup de paramètres peuvent être ajustés comme la salinité du milieu, le pH, la concentration en PLL-g-PEG, PLL-biotine ou encore NeutrAvidine. La reproductibilité a été un problème majeur puisqu'à chaque fois qu'un paramètre semblait réduire les adsorptions non spécifiques, il est souvent apparu que l'expérience d'après nous montrait le contraire. Pour les premiers tests, la non reproductibilité des expériences pouvaient également provenir du manque de précision dans la préparation des échantillons de poudre de PLL-g-PEG et PLL-biotine. En effet, les lots commandés sous la forme de 10 mg de poudre étaient répartis approximativement dans 20 tubes afin d'obtenir une masse de 0,5 mg pour chaque aliquot. Cependant, comme *SuSoS* précise que les poudres sont très sensibles à l'humidité et à l'oxygène, cette étape d'aliquotage a été réalisée dans une boîte à gants sous atmosphère inerte, ce qui a rendu encore plus approximative la répartition des 0,5 mg par aliquot. Ce n'est que plus tard, que nous avons commencé à stocker dans les tampons H1 ou H2, des solutions concentrées à 50 g/L en PLL-g-PEG et PLL-biotine.

Cette absence de continuité venait s'ajouter à notre manque de connaissance sur la vitesse de dégradation de la PLL-g-PEG et PLL-biotine, que ce soit en solution ou en poudre. Par ailleurs, les concentrations en PLL-g-PEG, PLL-biotine et NeutrAvidine sont

souvent sujettes à des modifications, notamment en cas de nouveau lot. En effet, début 2023, peu de temps après l'achat d'un nouveau lot de chimie de surface, j'ai conservé les concentrations en PLL-g-PEG et PLL-biotine mais je suis passé du tampon H1 au H2 afin de réduire les accroches non spécifiques. Plus tard, au début de l'année 2024 une nouvelle commande de chimie de surface a été réalisée, m'obligeant à multiplier par 3 la concentration massique en PLL-g-PEG pour atteindre 1,5 g/L. Ce paramètre est d'autant plus important qu'il dépend de l'architecture des chaînes de PLL-g-PEG. En effet, nous avons souhaité remplacer la PLL(20)-g(3,5)-PEG(2) par son homologue plus dense : PLL(20)-g(3,5)-PEG(5). Nous sommes partis du principe que des chaînes plus longues de PEG peuvent apparaître plus imposantes et donc bloquer plus facilement les sous-unités de CCMV. Cependant, il est apparu que la concentration massique à injecter devait être presque trois fois supérieure à celle en PLL(20)-g(3,5)-PEG(2), pour avoir les mêmes performances. Cette différence provient en réalité de l'écart en termes de masse molaire : 76 kDa pour la PLL(20)-g(3,5)-PEG(2) contre 160 kDa pour la PLL(20)-g(3,5)-PEG(5). Nous avons donc conservé la PLL(20)-g(3,5)-PEG(2). En parallèle, la concentration en NeutrAvidine qui a longtemps été fixé à 400 nM, varie désormais entre 80 et 20 nM.

Par ailleurs ces différents tests nous ont permis de remarquer un comportement assez particulier des protéines. Effectivement, il est apparu que si l'on utilise une même solution en protéines sur plusieurs cellules micro-fluidiques consécutives, le nombre d'accroches non spécifiques diminue petit à petit sans que le protocole de chimie de surface ait été modifié. Cette particularité, nous a permis de comprendre *a posteriori*, pourquoi les accroches non spécifiques pouvaient être si faibles dans le cas de la BSA-biotinylée (voir Figure 3.19). Nous avons mis en cause les cônes pour micro-pipette et les tubes mais aucune amélioration significative n'a été observée pour différents modèles utilisés. Pour contrer ce problème, j'ai donc décidé de préparer une nouvelle solution de protéines marquées à la concentration souhaitée pour chaque cellule micro-fluidique.

Ainsi, le protocole décrit en annexe C peut sembler assez fragile vu qu'il peut subir assez régulièrement certains ajustements. Cependant, même s'il ne nous permet pas d'éliminer entièrement les accroches non spécifiques des protéines marquées, il est assez efficace pour nous permettre d'observer des sites individuels d'ARN marqués, de dimères marqués et donc de travailler à l'échelle de la molécule unique. Pour terminer, je précise que ce protocole a été optimisé pour 5 nM en protéines marquées, ce qui nous permet d'avoir une concentration totale en protéines de 50 nM lorsque l'on considère un taux de marquage de 10 %. Comme j'en parle au chapitre 5, certaines expériences d'assemblage ont été effectuées pour une concentration en protéines marquées de 10 nM ce qui semble être très proche de la valeur limite acceptable afin de détecter des points individuels.

### 3.3 Retour sur la fluorescence

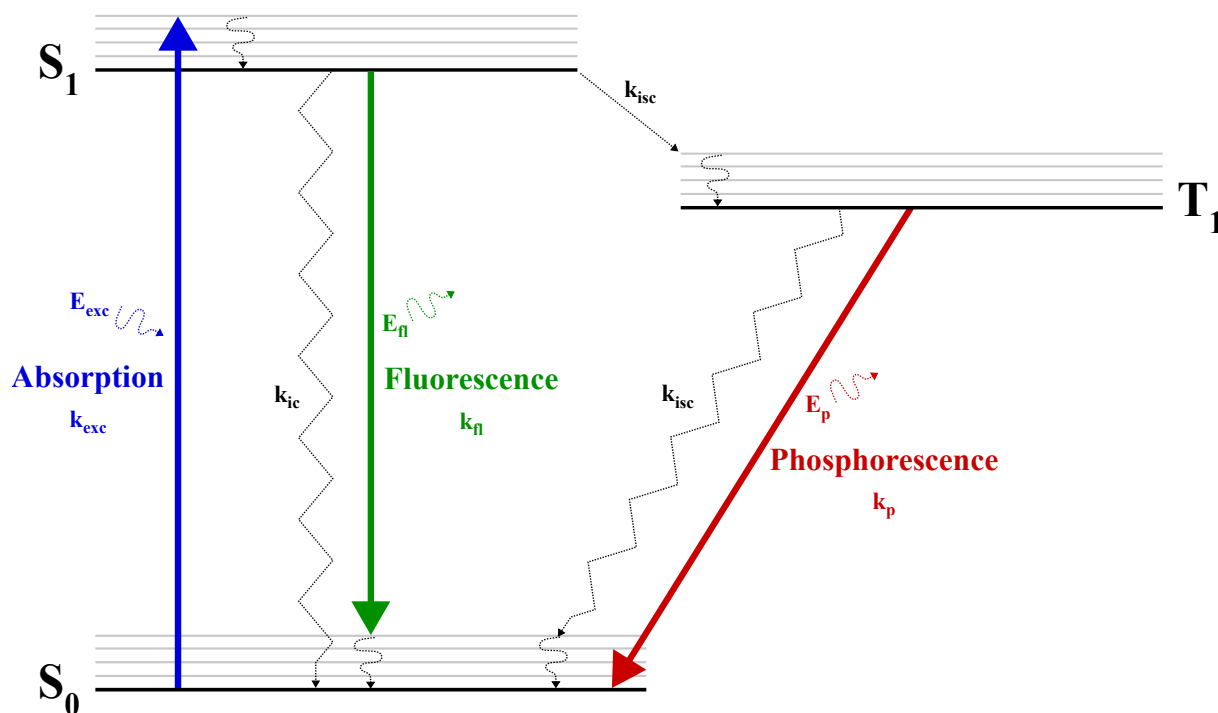
Les parties précédentes m'ont permis d'introduire deux bases importantes du projet de thèse. En les combinant, elles nous offrent la possibilité de détecter les échantillons qui nous intéressent, à savoir un ARN simple brin et des protéines de capsid. L'observation de ces derniers à partir de marqueurs fluorescents, tels que l'Atto647N ou l'Atto565, se base sur ce que l'on appelle la fluorescence. Ce principe physique a vu le jour au milieu du XIX<sup>e</sup> siècle notamment grâce aux travaux de John Herschel en 1845 puis de Georges Stokes en 1852 mais ce n'est qu'une cinquantaine d'années plus tard que les premiers microscopes à fluorescence ont été mis au point. Dès lors, de nombreuses évolutions ont été apportées à ce domaine que ce soit d'un point de vue optique, des sources d'illumination ou encore des stratégies de marquage.

La microscopie de fluorescence est une méthode d'imagerie très répandue et qui est à la base de plusieurs techniques optiques, dont le TIRFM. D'une manière générale, elles ont pour objectif d'étudier le comportement de protéines d'intérêt en les marquant avec des protéines fluorescentes telles que la *Green Fluorescent Protein* (GFP) ou des espèces chimiques beaucoup plus petites comme l'Atto647N et l'Atto565 que j'ai mentionnées sous le nom de marqueurs fluorescents ou fluorophores. De cette façon, c'est l'acquisition du signal de fluorescence qui va permettre de caractériser, par exemple, le déplacement des molécules d'intérêt au cours du temps.

Dans cette section, je commence par présenter le principe physique associé à la fluorescence ainsi que les mécanismes importants. Dans un second temps, je m'intéresse au marqueur fluorescent Atto647N qui est utilisé afin de marquer les protéines du CCMV. Je reviens notamment sur l'accroche protéine - marqueur, les propriétés photophysiques de ce dernier mais également sur la façon dont ses caractéristiques peuvent être améliorées.

#### 3.3.1 Principes et mécanismes importants

Le principe de fluorescence peut être décrit au regard des niveaux d'énergie d'un atome. Sans apport d'énergie, cet atome se trouve dans son état fondamental  $S_0$ . En revanche, lorsqu'un photon d'énergie  $E_{\text{exc}}$  est absorbé par le système, l'atome passera dans son premier niveau singulet excité  $S_1$ . Par la suite, en se désexcitant, l'atome va émettre un nouveau photon d'énergie  $E_{\text{fl}}$  plus faible que  $E_{\text{exc}}$ . Ce décalage énergétique entre l'absorption et l'émission, connu sous le nom de décalage de Stokes, est à la base du fonctionnement de la microscopie de fluorescence. De cette façon, il est possible de séparer le faisceau d'excitation qui apporte son énergie au marqueur fluorescent, du faisceau d'émission qui provient du fluorophore. Ainsi, en choisissant les filtres optiques adéquats selon les longueurs d'onde d'intérêt, seul la fluorescence est collectée par le système de détection (voir description du microscope TIRF à la section 3.1.2.1).



**FIGURE 3.33** – **Diagramme de Jablonski.** Les niveaux d'énergie S et T sont représentés par des traits horizontaux noirs et les niveaux vibrationnels en gris. Les flèches de couleur en trait plein décrivent des transitions radiatives (absorption, émission) et celles en pointillés noirs illustrent les différents mécanismes non radiatifs. Les grandeurs  $k_{exc}$ ,  $k_f$ ,  $k_p$ ,  $k_{ic}$  et  $k_{isc}$  caractérisent respectivement les taux d'excitation et de désexcitation de fluorescence, phosphorescence, conversion interne (S<sub>1</sub> vers S<sub>0</sub>) et intersystème (T vers S ou S vers T).

D'un point de vue cinétique, une transition radiative S<sub>1</sub> vers S<sub>0</sub> caractérisant la fluorescence se produit généralement sur des durées de l'ordre de la nanoseconde. En comparaison, les transitions non radiatives représentées par des flèches ondulées sur la Figure 3.33, ne durent que quelques picosecondes. Elles correspondent à la désexcitation d'un niveau d'énergie vibrationnelle supérieur vers son fondamental pour un même niveau d'énergie S. La perte d'énergie qui en découle est généralement libérée sous forme thermique. Par ailleurs, d'autres niveaux d'énergie, comme par exemple les états triplets, peuvent conduire à des transitions différentes. En effet, à partir de son premier niveau excité S<sub>1</sub>, un atome peut passer dans le premier état triplet noté T<sub>1</sub> dont l'énergie est plus faible. Dans cette configuration, la fluorescence n'est pas observée et la désexcitation vers S<sub>0</sub> est uniquement possible par une transition non radiative ou par émission d'un photon via le principe de phosphorescence. Dans ce cas, le photon a une énergie E<sub>p</sub> plus faible que celui émis par fluorescence. Il est important de noter que la durée de vie de cet état triplet est plus longue que celle de S<sub>1</sub> (de 10<sup>-6</sup> à 10<sup>-3</sup> s) ce qui sera visible au cours du suivi de données prises en fluorescence. Le diagramme de Jablonski représenté en Figure 3.33 reprend les principaux mécanismes mentionnés durant les deux derniers paragraphes.

En revanche, les protéines fluorescentes ou autres fluorophores organiques pâtissent de

deux mécanismes particuliers qui affectent leur photostabilité et peuvent rendre difficile l'interprétation des données obtenues :

- **Le photoblanchiment** : c'est généralement le phénomène photophysique qui vient en premier à l'esprit lorsqu'il s'agit d'énoncer les inconvénients de la microscopie de fluorescence. En effet, il caractérise tout simplement un état irréversible du fluorophore dans lequel plus aucune fluorescence ne peut être émise. Les processus mis en jeu ne sont pas encore parfaitement compris, mais les interactions entre le fluorophore présent dans son état excité et les molécules de son environnement, comme notamment l'oxygène, participent grandement au photoblanchiment. Il est important de comprendre que ce mécanisme photo-induit correspond à la durée de vie du fluorophore. Par conséquent, des paramètres tels que l'intensité d'excitation et le temps d'exposition de la molécule fluorescente contribuent à une apparition plus ou moins rapide du photoblanchiment, au même titre que l'environnement chimique. Ces deux points sont étudiés dans les sections 3.3.2.1 et 3.3.3.

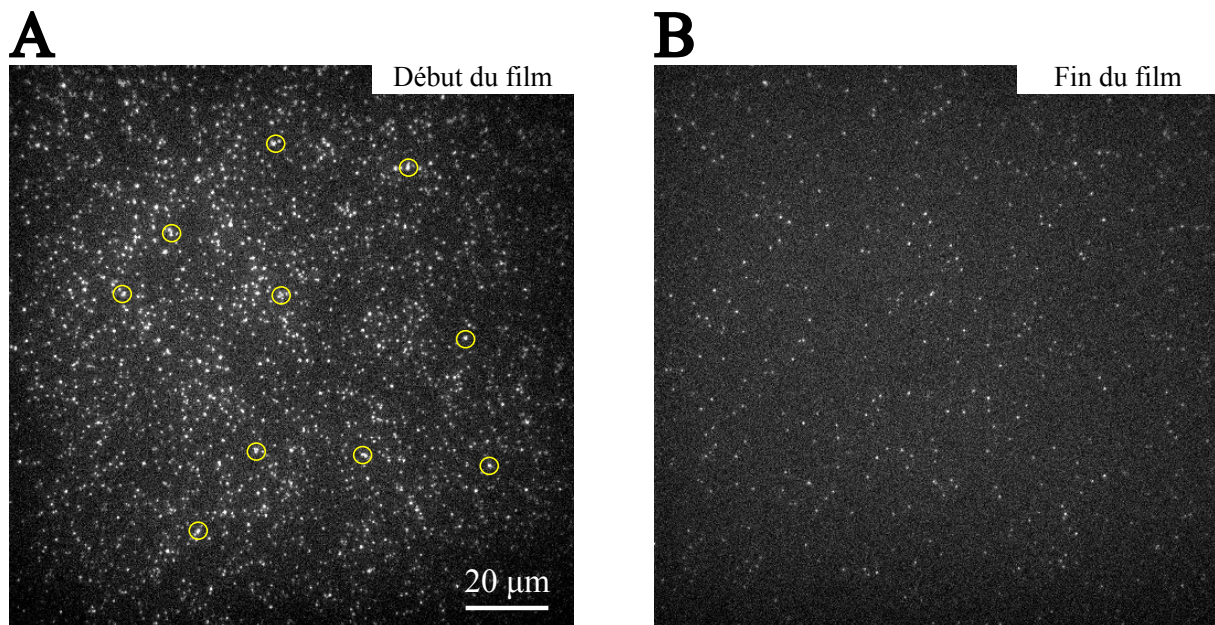
- **Le clignotement** : ce mécanisme illustre parfaitement l'effet que peut avoir l'état  $T_1$  sur l'acquisition d'un signal de fluorescence. Parmi les millions de transitions  $S_1$  vers  $S_0$ , le fluorophore peut se retrouver à plusieurs reprises bloqué dans cet état triplet également appelé état noir. En effet, jusqu'à la désexcitation vers  $S_0$ , aucune intensité de fluorescence n'est mesurée. Une fois que le retour au niveau fondamental est établi, la fluorescence peut de nouveau être mesurée jusqu'au prochain passage dans l'état triplet. Des études expérimentales ont montré que dans ce cas, l'oxygène dissous dans le milieu aide le fluorophore à quitter l'état triplet. En effet, un échange d'électrons entre l'état triplet de l'oxygène et celui du fluorophore permet de renvoyer ce dernier dans son état fondamental. En absence d'oxygène, Hubner et al. (2001) ont montré que la durée de vie de l'état triplet augmente. Dans la section 3.3.2.2 je reviens plus en détails sur l'observation expérimentale des clignotements avec l'Atto647N.

Ainsi, selon l'expérience de microscopie de fluorescence que l'on souhaite réaliser, il est important de connaître la photostabilité du fluorophore utilisé. C'est ce que je présente dans la section suivante pour l'Atto647N.

### 3.3.2 Caractérisation de la photostabilité de l'Atto647N

Lorsque l'on travaille à l'échelle de la molécule unique en microscopie de fluorescence, on s'intéresse aux fluctuations individuelles d'intensité de chaque fluorophore détecté. De cette façon, une arrivée ou un départ d'une molécule fluorescente dans le volume d'excitation va correspondre respectivement à une augmentation ou une diminution soudaine de l'intensité mesurée. C'est donc en détectant ce type d'événements qu'il est possible de décrire le comportement de la protéine d'intérêt liée au fluorophore. Dans le cas du suivi du mécanisme d'auto-assemblage de la capsid de CCMV, nous souhaitons observer plusieurs dimères marqués s'accrocher et quitter un même site d'ARN au cours du temps.

En d'autres termes, plusieurs augmentations et diminutions d'intensité vont caractériser les signaux obtenus et c'est à partir de ces variations soudaines que l'on souhaite en apprendre davantage sur la cinétique et la dynamique d'assemblage de la capsid du CCMV. Cependant, les perturbations de la photostabilité de l'Atto647N vont venir complexifier l'interprétation des événements d'arrivée (clignotements) et de départ (photoblanchiment et clignotements) des dimères marqués. Il est donc important de caractériser les mécanismes de photoblanchiment et clignotements pour l'Atto647N. Pour ce faire, j'ai réalisé le plus souvent possible des contrôles dits de "photophysique" dans les mêmes conditions expérimentales (illumination et solution tampon) que les expériences d'assemblage. Pour ces tests, la chimie de surface décrite à la section précédente est conservée, mais au lieu d'accrocher des ARN hybridés, ce sont des Atto647N-biotinylés qui se fixent aux sites de NeutrAvidine. Dans toute cette section, je présente des résultats pour lesquels les fluorophores ont été incubés deux minutes avant d'être rincés par du tampon dans lequel ils ont été préparés. La Figure 3.34 montre la première (A) et la dernière (B) image observée sur une cellule de photophysique.



**FIGURE 3.34** – Illustration d'une couverture en Atto647N pour une expérience de photophysique. Les accroches correspondent aux Atto647N repérés en début (A) et fin (B) d'acquisition. Les cercles jaunes indiquent des agrégats (précisions à la page suivante). Conditions d'illumination : 1600 images, 50 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 5,4 mW.

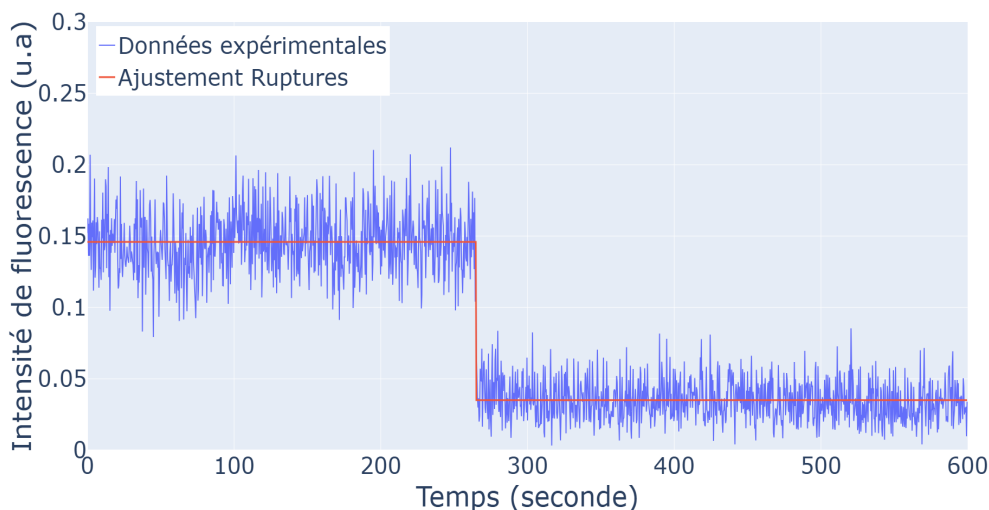
Ici, les marqueurs fluorescents ont été éclairés pendant 1600 images avec une puissance laser mesurée après le cube PBS de 5,4 mW selon le cycle suivant : 50 ms d'exposition et 375 ms de cycle. Sur 10 minutes, nous pouvons voir que le nombre de fluorophores est passé de plus de 1500 à près de 200 points. Cette forte diminution du nombre de points détectés est principalement due au mécanisme de photoblanchiment. Par ailleurs, toutes



les données de photophysique acquises suivent un premier protocole d'analyse afin de repérer les fluorophores et tracer les signaux d'intensité. Cette première étape d'analyse, commune aux expériences d'assemblage, est décrite à la section 3.4. Une analyse spécifique aux données de photophysique concerne l'élimination d'agrégats comme par exemple ceux entourés en jaune sur la Figure 3.34.A. En effet, la caractérisation du photoblanchiment et des clignotements nécessite de travailler à l'échelle du fluorophore unique. Ainsi, j'ai mis en place des seuils de sélection afin de ne conserver que des signaux présentant une diminution unique d'intensité (photoblanchiment) ou une alternance d'augmentation et de diminution (clignotements). Le protocole d'analyse des expériences de photophysique est décrit en détails à l'annexe E.

### 3.3.2.1 Photoblanchiment de l'Atto647N

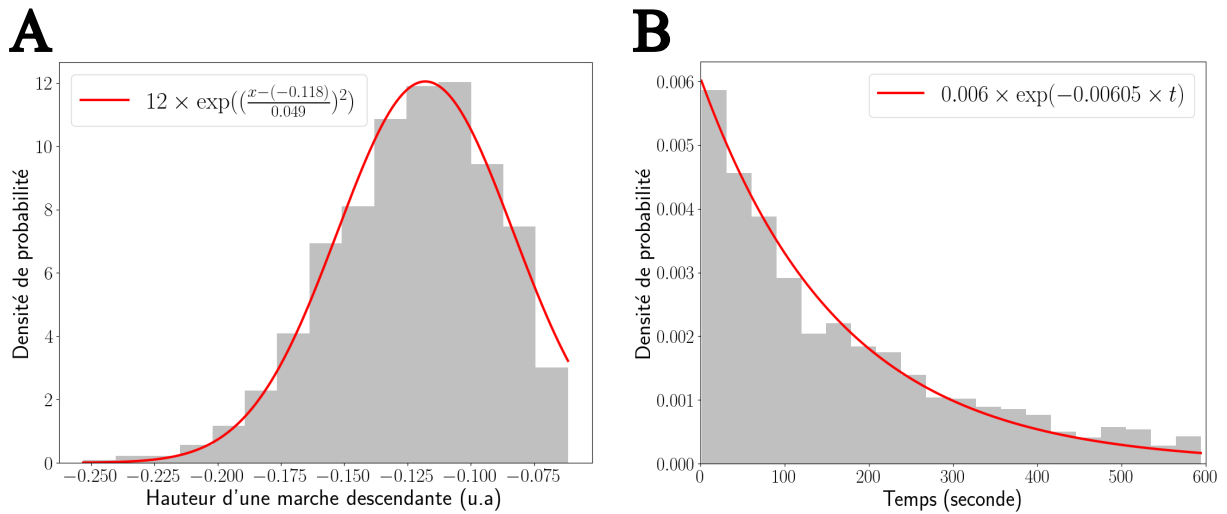
La Figure 3.35 montre un exemple de signal qui illustre le photoblanchiment de l'Atto647N :



**FIGURE 3.35** – **Observation du photoblanchiment sur l'Atto647N.** Évolution de l'intensité émise par un Atto647N au cours du temps : données expérimentales en bleu et ajustement en marches d'escalier en rouge. Conditions d'illumination : 1600 images, 50 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 5,4 mW.

Le signal rouge est obtenu en utilisant un algorithme de détection de marches sur lequel je reviens en détails au chapitre 4. Cet ajustement en marches d'escalier est très utile puisqu'il est appliqué automatiquement à tous les fluorophores repérés sur un champ de vue. Ici, cela me permet de repérer facilement l'instant auquel le photoblanchiment a lieu mais également d'estimer l'intensité émise par chaque molécule fluorescente en mesurant la hauteur de la marche descendante. En collectant ces deux informations pour chaque fluorophore repéré sur une cellule micro-fluidique, j'obtiens généralement les deux histogrammes de la Figure 3.36. J'ai également calculé le pourcentage de molécules fluo-

rescentes uniques qui ne s'éteignent jamais, noté  $F_\infty$ , en ne conservant que les signaux pour lesquels l'intensité mesurée varie peu par rapport à sa valeur en début d'acquisition.



**FIGURE 3.36** – **Caractérisation du photoblanchiment sur l'Atto647N.** A) Distribution de la hauteur d'une marche descendante : données expérimentales en gris et ajustement gaussien en rouge. B) Distribution des instants avant photoblanchiment : données expérimentales en gris et ajustement exponentiel en rouge. Conditions d'illumination : 1600 images, 50 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 5,4 mW.

La durée moyenne avant photoblanchiment  $\tau_b$  est obtenue en réalisant un ajustement exponentiel de la distribution des instants collectés. De cette façon, il est possible d'obtenir la constante de temps qui vaut ici environ 165 secondes. Pour la hauteur moyenne d'une marche descendante  $h$ , un ajustement gaussien nous donne 0,118 u.a. Le fait que la distribution des valeurs de  $h$  soit assez large peut principalement s'expliquer par les différentes directions d'émission du fluorophore et donc par son orientation. Enfin, le pourcentage de marqueurs immortels vaut environ 6,3 % ici.

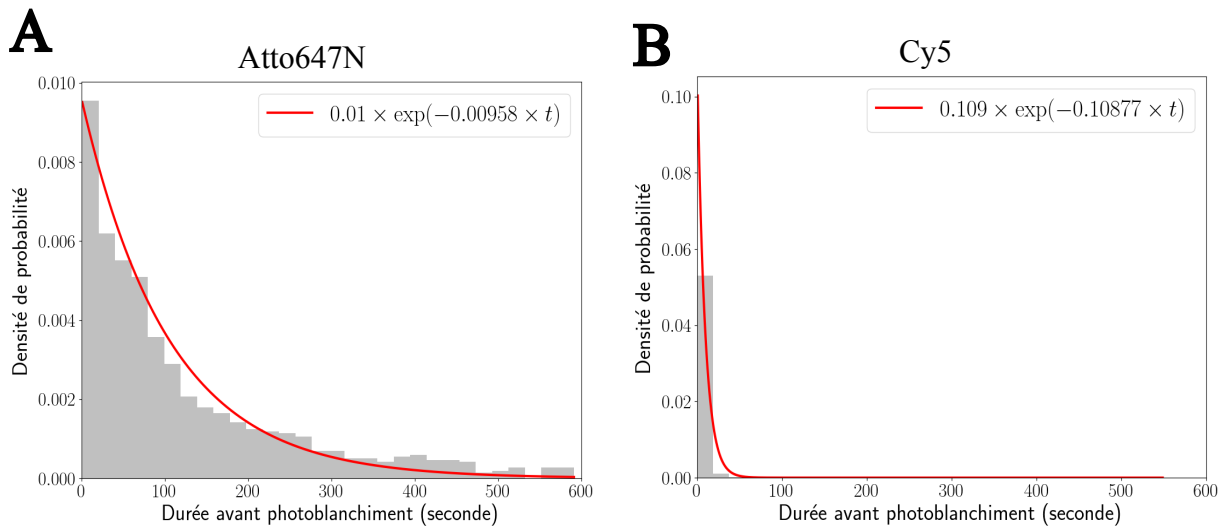
Afin de pouvoir interpréter ces résultats, le tableau 3.1 ci-dessous regroupe les résultats obtenus pour  $\tau_b$ ,  $h$  et  $F_\infty$  pour des conditions d'illumination différentes. **Pour faire simple, la durée moyenne avant photoblanchiment diminue et la hauteur des marches descendantes augmente lorsque le nombre de photons incidents croît.** En effet, pour une même durée de cycle, si le temps d'exposition ou la puissance incidente augmente, les fluorophores vont recevoir plus de photons et donc en absorber plus, ce qui va conduire à un photoblanchiment plus rapide et une émission plus importante (marches plus hautes). Par ailleurs, on peut remarquer que le pourcentage de fluorophores immortels suit la même évolution que la durée moyenne avant photoblanchiment. Ainsi, en suivant ce cheminement, réduire la durée d'illumination et/ou la puissance incidente permet de repousser le photoblanchiment et faire des films plus longs. Cependant, il faut faire attention de ne pas utiliser des conditions d'excitation trop faibles afin de conserver un rapport signal sur bruit convenable et de pouvoir détecter facilement les fluorophores

individuels. En général, les conditions d'illumination ont été choisies afin d'avoir une hauteur de marche minimale de 0,10 u.a. Pour la totalité des données du tableau, les valeurs de  $\tau_b$  et  $h$  correspondent aux paramètres caractéristiques des ajustements exponentiel et gaussien.

$N_{\text{images}}$	Temps de cycle	Temps d'exposition	$P_R$	$\tau_b$	$h$	$F_\infty$
1600	0,375 s	50 ms	3,7 mW	221 s	0,082 u.a	12 %
1600	0,375 s	50 ms	5,4 mW	165 s	0,118 u.a	6,3 %
1600	0,375 s	100 ms	3,7 mW	131 s	0,129 u.a	3,5 %
1600	0,375 s	100 ms	5,4 mW	71 s	0,153 u.a	1,0 %

**TABLEAU 3.1** – Influence des conditions d'illumination sur les paramètres de photoblanchiment de l'Atto647N.  $N_{\text{images}}$  correspond au nombre d'images par signal,  $P_R$  indique la puissance du laser rouge mesurée en sortie du cube PBS,  $\tau_b$  correspond à la durée moyenne avant photoblanchiment,  $h$  à la hauteur moyenne d'une marche descendante et  $F_\infty$  au pourcentage moyen de fluorophores immortels.

Pour conclure sur le photoblanchiment, je voulais insister sur la très bonne photostabilité de l'Atto647N en comparant ce dernier à un autre fluorophore fréquemment utilisé, le Cy5. Ce marqueur fluorescent, plus petit que l'Atto647N, possède également un pic d'absorption situé aux alentours de 647 nm. Ces deux fluorophores ont été comparés dans des conditions classiques d'illumination : 1600 images, 100 ms de temps d'exposition et 375 ms de temps de cycle.



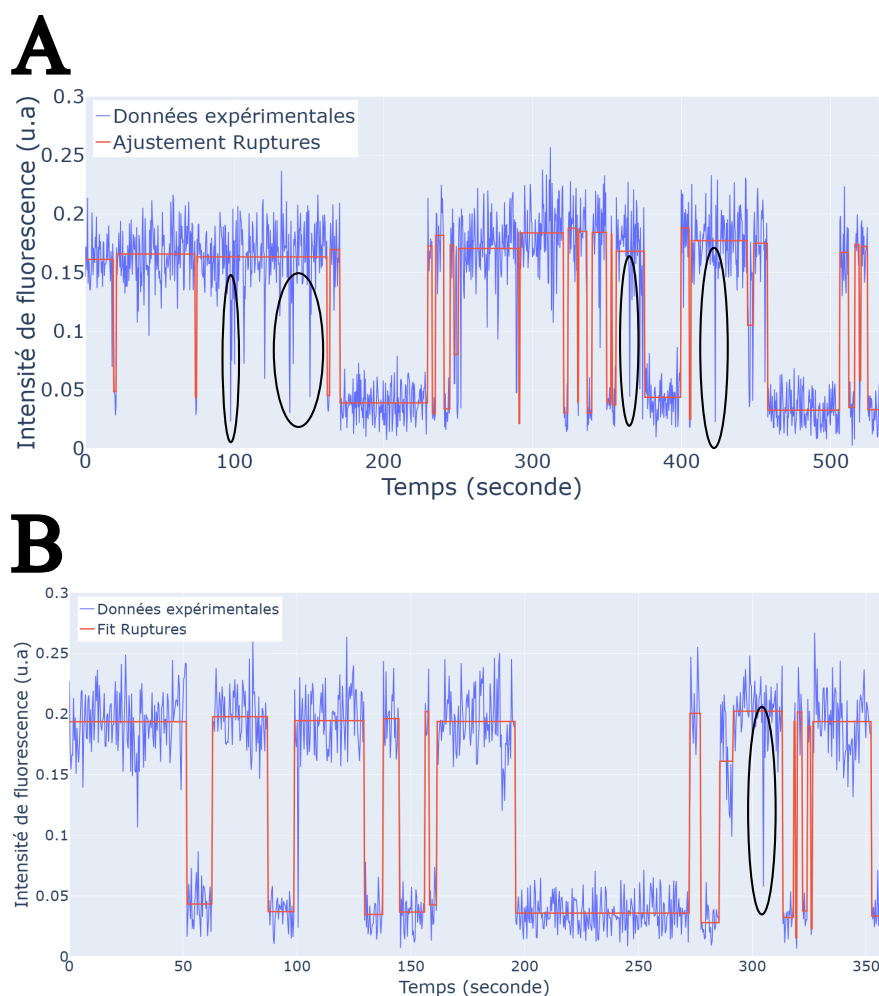
**FIGURE 3.37** – Comparaison de la durée moyenne avant photoblanchiment de l'Atto647N et du Cy5. Distributions des instants avant photoblanchiment de A) l'Atto647N et B) du Cy5 avec les données expérimentales en gris et les ajustements exponentiels en rouge. Conditions d'illumination : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 6,3 mW.

Dans cette expérience, la hauteur moyenne d'une marche descendante est relativement comparable dans les deux cas avec 0,162 u.a pour l'Atto647N et 0,135 u.a pour le Cy5.

En revanche, comme le montre la Figure 3.37, le paramètre  $\tau_b$  est très différent. En effet, la durée moyenne avant photoblanchiment est d'environ 104 secondes pour l'Atto647N contre seulement 9 secondes pour le Cy5. Les instants de départ sont tellement rapides que l'algorithme de détection utilisé a beaucoup de difficultés à détecter toutes les marches descendantes. Ainsi, même si la faible photostabilité du Cy5 empêche l'observation de clignotements, elle rend plus compliquée la détection d'événements d'accroche de protéines marquées. L'Atto647N est donc un très bon candidat pour l'étude du mécanisme d'auto-assemblage du CCMV.

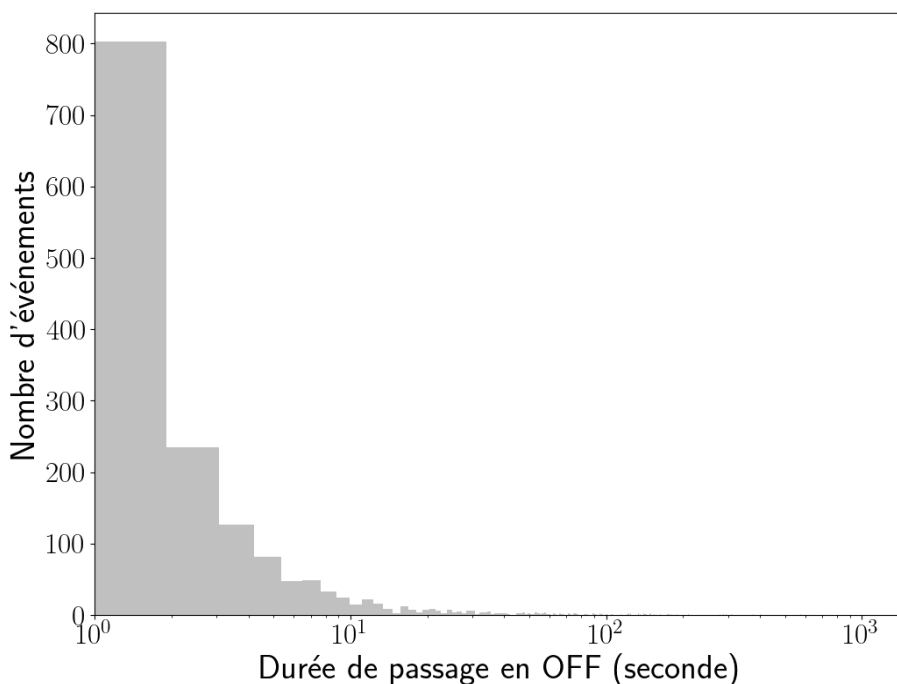
### 3.3.2.2 Clignotements de l'Atto647N

En parallèle du photoblanchiment, j'ai également pu observer plusieurs signaux dont l'évolution de l'intensité de fluorescence ressemble à celle représentée par la Figure 3.38 :



**FIGURE 3.38 – Observation de clignotements avec l'Atto647N.** A-B) Signaux expérimentaux avec des clignotements. En bleu, les données expérimentales et en rouge, les ajustements en marches d'escaliers. Conditions d'illumination : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW. Les événements entourés en noir représentent les passages en OFF qui ne durent qu'une image et qui n'ont pas été détectés par l'algorithme.

On remarque essentiellement deux choses en regardant cette figure : le nombre de passages en OFF, noté  $n_{cl}$ , peut être important avec 19 événements repérés en haut contre 11 en bas, et les durées d'absence d'émission, notées  $t_{cl}$ , peuvent aller de quelques images à plusieurs centaines. Cette seconde observation est particulièrement gênante pour l'analyse des expériences d'assemblage puisque les clignotements peuvent être courts (moins de 5 images), assez longs (entre 5 et 50 images) voire également très longs (plus de 50 images). En effet, avec une telle dispersion, il n'est pas possible de seulement lisser les signaux afin de retirer les événements courts. Cette hétérogénéité est illustrée tout aussi visuellement par la Figure 3.39.



**FIGURE 3.39** – **Caractérisation des durées de passages en OFF pour l'Atto647N.** Conditions d'illumination : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

Ici, la valeur moyenne de  $t_{cl}$  est de 35 secondes, soit environ 93 images. Par ailleurs, 48 % des clignotements repérés durent moins de 5 images alors qu'ils sont 89 % à être plus courts que 50 images. Cela signifie que les événements très longs sont loins d'être majoritaires mais qu'ils doivent être pris en compte. Plus globalement, dans ces conditions expérimentales, le pourcentage de fluorophores avec au moins un clignotement, noté  $Atto_{cl}$ , est de 16 % : il semble donc difficile de négliger ce mécanisme. **Le tableau 3.2 regroupe les principales informations collectées sur le clignotement pour différentes illuminations.** On remarque qu'il est difficile d'observer une tendance particulière en fonction des conditions d'illumination que ce soit pour  $t_{cl}$ ,  $n_{cl}$  ou  $Atto_{cl}$ .

Pour finir, je précise qu'ici, l'algorithme de détection ne permet pas de repérer des ruptures séparées par moins de deux images. En effet, la Figure 3.38 montre des dimi-

nutions d'intensité entourées en noir qui ne durent qu'une image mais qui n'ont pas été détectées par l'algorithme. Si l'on souhaite prendre en compte ces événements, une méthode simple serait de faire de nouvelles expériences mais en réduisant la durée de cycle afin de mieux résoudre ces passages en OFF qui ne durent qu'une image. Cela pourrait permettre d'avoir accès à une population plus importante de la distribution totale des clignotements et de mieux caractériser ces derniers.

$N_{\text{images}}$	Temps de cycle	Temps d'exposition	$P_R$	$t_{cl}$	$n_{cl}$	$Atto_{cl}$
1600	0,375 s	50 ms	3,7 mW	33 s	1,8	21 %
1600	0,375 s	50 ms	5,4 mW	26 s	2,1	18 %
1600	0,375 s	100 ms	3,7 mW	35 s	1,8	16 %
1600	0,375 s	100 ms	5,4 mW	27 s	2,6	21 %

**TABLEAU 3.2 – Influence des conditions d'illumination sur les paramètres de clignotements de l'Atto647N.**  $N_{\text{images}}$  correspond au nombre d'images par signal,  $P_R$  indique la puissance du laser rouge mesurée en sortie du cube PBS,  $t_{cl}$  correspond à la durée moyenne de passage en OFF,  $n_{cl}$  au nombre moyen de passages en OFF et  $Atto_{cl}$  au pourcentage de fluorophores qui clignent.

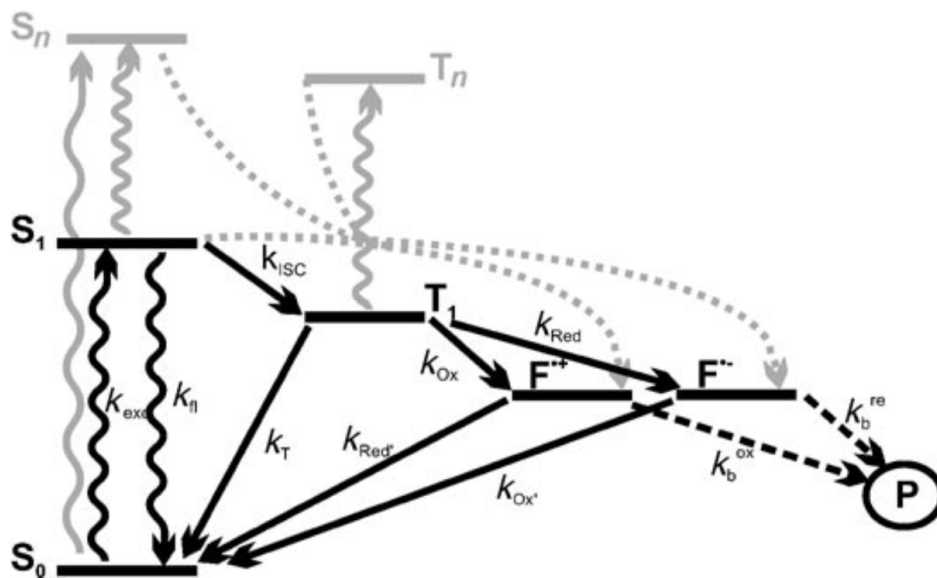
### 3.3.3 Vers l'amélioration de la photostabilité de l'Atto647N

Les mécanismes de photoblanchiment et clignotements caractérisés chez l'Atto647N risquent donc de rendre plus difficile l'analyse des signaux obtenus après une expérience d'assemblage. De plus, alors que les conditions d'illumination permettent de contrôler assez facilement le temps caractéristique de photoblanchiment, elles ne semblent pas avoir un effet clair sur les clignotements. En cherchant dans la littérature, il apparaît qu'une autre façon d'améliorer la photostabilité d'un fluorophore est de modifier l'environnement de ce dernier. En effet, comme je l'ai mentionné à la section 3.3.1, les interactions chimiques avec les molécules présentes en solution participent au passage d'un fluorophore dans des niveaux d'énergie différents de  $S_1$ , affectant ainsi son émission de photons. Dans les deux sous-parties suivantes, je m'intéresse à la modification de la photostabilité de l'Atto647N quand l'oxygène du milieu est retiré.

#### 3.3.3.1 Mécanisme général d'un système désoxygénant

Le rôle de l'oxygène dans la modification des propriétés photophysiques des fluorophores est assez complexe. En effet, il contribue à l'apparition du photoblanchiment, mais il libère les marqueurs fluorescents piégés dans l'état triplet, réduisant ainsi le temps de résidence dans un état noir durant un clignotement. Cela signifie que si l'on souhaite éliminer l'oxygène du milieu pour repousser le photoblanchiment, il faut également lui trouver une alternative afin de ne pas favoriser les clignotements. Vogelsang et al. (2008) ont proposé une solution générale du nom de ROXS (*Reducing and OXidizing System*) dont le fonctionnement est décrit par la Figure 3.40.

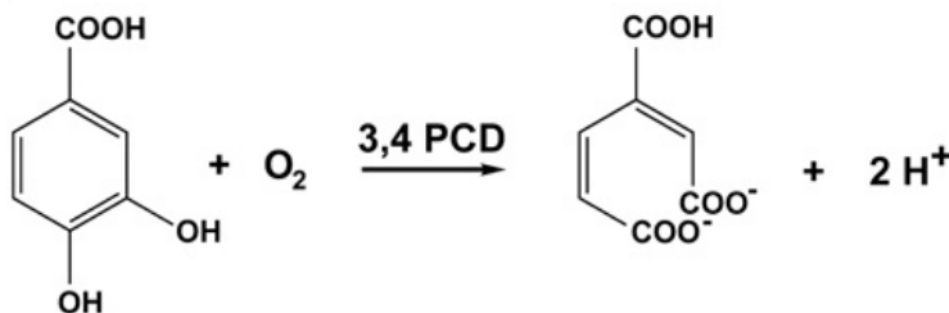
En adoptant cette stratégie, si le fluorophore se retrouve bloqué dans l'état  $T_1$  et que l'oxygène a été éliminé du milieu, un couple oxydant-réducteur va pouvoir agir à sa place. Le marqueur fluorescent peut soit être oxydé et se retrouver à l'état de radical cationique  $F^{\bullet+}$ , soit être réduit et ainsi passer dans un état de radical anionique  $F^{\bullet-}$ . Par la suite, le radical cationique pourra alors être réduit et le radical anionique oxydé, afin de retourner dans l'état fondamental  $S_0$ . Ces réactions sont considérées comme étant très rapides, ce qui limite les effets du clignotement et rend peu favorables les potentiels mécanismes à l'origine du photoblanchiment.



**FIGURE 3.40 – Fonctionnement du système ROXS.** Les niveaux d'énergie  $F^{\bullet+}$  et  $F^{\bullet-}$  correspondent respectivement aux radicaux cationiques et anioniques obtenus via des réactions d'oxydation  $k_{Ox}$  et de réduction  $k_{Red}$ . Les réactions notées  $k_b^{ox}$  et  $k_b^{re}$  caractérisent les mécanismes à l'origine du photoblanchiment des radicaux cationiques et anioniques. Le taux de désexcitation de l'état triplet est noté  $k_T$ . Les réactions représentées en gris correspondent à des mécanismes qui impliquent des niveaux d'énergie supérieurs  $S_n$  et  $T_n$ . Figure adaptée de Vogelsang et al. (2008).

Dans les travaux de Vogelsang et al. (2008), l'oxygène a été éliminé grâce au système enzymatique GODCAT (glucose oxydase et catalase) et les conditions ROXS ont été obtenues à partir du couple oxydant-réducteur composé du méthyl-viologène (MV) et de l'acide ascorbique (AA). Dans ces conditions, ils ont montré que le nombre moyen de photons émis par l'Atto647N est multiplié par un facteur 800 et que les durées de passage en OFF sont quasiment nulles pour des concentrations d'une dizaine de micromolaires en AA et MV. Cependant, comme le mécanisme réactionnel du GODCAT peut entraîner la formation de peroxyde d'hydrogène résiduel hautement réactif (Englander et al. (1987)), d'autres stratégies de dépletion en oxygène ont été adoptées. Aitken et al. (2008) ont montré que le couple PCA/PCD (acide protocatéchique / protocatéchuate-3,4-dioxygénase) permet de prolonger plus longtemps la durée avant photoblanchiment des fluorophores

Cy3, Cy5 et Alexa488. Par ailleurs, que ce soit pour le GODCAT ou le PCA/PCD, le couple AA/MV a été remplacé par l'agent réducteur du nom de Trolox. Ce dernier a également été utilisé durant les travaux de Vogelsang et al. (2008) et ses performances sont quasiment comparables à celles du système ROXS obtenu par l'association AA/MV. Le mécanisme réactionnel qui se produit à partir du couple PCA/PCD est illustré à la Figure 3.41.



**FIGURE 3.41** – Mécanisme réactionnel entre le PCA et l'oxygène catalysé par le PCD. Figure adaptée de Aitken et al. (2008).

Si la stratégie de déplétion en oxygène semble pouvoir être facilement adaptée, le choix du couple oxydant-réducteur est très dépendant du fluorophore utilisé. En effet, Vogelsang et al. (2008) ont observé que l'amélioration de la photostabilité n'est pas obtenue avec la même efficacité pour toutes les molécules fluorescentes et que dans le cas du MR121, le système ROXS ne semble pas avoir d'effet. Les auteurs ont proposé une explication en se basant sur des calculs thermodynamiques et ont souligné l'importance de comparer les potentiels redox du fluorophore étudié avec ceux des agents oxydants et réducteurs. **Leurs calculs ont montré que les réactions d'oxydation et de réduction de l'état triplet de l'Atto647N, mais également les réactions complémentaires permettant le retour dans l'état fondamental  $S_0$ , présentent toutes une énergie libre négative, signe de réactions favorables.** En revanche, le MR121 renvoie une énergie libre positive, ce qui corrobore l'absence d'effet du système ROXS à partir du couple AA/MV. Dans la thèse d'Antoine Le Gall, un tableau récapitulatif donne les valeurs d'énergie libre de plusieurs fluorophores lorsque l'acide ascorbique et le méthylviologène sont utilisés en tant qu'oxydant et réducteur (Le Gall (2011)). En outre, l'étude menée par Vogelsang et al. (2009) a également mis en lumière que la modulation des concentrations d'agents oxydants et réducteurs (AA et MV) permet de contrôler la durée de passages en OFF pendant un clignotement. En résumé, il est donc important de bien connaître les différentes espèces chimiques que l'on souhaite utiliser afin d'améliorer la photostabilité d'un fluorophore.

Dans le tableau 3.3, je renseigne les compositions des différentes solutions tampons



testées afin d'appliquer le système ROXS à l'Atto647N. Toutes ces solutions utilisent le système désoxygénant PCA/PCD. Les concentrations en Trolox réellement présentes dans les tampons ROXS-A' et ROXS-C' ne sont peut-être pas conformes avec celles énoncées dans le tableau ci-dessous. En effet, j'ai préparé une solution à 75 mM en MES et 3 mM en Trolox à pH relativement acide (6,1) afin de favoriser la dissolution de l'agent réducteur. Après agitation, le Trolox semblait s'être totalement dissous mais une fois la filtration terminée, le filtre utilisé était entièrement gris, qui est également la couleur de la poudre de Trolox initiale. Cette observation doit permettre de relativiser les résultats obtenus avec les tampons ROXS-A' et ROXS-C' que je présente dans la sous-partie suivante.

NOM	COMPOSITION	pH
Tampon B'	50 mM MES, 100 mM NaCl	6,1
Tampon ROXS-A	50 mM MES, 100 mM NaCl, 10 mM PCA, 200 nM PCD, 1 mM AA, 1 mM MV	6,1
Tampon ROXS-A'	39 mM MES, 100 mM NaCl, 10 mM PCA, 200 nM PCD, 1,6 mM Trolox	6,1
Tampon ROXS-C	50 mM MES, 100 mM NaCl, 2,5 mM PCA, 50 nM PCD, 1 mM AA, 1 mM MV	6,1
Tampon ROXS-C'	50 mM MES, 100 mM NaCl, 2,5 mM PCA, 50 nM PCD, 2 mM Trolox	6,1

**TABLEAU 3.3** – **Système ROXS : composition des solutions tampons utilisées.**

### 3.3.3.2 Application expérimentale sur l'Atto647N

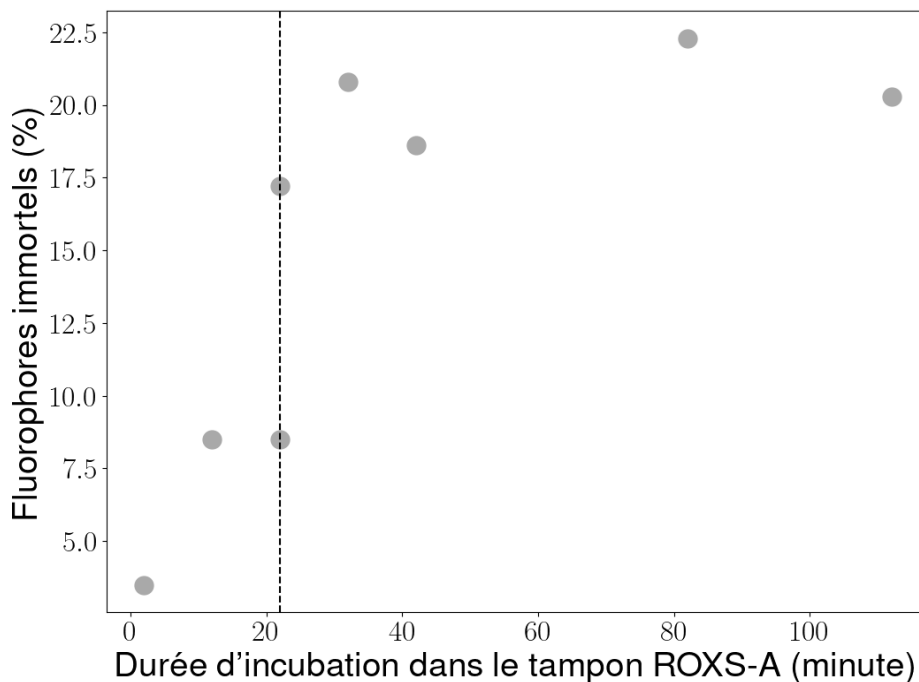
Les quatre tampons ROXS ont été utilisés afin de préparer des solutions à 1 nM d'Atto647N avec une modification de type biotine. La photostabilité de ce fluorophore a ensuite été évaluée pour quatre conditions différentes d'illumination qui sont décrites dans le tableau 3.4 ci-dessous :

Nom	N <sub>images</sub>	Temps de cycle	Temps d'exposition	P <sub>R</sub>
Illum-1	1600	0,375 s	50 ms	3,7 mW
Illum-2	1600	0,375 s	50 ms	5,4 mW
Illum-3	1600	0,375 s	100 ms	3,7 mW
Illum-4	1600	0,375 s	100 ms	5,4 mW

**TABLEAU 3.4** – **Système ROXS : conditions expérimentales d'illumination.** N<sub>images</sub> correspond au nombre d'images par signal et P<sub>R</sub> indique la puissance du laser rouge mesurée en sortie du cube PBS.

En ce qui concerne le protocole expérimental, les marqueurs fluorescents ont été incubés à température ambiante pendant des durées variables. De plus, pour une solution tampon donnée, j'ai réalisé plusieurs acquisition de dix minutes en changeant de champ de vue entre chaque film afin de repartir avec des fluorophores qui n'ont pas été éclairés. Les

performances des systèmes ROXS ont été évaluées à partir des paramètres caractéristiques du photoblanchiment et des clignotements mentionnés dans les sections précédentes.



**FIGURE 3.42** – Évolution de  $F_\infty$  dans le tampon ROXS-A en fonction de la durée d'incubation. Les résultats obtenus pour chaque point correspondent à une seule acquisition effectuée dans les conditions d'illumination 4. Les pointillés présents à 22 minutes mettent en évidence la transition vers une augmentation significative de  $F_\infty$ .

Les résultats des expériences initiales étaient assez surprenants puisque pour une durée d'incubation de 2 minutes, la photostabilité de l'Atto647N n'a absolument pas évolué entre un des tampons classiques d'assemblage ( $B' = 50$  mM MES, 100 mM NaCl à pH 6,1) et ROXS-A. En effet, pour les conditions d'illumination 4, le pourcentage de fluorophores immortels est passé de 1,6 à 3,5 %. J'ai donc voulu voir si une durée d'incubation plus longue pouvait changer quelque chose. Sur la Figure 3.42 je trace l'évolution du pourcentage de fluorophores immortels  $F_\infty$  en fonction de la durée d'incubation. Chaque point observé sur cette figure correspond aux résultats obtenus à partir d'une acquisition, et **la durée d'incubation suffisante afin de maximiser  $F_\infty$  semble se situer aux alentours de 30 minutes**. En effet, les acquisitions faites après 32 - 42 - 82 et 112 minutes renvoient une valeur de  $F_\infty$  d'environ 20 % contre à peine 3 % après 2 minutes d'incubation. Pour les deux expériences avec un temps d'incubation de 22 minutes, les résultats varient, ce qui peut être le signe d'une efficacité du système ROXS qui n'est pas encore totalement atteinte.

Ainsi, il semblerait que le tampon ROXS-A permette de repousser efficacement le photoblanchiment. Il est donc intéressant de regarder plus en détails les autres paramètres de photostabilité et ce pour tous les tampons ROXS. Ces informations sont résumées dans

le tableau 3.5 pour lequel les durées d'incubation considérées sont d'au moins 30 minutes pour le tampon ROXS-A et au moins 20 minutes pour les trois autres.

Illumination	Tampon	$\tau_b$	$F_\infty$	$t_{cl}$	$n_{cl}$	$Atto_{cl}$
Illum-1	ROXS-A	361 s	47 %	30 s	2,3	38 %
	B'	221 s	12 %	33 s	1,8	21 %
Illum-2	ROXS-A	286 s	39 %	25 s	2,6	41 %
	B'	165 s	6,3 %	26 s	2,1	18 %
Illum-3	ROXS-A	305 s	28 %	24 s	2,9	33 %
	ROXS-C	159 s	8,3 %	21 s	2,0	22 %
	ROXS-C'	54 s	0,7 %	34 s	1,5	9,7 %
	B'	131 s	3,5 %	35 s	1,8	16 %
Illum-4	ROXS-A	244 s	20 %	23 s	3,1	31 %
	ROXS-A'	33 s	0,4 %	31 s	1,5	7,3 %
	ROXS-C	89 s	1,8 %	17 s	1,8	17 %
	ROXS-C'	36 s	0,5 %	26 s	1,7	11 %
	B'	71 s	1,0 %	27 s	2,6	21 %

**TABLEAU 3.5** – **Système ROXS : résultats expérimentaux.**  $\tau_b$  correspond à la durée moyenne avant photoblanchiment,  $F_\infty$  au pourcentage moyen de fluorophores immortels,  $t_{cl}$  correspond à la durée moyenne de passage en OFF,  $n_{cl}$  au nombre moyen de passages en OFF et  $Atto_{cl}$  au pourcentage de fluorophores qui clignotent.

Pour résumer, il semblerait bien que le tampon ROXS-A puisse ralentir l'apparition du photoblanchiment quelles que soient les conditions d'illumination utilisées. En effet, les durées moyennes avant photoblanchiment et les pourcentages de fluorophores suivent une amélioration non négligeable par rapport aux expériences réalisées dans le tampon B'. En ce qui concerne les clignotements, les conclusions sont moins réjouissantes. En effet, le nombre de passages en OFF ainsi que les durées de ces événements n'ont pas beaucoup évolué par rapport à ce qui avait été observé dans le tampon B'. Par ailleurs, le photoblanchiment ayant été repoussé, le nombre moyen de fluorophores qui clignotent a augmenté pour chacune des conditions d'illumination, en passant par exemple de 18 à 41 % pour la condition Illum-2. **Au vu de ces résultats, il semble assez évident que les conditions d'oxydo-réduction de ROXS-A ne permettent pas de réduire les clignotements de l'Atto647N.** Cela peut paraître surprenant au regard des résultats très positifs obtenus par Vogelsang et al. (2008) dans ces mêmes conditions.

Si l'on se penche désormais sur les autres tampons ROXS, on peut remarquer que peu importe les conditions d'illumination, la photostabilité de l'Atto647N n'a pas été améliorée. En effet, que ce soient les paramètres caractéristiques du photoblanchiment ou ceux des clignotements, ils sont tout à fait comparables à ceux obtenus dans le tampon B'. Pour les tampons ROXS-A' et ROXS-C', cette absence d'effet sur la photophysique de l'Atto647N peut provenir, comme je l'ai dit plus tôt, d'une méconnaissance de la concen-

tration réelle en Trolox. Il est intéressant de remarquer qu'avec ces deux tampons,  $\tau_b$ ,  $F_\infty$  et  $Atto_{cl}$  ont diminué par rapport à ce qui avait été observé dans le tampon B'. Ces trois observations semblent indiquer une absence de Trolox dans le milieu. Ainsi, sans espèce chimique capable de remplacer l'oxygène consommé par le système PCA/PCD, les fluorophores semblent se retrouver plus facilement bloqués dans l'état triplet  $T_1$ . En revanche, pour le tampon ROXS-C, la diminution des concentrations en PCA/PCD semble avoir affecté l'efficacité du mécanisme réactionnel mis en place. Il est possible que l'élimination de l'oxygène du milieu ne soit plus aussi effective ce qui fait que les différences avec le tampon B' classique sont très faiblement marquées. Pourtant, Aitken et al. (2008) ont montré qu'avec ces mêmes concentrations en PCA/PCD et 1 mM de Trolox, le photoblanchiment est retardé pour le Cy3, Cy5 et Alexa488, malgré une inefficacité vis-à-vis des clignotements.

**En somme, les tests réalisés avec les différents tampons ROXS n'ont pas donné les résultats escomptés. En effet, un seul tampon sur les quatre utilisés montre des effets positifs vis-à-vis du photoblanchiment mais aucun ne semble résoudre le problème des clignotements.** Selon les concentrations en PCA/PCD choisies ou les agents oxydants et réducteurs utilisés, des contradictions ont été obtenues par rapport à ce que Vogelsang et al. (2008) et Aitken et al. (2008) ont pu observer. Néanmoins, comme je l'ai dit à la section 3.3.1, les mécanismes affectant la photostabilité des molécules fluorescentes sont très dépendants des conditions d'illumination et des propriétés chimiques du milieu. En effet, même si les concentrations des tampons ROXS sont identiques à celles retrouvées dans la littérature, les expériences en elles-mêmes ont été réalisées dans des conditions expérimentales différentes (puissance laser, pH) afin de correspondre à celles utilisées durant les expériences d'assemblage. Cela peut expliquer les différences rencontrées ici.

Ainsi, l'absence d'effets notables sur les clignotements et le fait de devoir attendre 30 minutes<sup>3</sup> pour repousser au mieux le photoblanchiment ont fait que nous avons décidé de ne pas appliquer le système ROXS à nos expériences. La photostabilité de l'Atto647N présentée à la section 3.3.2 est celle qu'il faut donc considérer pour les expériences d'assemblage, auxquelles les clignotements sont totalement intégrés. Par ailleurs, le photoblanchiment est caractérisé à la fin de chacune des expériences décrites à la section 5.4.

### 3.3.4 Stratégie de marquage choisie

Pour terminer sur la caractérisation de l'Atto647N, je souhaite revenir sur le marquage des dimères. Avant réaction, le fluorophore possède une modification du type NHS-Ester à

---

3. Des tests réalisés en toute fin de thèse ont montré qu'en oxydant le Trolox (dissolution + agitation pendant 18 heures à température ambiante), 40 à 50 % des fluorophores éclairés ne blanchissent pas dans les conditions d'illumination 3, pour une préparation dans un tampon similaire à ROXS-A' (différences : 50 mM MES et 1 mM Trolox) et 2-3 minutes d'incubation dans la cellule. Aucun effet observé vis-à-vis des clignotements.

l'une de ses extrémités. Ensuite, durant la période d'incubation avec les virions de CCMV, la fonction NHS-Ester va réagir avec l'une des deux lysines exposées sur la partie extérieure du corps de la protéine de CCMV (acides aminés 106 et 131 sur 190). La concentration en protéines étant en excès par rapport à l'Atto647N au moment du marquage (voir annexe A), la majorité des dimères devraient être marqués par un seul fluorophore. Le mécanisme réactionnel est décrit par la Figure 3.43.

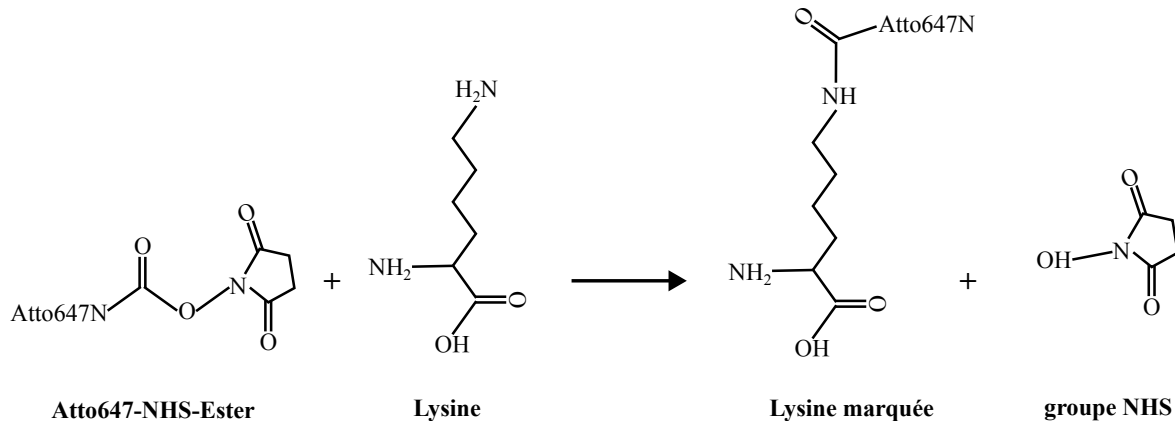
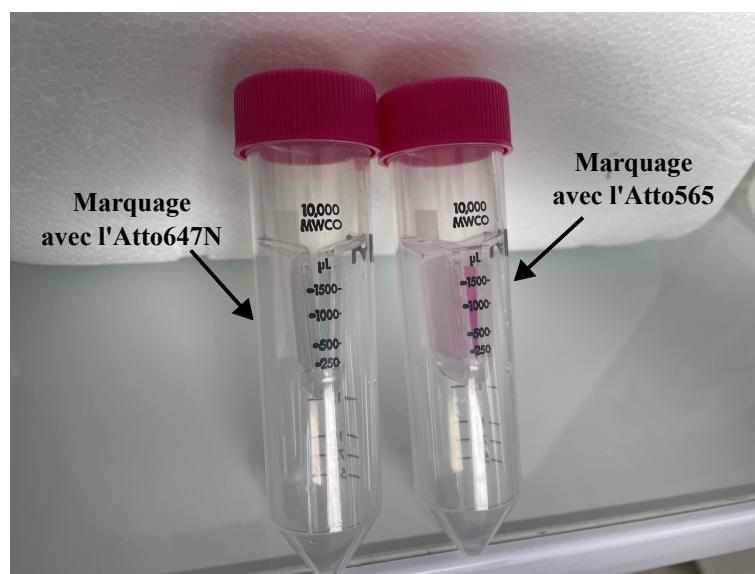


FIGURE 3.43 – Mécanisme réactionnel entre une fonction NHS-Ester et une lysine.

Il est important que cette réaction ait lieu avec l'une des lysine présentes sur le corps de la protéine et donc que le marquage soit réalisé lorsque les virions sont encore assemblés. En effet, si les dimères de capsidite sont marqués lorsqu'ils sont libres en solution, les trois lysines du domaine N-Terminal (Speir et al. (1995)) peuvent être pris pour cible par le marqueur. Or, nous avons vu à la section 2.1.4.2 que les interactions électrostatiques entre le matériel génétique chargé négativement et la partie N-terminale chargée positivement favorisent l'assemblage de la capsidite de CCMV. Si des marqueurs fluorescents viennent se fixer au niveau de ces trois lysines, il y a des risques pour que l'assemblage en soit perturbé. Par ailleurs, le fournisseur chez lequel nous achetons l'Atto647N-NHS-Ester, *ATTO-TEC*, propose un protocole afin de favoriser le marquage des protéines d'intérêt. Une solution tampon de pH compris entre 8 et 9 permet à l'amine de la chaîne latérale de la lysine d'être suffisamment déprotonée, ce qui amène une réaction rapide avec la fonction NHS-Ester. C'est donc pour cela que le pH en début de purification des protéines est légèrement supérieur à celui du reste du protocole (voir annexe A).

Enfin, le choix du fluorophore utilisé va jouer un rôle sur la charge du complexe formé après marquage. Par exemple, *ATTO-TEC* précise qu'un couplage entre une protéine et un Atto647N induit une augmentation de la charge d'une unité alors qu'un Atto565 n'amène aucun changement. Ces deux marqueurs fluorescents ont été utilisés conjointement lors d'une purification afin d'augmenter la concentration en protéines marquées sans que les accroches non spécifiques n'évoluent. De cette façon, il

aurait été possible de marquer deux lots différents de protéines avec un taux de marquage individuel de 10 % puis d'ajuster les concentrations *a posteriori* pour obtenir un taux de marquage global de 10 %. Seulement 5 % des protéines auraient été marquées dans chaque couleur, réduisant le nombre d'accroches non spécifiques visibles dans le rouge. Cependant, en fin de purification, en concentrant le lot de protéines marquées par l'Atto565 avec un Falcon de dialyse, la membrane de ce dernier était entièrement rose. En comparaison, la membrane était très légèrement bleutée pour le lot de protéines marquées par l'Atto647N (Figure 3.44). Ces observations nous indiquent que la concentration des espèces marquées en solution a probablement été réduite à la suite de l'étape de reconcentration, et ce principalement pour le lot contenant l'Atto565.



**FIGURE 3.44** – Comparaison des membranes de dialyse pour deux lots de protéines marquées différemment. Le Falcon de gauche a été utilisé pour le lot marqué avec l'Atto647N et celui de droite avec l'Atto565.

Les mesures d'absorbance faites au Nanodrop ont confirmé cette hypothèse. En effet, l'absorbance à 565 nm était de 0,15 contre 0,87 à 647 nm<sup>4</sup>, ce qui correspond à des taux de marquage respectifs de 2,6 et 12,1 %. Ainsi, la membrane de dialyse du Falcon semble avoir interagi différemment avec les protéines marquées par l'Atto565.

### 3.4 Premières étapes d'analyse des données brutes

Les résultats présentés dans la section précédente ont été obtenus après deux étapes communes à tout type d'expérience décrite dans ce manuscrit. La première consiste à retravailler les acquisitions afin de pouvoir les analyser plus facilement. Il s'agit notamment de la correction des dérives mécaniques mais également de la réduction du bruit retrouvé

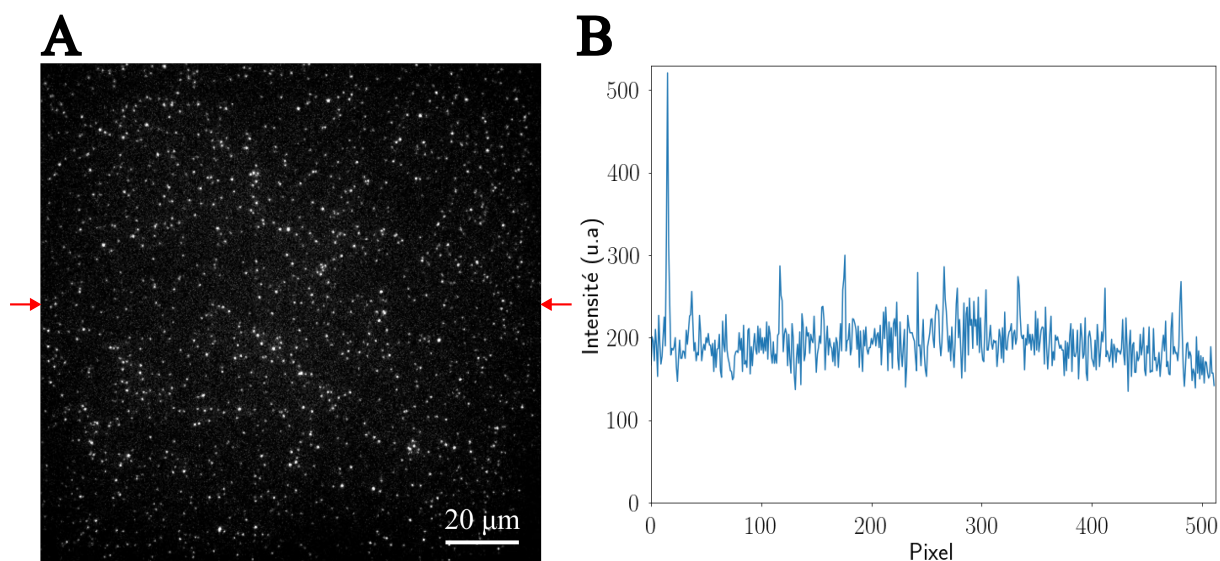
4. Le coefficient d'extinction molaire de l'Atto565 à sa longueur d'onde maximale d'absorption, noté  $\epsilon_{f,max}$ , vaut  $1,2 \times 10^5$  L.mol<sup>-1</sup>.cm<sup>-1</sup> contre  $1,5 \times 10^5$  L.mol<sup>-1</sup>.cm<sup>-1</sup> pour l'Atto647N.

au sein des images brutes. En ce qui concerne la seconde étape, elle correspond à la détection des molécules fluorescentes via un algorithme Python. L'objectif final est de colocaliser les particules détectées dans le rouge et le vert, afin d'avoir accès au suivi de l'intensité de fluorescence émise par des protéines marquées sur un site d'ARN.

Dans cette section je décris plus en détails ces deux étapes réalisées à la suite des acquisitions.

### 3.4.1 Traitement des films enregistrés

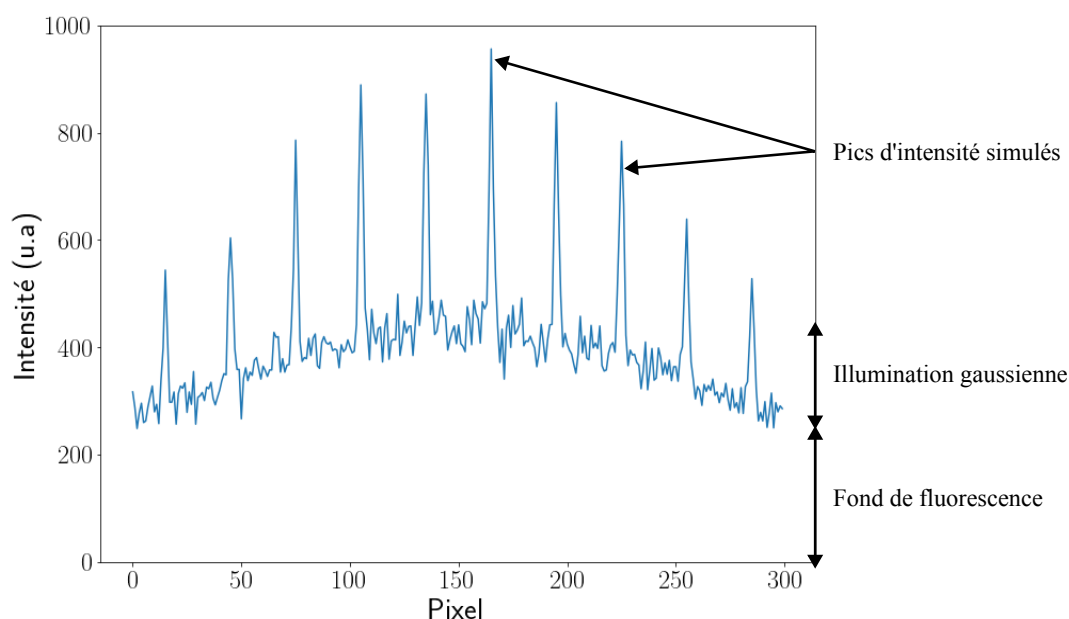
Les différents champs de vue observés jusqu'à présent correspondent aux acquisitions brutes enregistrées pendant une expérience. Ces images paraissent assez satisfaisantes afin de les utiliser pour des analyses plus précises puisqu'à l'oeil, il semble que nous soyons capables de repérer facilement les échantillons marqués à la surface. De plus, la couverture en marqueurs fluorescents de la Figure 3.45.A nous indique que l'intensité du faisceau d'excitation semble uniforme sur la totalité du champ de vue, et ce grâce à l'illumination *Flat-top* présentée à la section 3.1.2.1. Cependant, en réalisant une coupe selon la direction horizontale de l'image, cette affirmation n'est plus tout à fait vraie comme le montre la Figure 3.45.B



**FIGURE 3.45** – **Observation du fond de fluorescence sur une image brute.** A) Image brute où les flèches rouges montrent la position de la coupe d'intensité (B) réalisée selon la direction horizontale du champ de vue. Il s'agit d'une image prise 5 minutes après le début d'une acquisition de 10 minutes, au cours de laquelle la photophysique de l'Atto647N a été étudiée. Les conditions d'acquisition sont les suivantes : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

Grâce à cette coupe horizontale, nous pouvons repérer deux éléments importants : une ligne de base dont l'intensité varie entre 150 et 200 niveau de gris et des pics qui caractérisent principalement les marqueurs fluorescents repérés à la surface. Premièrement, le fait

que le niveau moyen d'intensité ne soit pas plus bas correspond au fond de fluorescence. Ce dernier peut s'expliquer par une fluorescence résiduelle de la lamelle de verre ou des espèces chimiques non marquées. Cependant l'enveloppe moyenne du signal mesurée au niveau de la coupe est légèrement plus faible sur les côtés qu'au centre. Ce manque d'uniformité vis-à-vis de l'illumination met en évidence certaines limites de notre configuration *Flat-top* et peut avoir des conséquences sur l'analyse. Comme je l'ai expliqué au début de la section 3.3.2, nous sommes intéressés par les variations soudaines d'intensité de fluorescence provenant de l'interaction entre protéines marquées et ARN. Il est donc important que ces événements, issus d'un fluorophore unique, aient la même hauteur quelle que soit la position de l'ARN simple brin sur le champ de vue. Or, la Figure 3.36.A a bien montré que la dispersion de l'intensité moyenne émise par un fluorophore unique est déjà suffisamment importante du fait de l'hétérogénéité de son orientation d'émission. Cette disparité ne devant pas être accentuée par le profil du faisceau d'excitation, l'illumination dite *Flat-top* a été mise en place sur le montage. Néanmoins, malgré cette précaution, l'intensité n'est toujours pas parfaitement uniforme sur la totalité du champ de vue.



**FIGURE 3.46** – Simulation de l'intensité d'une image brute. Ce signal a été simulé à partir d'un fond de fluorescence et d'un bruit gaussiens.

La Figure 3.46 illustre les trois composantes prédominantes observées sur une image brute : des pics d'intensité relatifs aux marqueurs, une inhomogénéité de l'illumination (ici simulée comme étant gaussienne) et un fond de fluorescence. Ces données ont été simulées par Véronique Vienne au cours de sa thèse. La méthodologie détaillée ci-dessous a donc pour but d'aplatir le fond de fluorescence et faire en sorte que l'illumination soit encore plus homogène sur l'entièreté du champ de vue. Le traitement d'images contient quatre étapes successives, toutes réalisées sur Python.



- Blurring :

La première étape consiste à renormaliser l'enveloppe moyenne du fond afin d'avoir une illumination homogène. Pour ce faire, il faut commencer par flouter la totalité de l'image brute en moyennant localement l'intensité sur une zone carrée de taille  $N \times N$  pixels. Par la suite, l'intensité de chaque pixel de l'image bruitée est divisée par la valeur d'intensité présente aux mêmes coordonnées mais appartenant à l'image floutée. La renormalisation est donc dépendante de la moyenne locale faite en chaque zone du champ de vue. Le niveau de l'enveloppe étant inhomogène sur toute l'image, les pics seront donc plus aplatis là où l'intensité moyenne est la plus forte, à savoir au centre. Cela nous permet donc de travailler avec des molécules uniques dont les pics sont quasiment de même hauteur quelle que soit leur position sur le champ.

- Rolling Ball :

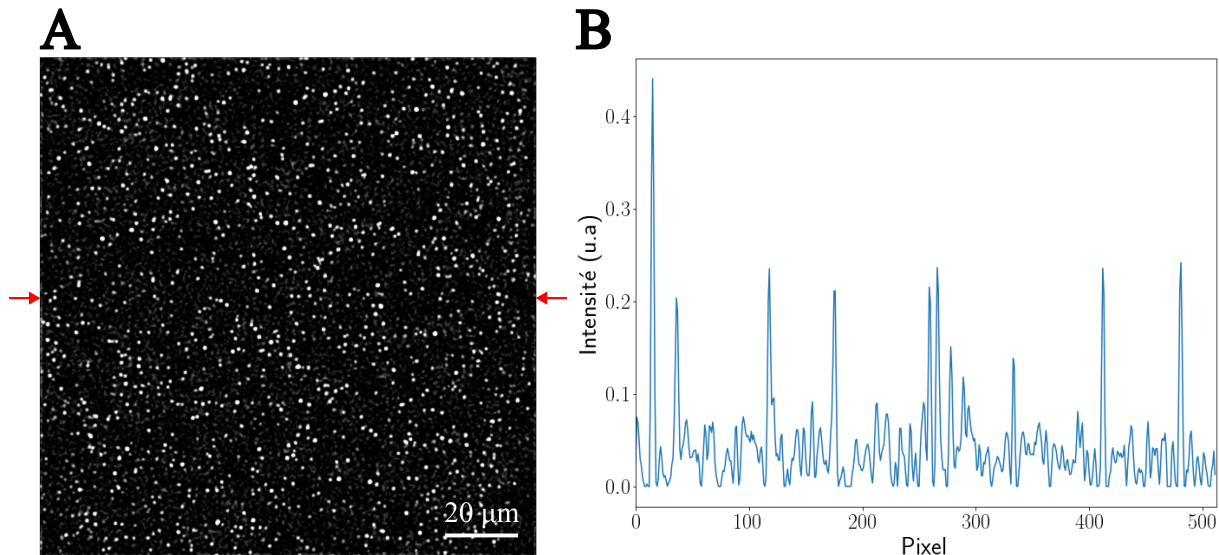
Après renormalisation, l'enveloppe moyenne du signal obtenu est constante sur la totalité du champ de vue mais le fond de fluorescence est toujours présent. Afin de le retirer de l'image normalisée, on utilise un traitement appelé *Tophat* ou *Rolling Ball*. Cette étape consiste à cartographier grossièrement l'enveloppe moyenne du signal via une sphère d'une certaine taille. Cette dernière doit être suffisamment petite pour évaluer le plus correctement possible les variations du fond de fluorescence mais aussi suffisamment grosse pour ne pas repérer les pics d'intensité. En effet, une fois que la sphère est passée sur toute l'image, l'estimation du fond fait par cette dernière est soustraite à l'image normalisée. Si les pics sont pris en compte dans l'estimation, la soustraction va les retirer du champ de vue ce qui nous ferait perdre les signaux d'intérêt. En revanche, si de gros agrégats sont présents sur l'image, une sphère dont la taille est plus petite que ces derniers, mais plus grande que les particules d'intérêt, va permettre de les éliminer.

- Wavelet :

La dernière partie de ce traitement se base sur l'application d'un filtre passe-bas. De cette façon, les faibles variations du signal présentes sur le fond seront dans un premier temps sélectionnées avant d'être retirées par soustraction de l'image obtenue après *Tophat*. Grâce à cette soustraction, les pics d'intérêt sont encore plus facilement détectés par rapport au fond.

À la suite de ces trois opérations de traitement, l'étape de *Rolling Ball* est une nouvelle fois utilisée afin de corriger les valeurs négatives d'intensité apparues avec le traitement en *Wavelet*. Une fois ce protocole appliqué, un champ de vue typique qu'il est possible d'observer est illustré par la Figure 3.47.A. **Cette fois-ci, le fond de fluorescence est bien uniforme sur toute la largeur du champ de vue, comme le montre la**

**Figure 3.47.B.** Il est important de remarquer que les intensités des pics sont à peu près identiques peu importe leur position dans la direction horizontale du champ. À ce stade, les différences restantes en termes de hauteur peuvent s'expliquer soit par la présence d'agrégats à la place de fluorophores uniques soit par les différentes orientations d'émission de ces derniers.



**FIGURE 3.47** – **Résultat du traitement d'images.** A) Correction du bruit de fond via un algorithme Python où les flèches rouges montrent la position de la coupe d'intensité (B) réalisée selon la direction horizontale du champ de vue. Il s'agit d'une image prise 5 minutes après le début d'une acquisition de 10 minutes, au cours de laquelle la photophysique de l'Atto647N a été étudiée. Les conditions d'acquisition sont les suivantes : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

### 3.4.2 Détection des fluorophores et suivi de l'intensité

Une fois le traitement d'images réalisé, le module Python du nom de *trackpy* est utilisé et plus particulièrement la fonction *trackpy.locate* afin de détecter les marqueurs fluorescents. Cette fonction repère les particules d'intérêt en comparant le signal provenant de ces dernières à un ajustement gaussien 2D : le centre de cette gaussienne correspond aux coordonnées (X,Y) du fluorophore. Pour affiner la détection des ARN ou protéines marquées, il existe notamment deux paramètres sur lesquels nous pouvons nous appuyer : le diamètre  $d$  des points à repérer et l'intensité minimale intégrée sur cette zone de taille  $d$ , qui est appelée *minmass*. Généralement, je m'appuie sur le nombre de points détectés par le module *Trackmate* du logiciel *ImageJ* afin de choisir le bon seuillage. Par ailleurs, en comparant les deux types de détection, j'ai remarqué que pour une couverture de quelques centaines de marqueurs fluorescents, *trackpy* repère toujours un plus grand nombre de points que *Trackmate*. En revanche, à partir de plus de 1800-2000 points le comportement s'inverse et *trackpy* pourra difficilement détecter plus de 3000 points.

La fonction *trackpy.locate* m'est surtout utile afin de repérer les ARN simples brins

et les Atto647N-biotinylés qui sont injectés respectivement en début d'expérience d'assemblage ou durant une cellule de photophysique. En ce qui concerne la détection des protéines marquées, je me base sur les positions des ARN afin de suivre l'intensité de fluorescence des Atto647N au cours du temps. En effet, avant mon arrivée en thèse, des tests ont montré que repérer les protéines marquées présentes à chaque image et comparer leurs coordonnées à celles des ARN pouvait prendre plusieurs dizaines d'heures pour une seule journée d'expérience. Afin de rendre l'analyse plus rapide, la méthode que nous avons adoptée est désormais de collecter les positions des ARN marqués et de voir si l'on détecte, sur ces coordonnées, un signal de fluorescence émis par les Atto647N. De cette façon, je m'intéresse directement aux ARN et protéines qui sont colocalisés sans regarder les sous-unités de capsid qui ne se fixent pas sur leur matériel génétique.

A						B					
	x	y	frame	particle	signal		x	y	frame	particle	signal
300	466.098039	102.882353	0	300	0.065893	1205900	466.098039	102.882353	800	300	0.022990
301	42.116667	103.933333	0	301	0.169058	1205901	42.116667	103.933333	800	301	0.052799
302	151.208333	104.270833	0	302	0.201811	1205902	151.208333	104.270833	800	302	0.094666
303	239.349398	103.903614	0	303	0.266648	1205903	239.349398	103.903614	800	303	0.274545
304	396.996337	104.021978	0	304	0.202660	1205904	396.996337	104.021978	800	304	0.214348
305	16.163934	104.836066	0	305	0.015996	1205905	16.163934	104.836066	800	305	0.014123
306	256.034483	104.620690	0	306	0.111789	1205906	256.034483	104.620690	800	306	0.188959
307	347.808118	104.937269	0	307	0.231159	1205907	347.808118	104.937269	800	307	0.103397
308	356.758389	105.087248	0	308	0.523244	1205908	356.758389	105.087248	800	308	0.044508
309	89.600000	106.050000	0	309	0.096759	1205909	89.600000	106.050000	800	309	0.018551
310	322.011561	105.832370	0	310	0.537304	1205910	322.011561	105.832370	800	310	0.057712
311	378.893204	106.349515	0	311	0.193813	1205911	378.893204	106.349515	800	311	0.202016
312	330.838150	106.872832	0	312	0.264874	1205912	330.838150	106.872832	800	312	0.142111
313	389.025000	107.150000	0	313	0.178911	1205913	389.025000	107.150000	800	313	0.015935
314	483.895238	107.038095	0	314	0.346409	1205914	483.895238	107.038095	800	314	0.065423

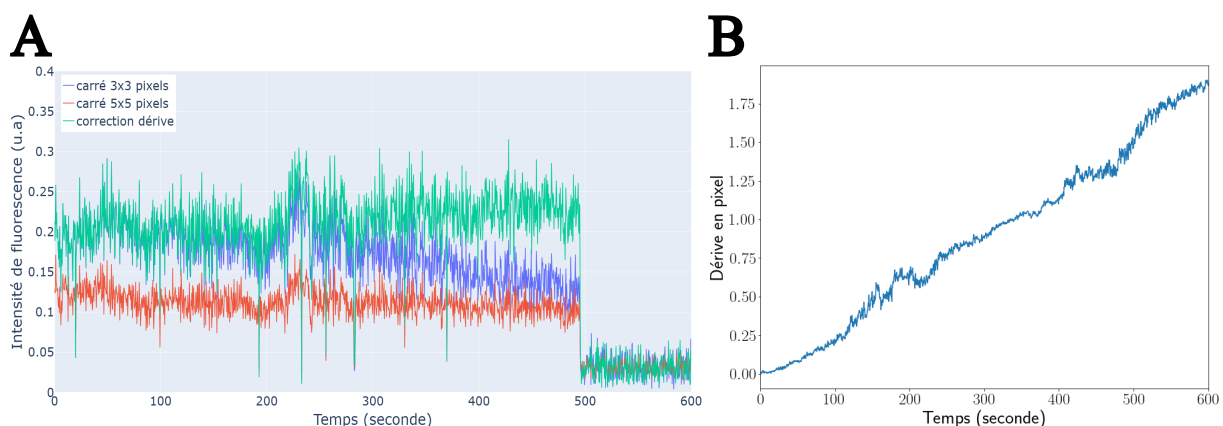
**FIGURE 3.48** – **Tableau Python regroupant les intensités de fluorescence de sites d'assemblage.** Colocalisations enregistrées A) en tout début de film et B) au bout de 5 minutes. Ces données expérimentales proviennent d'une expérience d'assemblage à l'équilibre au cours de laquelle une acquisition de 10 minutes a débuté 15 minutes après l'injection de 50 nM en protéines ayant un taux de marquage de 10 %. Les paramètres d'acquisition sont les suivants : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le PBS de 3,7 mW.

Par ailleurs, la taille de la zone sur laquelle l'intensité de fluorescence est intégrée pour chaque point repéré dans la voie de détection rouge est directement liée à la limite de diffraction d'une molécule unique. En effet, le diamètre  $D$  de la tache de diffraction d'un marqueur unique est donné par la relation suivante :

$$D = \frac{1,22 \times \lambda}{ON} \quad (3.4)$$

avec  $\lambda$  la longueur d'onde d'émission du fluorophore concerné. Dans le cas de l'Atto647N, cette dernière vaut environ 664 nm et avec une ouverture numérique  $ON$  de 1,50 pour l'objectif de microscope *Olympus*, on trouve un diamètre de 540 nm. Or, comme un pixel mesure 267 nm, un marqueur unique ne devrait pas former une tache de diffraction supérieure à un carré de 3x3 pixels. Ainsi, lorsque l'on intègre l'intensité émise par un Atto647N, la zone considérée est un carré de 3x3 pixels centré sur les coordonnées des ARN repérés au préalable. La Figure 3.48 montre un exemple de tableau de colocalisations sur lequel sont présentes les coordonnées des ARN détectés et les intensités mesurées dans la voie rouge en tout début d'acquisition et au bout de 5 minutes.

Pour résumer, le suivi de l'intensité de fluorescence issue des protéines marquées se fait à partir de la détection des ARN marqués et dont les coordonnées sont considérées comme fixes puisqu'ils sont repérés uniquement au début et à la fin d'une expérience d'assemblage. Cette manière de colocaliser ARN et protéines nous permet de gagner du temps au niveau de l'analyse mais nous rend très sensible aux potentielles dérives mécaniques ayant lieu dans le plan de l'échantillon. En effet, en augmentant les durées d'acquisition j'ai commencé à observer une diminution progressive de l'intensité de fluorescence avant que le photoblanchiment ait lieu. Le tracé bleu de la Figure 3.49.A est un exemple de cette observation. Dans ce cas, le signal de fluorescence avait été intégré sur un carré de taille 3x3 pixels. En revanche, lorsque la zone d'intégration passe de 3x3 pixels à 5x5, l'intensité est constante jusqu'à ce qu'elle diminue soudainement au moment du photoblanchiment du fluorophore. Néanmoins, comme l'intensité est moyennée sur une zone plus grande, des pixels caractérisés uniquement par du bruit de fond sont pris en compte, ce qui réduit l'intensité totale.

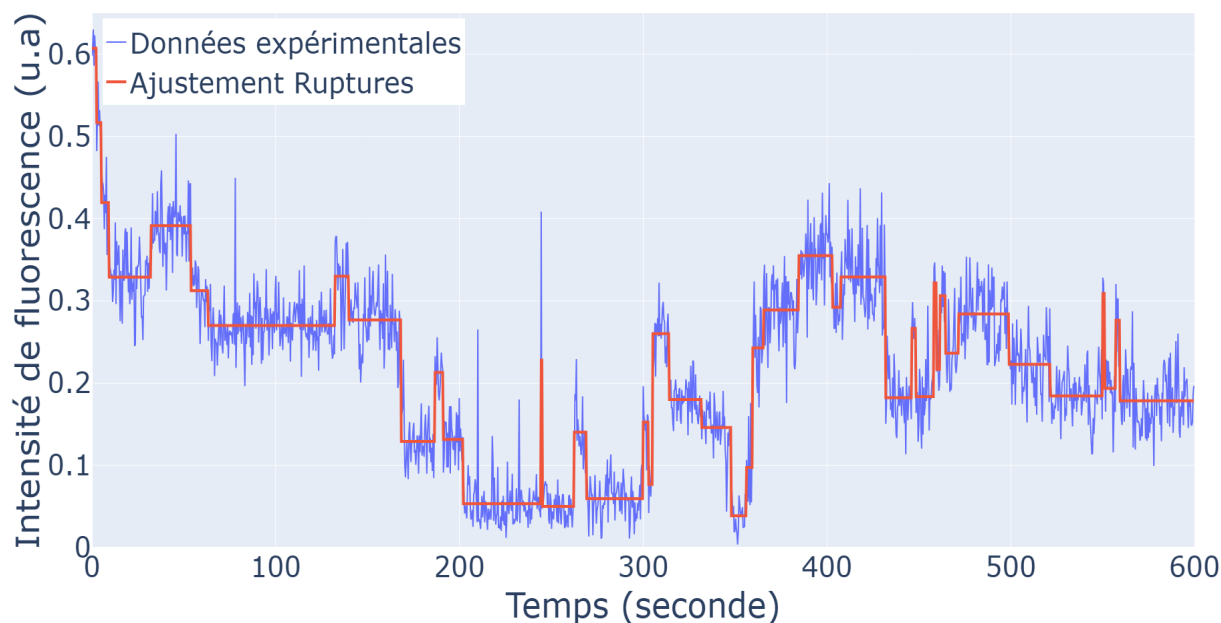


**FIGURE 3.49** – **Caractérisation de la dérive mécanique observée durant une acquisition.**

A) Suivi de l'intensité de fluorescence d'un même fluorophore unique pour trois configurations d'analyse différentes : en bleu, l'intensité est intégrée sur un carré de 3x3 pixels, en rouge le carré mesure 5x5 pixels et en vert la dérive a été corrigée en amont de l'analyse (carré de 3x3 pixels). B) La dérive a été mesurée dans le plan XY de l'échantillon sur un film de 10 minutes à partir d'Atto647N-biotinylés. Les paramètres d'acquisition sont les suivants : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

En réalité, l'agrandissement de la zone d'intégration a permis de compenser ce que nous avons constaté à l'aide du module *Fast4DReg* présent dans *ImageJ*. En effet, la Figure 3.49.B montre le résultat renvoyé par ce module, à savoir un déplacement dans le plan XY de l'échantillon de près de 2 pixels au bout de 10 minutes. Cela signifie qu'au fur et à mesure de l'acquisition, le fluorophore se décale progressivement par rapport à ses coordonnées obtenues à l'image zéro, ce qui conduit à une diminution de l'intensité totale collectée. Fort heureusement, le module *Fast4DReg* nous permet de corriger ce problème, ce qui nous assure de collecter la totalité de l'intensité de fluorescence provenant de chaque coordonnée d'intérêt comme le montre le tracé vert de la Figure 3.49.A. Je précise que cette étape a lieu juste avant le traitement d'images.

**En somme, ces quelques étapes d'analyse effectuées sur *ImageJ* et Python nous permettent d'avoir accès au suivi de l'intensité de fluorescence au cours du temps.** Les Figures 3.35 et 3.38 de la section précédente ont suivi ces différentes étapes de traitement, tout comme la Figure 3.50. Cette dernière, provient d'une expérience d'assemblage à l'équilibre, au cours de laquelle une acquisition de 10 minutes a été enregistrée après avoir laissé incuber les protéines marquées pendant 15 minutes au contact des ARN, dans une cellule micro-fluidique.



**FIGURE 3.50** – Suivi de l'intensité de fluorescence provenant de l'accroche de protéines marquées sur un ARN simple brin. Cette courbe d'évolution de l'intensité de fluorescence provient d'une expérience d'assemblage à l'équilibre au cours de laquelle une acquisition de 10 minutes a débuté 15 minutes après l'injection de 50 nM en protéines dont 10 % sont marquées. En bleu, il s'agit des données expérimentales et en rouge de l'ajustement en marches d'escaliers. Les paramètres d'acquisition sont les suivants : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

### 3.5 Conclusion

Au cours de ce chapitre, j'ai pu expliquer les différentes composantes essentielles à l'imagerie d'un virus en microscopie de fluorescence à l'échelle de la molécule unique. Le microscope TIRF est plus qu'important puisqu'il nous permet d'éclairer les échantillons d'intérêt comme nous le souhaitons mais je pense qu'il est nécessaire d'insister sur notre ultra-dépendance vis-à-vis de la chimie de surface. En effet, comme je l'ai déjà mentionné durant la section 3.2, la façon dont les lamelles de microscope sont fonctionnalisées détermine entièrement une journée d'expérience. Du fait de sa dépendance envers plusieurs paramètres, cette étape est très fragile et certaines expériences ont dû être repoussées lorsque le nombre d'accroches non spécifiques en protéines marquées était trop important. Par ailleurs, même encore en fin de thèse le protocole de chimie de surface a dû être adapté, preuve que notre connaissance de son mode d'action n'est pas totale. Malgré tout, l'architecture de type PLL-g-PEG et PLL-biotine semble être la plus adaptée à notre projet et nous a permis de réaliser un grand nombre d'expériences.

Ces deux composantes, en plus du protocole de traitement d'images que je viens de décrire ont été combinées afin de caractériser le comportement de fluorophores uniques tels que l'Atto647N. Cette étape de caractérisation est importante puisque nous avons pu mieux comprendre la photophysique de ce marqueur fluorescent. Des informations telles que la hauteur de marche ou la durée moyenne avant photoblanchiment ont pu être collectées pour différentes conditions d'illumination afin de voir quelles sont celles qui sont le plus adaptées aux expériences d'assemblage que nous souhaitons mettre en place. En parallèle, l'observation du comportement de l'Atto647N nous a permis de mettre en lumière un aspect de sa photophysique qui peut altérer notre interprétation des données collectées au cours d'une expérience. En effet, le mécanisme de clignotements a pu être observé à plusieurs reprises mais son comportement difficilement prévisible a rendu sa caractérisation complexe. De plus, les solutions envisagées pour réduire ses effets n'ont pas fonctionné comme attendu puisque seule la durée moyenne avant photoblanchiment a pu être améliorée. Néanmoins, le système ROXS ayant prouvé son efficacité dans de nombreuses études, le but serait de se pencher sur ce problème dans un avenir proche.

Tout ce qui est détaillé dans ce chapitre nous amène à l'enregistrement de signaux tels que celui présenté à la Figure 3.50. Sur cette dernière, on peut voir que l'intensité de fluorescence issue de protéines marquées varie assez nettement au niveau d'un site d'ARN. Avant de rentrer plus en détails dans l'interprétation de cette dynamique, je présente dans le prochain chapitre la dernière brique importante du projet de thèse : l'algorithme de détection de marches.



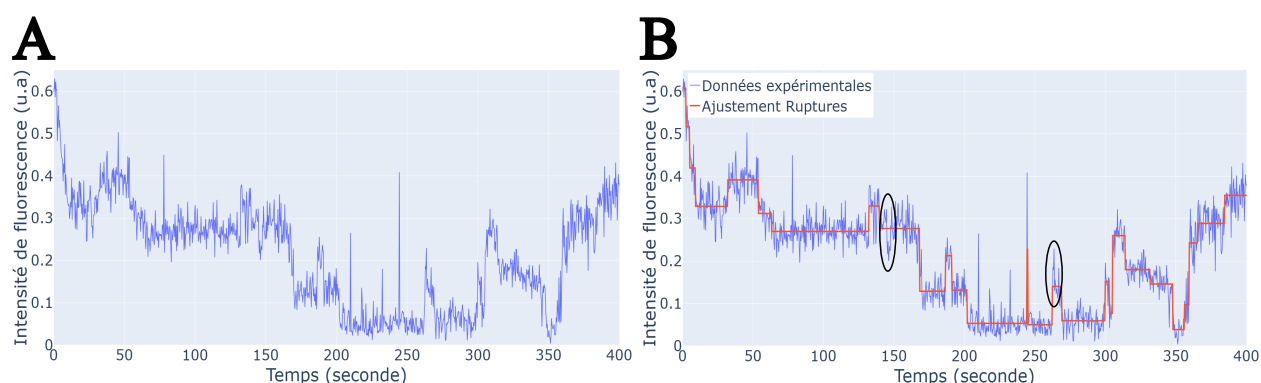
## Chapitre - 4

---

### Détection de marches sur des données obtenues en microscopie de fluorescence

---

Dans cette nouvelle partie, je présente en détails l'algorithme de détection de marches qui nous permet d'analyser plus facilement les données acquises durant une expérience d'assemblage (Figure 4.1). Pour commencer, je montre que l'utilisation d'un ajustement en marches d'escalier est nécessaire puis je présente les deux algorithmes avec lesquels j'ai travaillé. Dans un second temps, j'introduis la façon dont les signaux simulés ont été générés, notamment en mettant en lumière les critères utilisés afin de comparer les ajustements prédits par les algorithmes. Pour finir, je reviens sur les performances de ces derniers quant à leur capacité à bien détecter les variations d'intensité et l'impact que cette estimation a sur le calcul des paramètres d'assemblage que l'on souhaite mesurer en expérience. Cette comparaison a été menée pour des signaux simulés avec et sans clignotements de l'Atto647N.



**FIGURE 4.1** – **Détection de marches dans un signal bruité.** Suivi de la dynamique d'assemblage de sous-unités marquées autour d'un ARN simple brin : A) sans et B) avec un ajustement estimé par l'algorithme Ruptures. Les zones entourées en noir montrent des événements que Ruptures aurait pu (dû ?) repérer.



## 4.1 Les algorithmes de détection

Dans cette première section, j'explique en quoi il est important pour nos expériences de travailler sur un ajustement en marches d'escalier plutôt qu'un signal bruité. J'introduis dans un second temps les deux algorithmes de détection avec lesquels j'ai travaillé et dont j'ai comparé les performances : Ruptures et quickPBSA (*quick Photobleaching Step Analysis*).

### 4.1.1 Nécessité d'une telle méthode pour nos expériences

À la fin du chapitre 2, je mets en avant le fait que les différentes techniques utilisées afin de suivre la cinétique d'assemblage d'une capsid virale autour de son matériel génétique ne permettent pas de travailler à l'échelle de la sous-unité unique. Il n'est donc pas possible d'observer l'arrivée et le départ des dimères de capsid, que ce soit au fur et à mesure de la croissance de cette dernière ou lorsque des réarrangements ont lieu en fin d'assemblage. C'est dans le but de pouvoir détecter ces événements que nous avons choisi de marquer les sous-unités de CCMV avec un fluorophore. En effet, comme décrit durant le chapitre 3, la microscopie TIRF nous permet non seulement de travailler à l'échelle du virus unique mais également de caractériser les propriétés photophysiques de marqueurs fluorescents individuels. Grâce à ces observations, nous sommes en mesure de dire que les variations d'intensité de fluorescence détectées dans la voie d'excitation rouge qui sont illustrées aux Figures 3.1, 3.50 et 4.1 correspondent à des dimères marqués interagissant avec un ARN.

Les signaux bruités présentent plusieurs variations d'intensité qui correspondent, en partie, à des sous-unités fluorescentes qui arrivent ou quittent l'ARN. Cependant, les expériences décrites à la section 3.3.2 ont montré que les fluorophores peuvent photoblanchir et clignoter, créant ainsi des variations d'intensité n'ayant pas une origine biologique. Les dynamiques de ces mécanismes sont assez dispersées (Figures 3.36.B et 3.39), puisque par exemple, pour une même condition expérimentale, la durée de passages en OFF peut aller de moins d'une seconde (2 images) à plusieurs minutes (centaines d'images). De plus, que ces variations soient d'origine photophysique ou biologique, la distribution d'intensité émise par un marqueur fluorescent est elle aussi assez large. En effet, la Figure 3.36.A montre que l'intensité moyenne émise par un fluorophore unique est de 0,118 u.a et son écart-type vaut 0,049 u.a.

Pour continuer dans ce sens, le Tableau 4.1 regroupe plusieurs informations caractérisant l'hétérogénéité des variations d'intensité observées au sein d'un signal typique d'une expérience d'assemblage. Ces résultats ont été obtenus à partir des sites d'ARN détectés parmi deux cellules micro-fluidiques étudiées dans les mêmes conditions. Dans les deux cas, nous pouvons remarquer qu'il y a plus de variations négatives, mais cela est dû aux conditions expérimentales de l'expérience sur lesquelles je reviens dans le chapitre 5. Par

ailleurs, même si les variations positives et négatives d'intensité ont des valeurs moyennes similaires, les valeurs sont assez dispersées comme en témoignent les écarts-types relevés. La même observation peut être faite vis-à-vis des durées de plateaux entre deux variations successives.

Essai	Sites d'ARN	Événements	Nombre	Moyenne	Écart-type
1	1332	Variations positives	5554	0.0876 u.a	0.0650 u.a
		Variations négatives	6796	-0.0897 u.a	0.0582 u.a
		Plateaux	13682	58.4 s	98.4 s
2	718	Variations positives	4576	0.0885 u.a	0.0557 u.a
		Variations négatives	6545	-0.0924 u.a	0.0564 u.a
		Plateaux	11839	36.4 s	59.2 s

**TABLEAU 4.1 – Hétérogénéité des variations d'intensité détectées au cours d'une expérience d'assemblage.** Pour cette expérience, 50 nM en protéines, dont 10 % sont marquées, ont été injectés après avoir repéré les ARN. L'incubation des protéines a duré 15 minutes avant que l'acquisition dans la voie rouge soit lancée : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

Ces hétérogénéités dans la nature des variations d'intensité, leur hauteur et leur dynamique ont été obtenues grâce à l'algorithme Ruptures. Elles montrent surtout qu'il aurait été très difficile de détecter tous ces événements en utilisant des seuils, que ce soit sur l'écart temporel entre deux variations consécutives ou l'intensité émise par un fluorophore. Les ajustements en marches d'escalier, estimés à partir des détections faites par Ruptures, permettent d'être plus précis dans la façon de repérer ces événements. De cette façon, de nouvelles informations concernant la dynamique d'arrivée et de départ des sous-unités peuvent être récoltées. Il est donc important de comprendre comment ces ajustements ont été obtenus puis caractérisés.

#### 4.1.2 Présentation de l'algorithme Ruptures

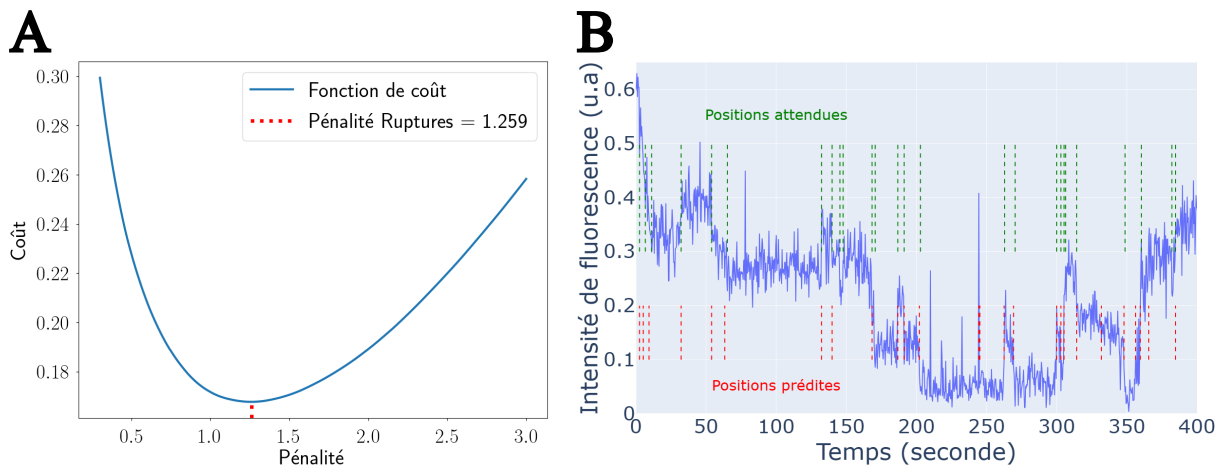
L'algorithme Ruptures m'a permis d'analyser les milliers de signaux collectés durant chaque expérience de photophysique ou d'assemblage, et dans cette section je décris rapidement son fonctionnement. Tout ce que je présente ici est expliqué plus en détails par les auteurs de l'article suivant : Truong et al. (2020).

En quelques mots, l'objectif de Ruptures est de repérer des changements soudains (ou des ruptures) au sein de la valeur moyenne d'un signal temporel. Il est donc tout à fait adapté pour travailler sur les signaux auxquels nous nous intéressons. Cependant, comme nous ne connaissons pas en amont le nombre de ruptures, l'algorithme va appliquer des centaines de valeurs strictement positives appelées **pénalités** afin de détecter au mieux les changements soudains présents dans les signaux. Cette pénalité va orienter Ruptures sur le nombre de variations à repérer :

- Une faible pénalité sera choisie lorsqu'un grand nombre de variations doit être détecté. C'est notamment le cas pour une expérience d'assemblage ou de photophysique visant à caractériser les clignotements de l'Atto647N. Le risque ici est de favoriser la détection de ruptures parmi du bruit.

- Au contraire, une pénalité élevée sera préférée afin de mettre l'accent sur les changements les plus significatifs. C'est par exemple ce que l'on cherche à faire afin de caractériser le photoblanchiment de l'Atto647N.

Afin de trouver la pénalité la plus adaptée, Ruptures se base sur un système d'apprentissage supervisé à partir duquel il va être entraîné. En effet, la première étape d'utilisation de l'algorithme est de lui fournir des signaux ainsi qu'une liste des positions des variations repérées manuellement pour chacun d'entre eux : on parle ici de **données d'entraînement**. Par exemple, que ce soit pour une expérience d'assemblage ou de photophysique, j'utilise généralement 50 signaux choisis aléatoirement parmi tous ceux enregistrés ce jour-là. La seule contrainte ici est que les positions des variations d'intensité doivent être séparées au minimum de deux images pour être détectées par Ruptures.



**FIGURE 4.2 – Phase d'entraînement de Ruptures.** A) Fonction de coût tracée à partir de 50 signaux d'entraînement et 1000 pénalités choisies entre 0,3 et 3. B) Comparaison entre les positions des ruptures attendues (en vert) et prédites par l'algorithme (en rouge). Les conditions d'acquisition des signaux d'entraînement sont les suivantes : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

Durant son entraînement, l'algorithme va utiliser 1000 pénalités réparties sur une plage de valeurs assez larges, généralement entre 0,3 et 3, et ce pour les 50 signaux d'entraînement. Pour chaque couple pénalité - signal, Ruptures prédit la position des variations soudaines et compare sa prédiction à la liste que nous avons créée manuellement. De cette comparaison, un coût est calculé afin d'estimer l'écart entre la prédiction de l'algorithme et la liste attendue. Ainsi, pour chaque pénalité, 50 valeurs de coût sont obtenues puis moyennées afin d'obtenir une fonction de coût comme celle illustrée par

la Figure 4.2.A. **Pour Ruptures, la meilleure pénalité est celle qui minimise la fonction de coût**, à savoir 1,26 ici. Par ailleurs, pour un entraînement typique, le nombre de marches attendu est assez variable selon le signal considéré mais nous nous repérons généralement une trentaine de marches la plupart du temps.

Une fois la pénalité déterminée, nous pouvons l'utiliser sur l'ensemble des signaux afin de connaître les positions des ruptures estimées par l'algorithme. Pour les signaux d'entraînement, il est même possible de comparer les positions prédites (en rouge) à celles attendues (en vert), comme illustré en Figure 4.2.B. Si les estimations ne sont pas satisfaisantes, on peut toujours affiner la recherche de la pénalité en rentrant manuellement des valeurs dans l'algorithme. En effet, la Figure 4.2.A montre qu'autour de son minimum, la fonction de coût ne varie pas énormément, ce qui signifie que la pénalité la plus adaptée peut se situer dans cette zone. Enfin, une fois que les estimations faites par Ruptures nous conviennent, la dernière étape consiste à calculer la valeur moyenne du signal entre deux variations consécutives : **l'ajustement en marches d'escalier est réalisé**.

Pour terminer, outre le nombre de ruptures, d'autres caractéristiques propres aux signaux étudiés vont influencer la valeur de pénalité choisie : la longueur du signal, la hauteur des variations ainsi que le bruit. Il est donc important d'entraîner une nouvelle fois Ruptures si les conditions expérimentales sont modifiées.

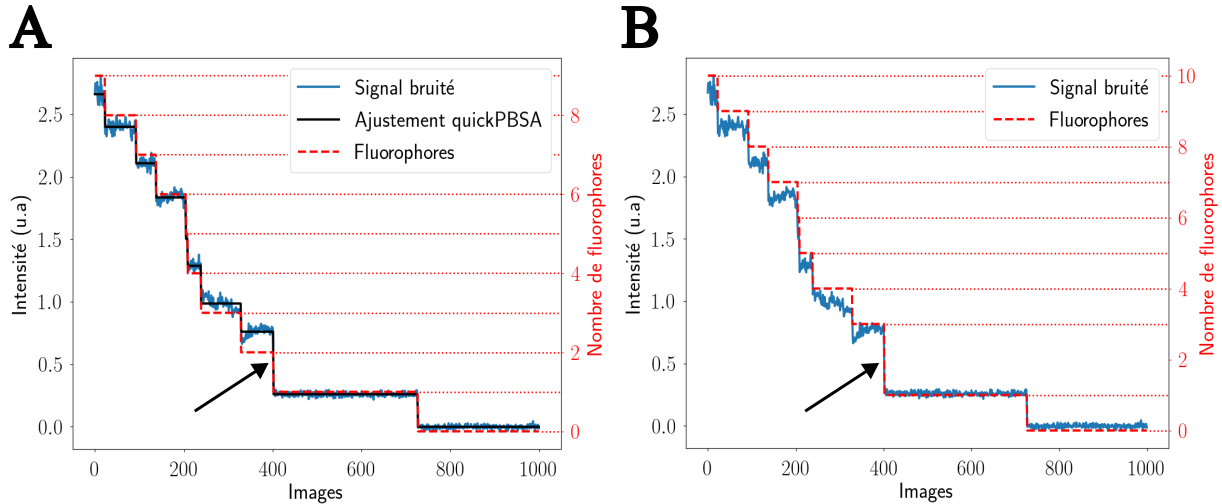
### 4.1.3 Comparaison avec un autre algorithme : quickPBSA

Durant ma thèse, un second algorithme de détection a été utilisé : quickPBSA. Cet algorithme, qui s'intitule *quick Photobleaching Step Analysis*, a été développé dans le but de repérer des variations d'intensité caractéristiques du photoblanchiment. Une explication détaillée du fonctionnement et des performances de quickPBSA est donnée par l'article suivant : Hummert et al. (2021). L'algorithme fonctionne en deux étapes, que je décris en m'appuyant sur la Figure 4.3 :

- **Détection préliminaire** : Ici, un seuil sur la hauteur des variations d'intensité est choisi. Les auteurs conseillent de prendre une valeur comprise entre le tiers et la moitié d'un changement d'intensité caractéristique du signal étudié. Si le seuil est pris trop petit, des événements supplémentaires, non représentatifs du photoblanchiment seront sélectionnés alors que s'il est trop grand, certaines variations d'intensité seront omises. À la suite de cette étape, l'ajustement représenté en noir sur la Figure 4.3.A est obtenu tout comme la correspondance en nombre de fluorophores en rouge.

- **Correction / perfectionnement** : Au moment de la correspondance entre l'ajustement et le nombre de fluorophores, l'algorithme associe tout simplement un événement de photoblanchiment à un marqueur fluorescent. Il ne tient pas compte de la probabilité que deux fluorophores puissent photoblanchir en même temps comme illustré en Figure 4.3.A par la flèche noire. Ici, le nombre de fluorophores est donc sous-estimé d'une unité.

Pour corriger cet oubli, l'algorithme va minimiser le logarithme d'un paramètre mathématique appelé "probabilité *a posteriori*". Ainsi, en fonction de la hauteur de la variation d'intensité repérée, le nombre de fluorophores ayant photoblanchi pourra être pris supérieur à 1. Le résultat de cette correction est représenté par la Figure 4.3.B.



**FIGURE 4.3** – **Exemple de détection par quickPBSA.** Estimation du nombre de fluorophores ayant photoblanchi A) avant et B) après l'étape de correction appliquée par quickPBSA. Les flèches noires montrent une variation double *i.e.* durant laquelle deux fluorophores photoblanchissent en même temps. Les pointillés rouges permettent de compter le nombre de plateaux.

Pour les tests que j'ai effectués avec quickPBSA, seule la première étape a été testée. En effet, l'objectif était de voir si, dans un premier temps, l'algorithme pouvait repérer correctement les différentes variations d'intensité. L'estimation du nombre de fluorophores n'était pas recherchée durant les tests préliminaires.

## 4.2 Introduction à la simulation de données

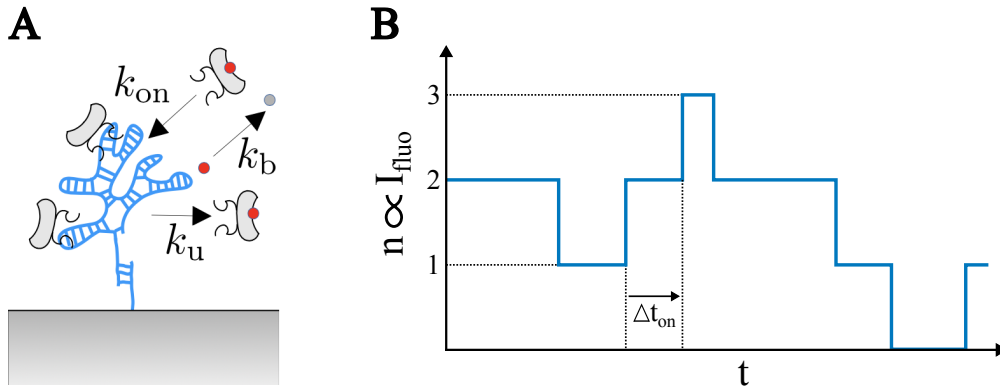
Le but des deux algorithmes présentés plus haut est donc de faire apparaître des marches bien distinctes à partir de la détection de variations d'intensité. Au moment de la prise en main de Ruptures et quickPBSA, je les ai directement utilisés sur des données expérimentales. Cependant, il est rapidement apparu que les ajustements en marches d'escalier obtenus n'étaient pas totalement représentatifs des signaux bruités. En d'autres termes, les algorithmes pouvaient rajouter des marches inexistantes ou en omettre certaines, malgré des estimations globales très satisfaisantes. Il était donc important de se poser la question suivante : **À quel point pouvons-nous faire confiance aux ajustements en marches d'escalier ?** Pour répondre à cette interrogation, nous avons simulé nos propres signaux, puisque contrairement aux données expérimentales, les vraies positions des marches sont connues.

Ainsi, dans cette section je commence par détailler le modèle mathématique qui a

été utilisé afin de simuler des signaux typiques de l'assemblage de capsides à l'équilibre. Ensuite, je décris comment les signaux ont été simulés sur Python afin de respecter le modèle mathématique choisi. Pour finir, je présente les métriques d'analyse utilisées afin d'évaluer les performances des deux algorithmes. **Message important pour le lecteur.** Jusqu'à présent je ne faisais pas de distinction entre protéines, sous-unités ou dimères marqués. À partir de maintenant et pour les chapitres suivants, un événement d'accroche, *i.e.* une marche montante, correspond à l'arrivée d'un dimère marqué, aussi appelé sous-unité marquée.

#### 4.2.1 Modélisation de la dynamique d'assemblage à l'équilibre

La majorité des expériences d'assemblage réalisées durant ma thèse l'ont été à l'équilibre. Cela signifie que les protéines ont été incubées pendant 15 minutes dans la cellule micro-fluidique, au contact des ARN simples brins, avant que l'acquisition ne commence. Des sous-unités marquées et non marquées peuvent donc déjà être présentes sur l'ARN dès le début du film, comme représenté en Figure 4.4.A. Nous nous attendons à observer des capsides partiellement formées dont la taille n'évolue plus dans le temps, mais avec des dimères pouvant partir de cette dernière ou bien s'y fixer. En raisonnant en termes de sous-unités marquées et donc d'intensité de fluorescence, les échanges attendus au niveau d'un ARN simple brin peuvent être matérialisés par le signal illustré à la Figure 4.4.B.



**FIGURE 4.4** – Illustration de la dynamique d'assemblage d'une capside à l'équilibre. A) La capside partiellement formée est un lieu d'échanges de sous-unités où, pour les dimères marqués (avec un rond rouge) :  $k_{on}$  est la constante d'association,  $k_u$  est la constante de dissociation et  $k_b$  correspond à la constante de photoblanchiment. B) Évolution de l'intensité de fluorescence  $I_{fluo}$  détectée dans la voie rouge en fonction des échanges de dimères marqués où :  $n$  correspond au nombre de sous-unités fluorescentes présentes sur l'ARN et  $\Delta t_{on}$  est la durée entre deux marches montantes consécutives.

Sur ce signal, nous pouvons voir des marches descendantes qui peuvent correspondre à un départ d'un dimère marqué mais aussi au photoblanchiment ou au début de clignotement d'un Atto647N dont la sous-unité est toujours sur l'ARN. De la même façon, une marche montante va caractériser l'accroche d'un dimère marqué ou la fin d'un clignotement d'un marqueur fluorescent qui a pu quitter l'état triplet  $T_1$ . **Dans notre modèle,**

nous n'allons pas inclure les clignotements du fluorophore ce qui signifie qu'ils ne participent pas aux événements montants et descendants. Leur influence est discutée en fin de chapitre. Le Tableau 4.2 regroupe les notations des différents paramètres utilisés dans le modèle.

Notation	Paramètre physique correspondant
$n$	Nombre de sous-unités marquées sur l'ARN
$\Delta t_{on}$	Durée entre deux événements montants consécutifs ( $s$ )
$k_{on}$	Constante d'association ( $s^{-1}$ )
$k_{off}$	Constante de temps des marches descendantes ( $s^{-1}$ )
$k_u$	Constante de dissociation ( $s^{-1}$ )
$k_b$	Constante de temps de photoblanchiment ( $s^{-1}$ )

**TABLEAU 4.2** – Paramètres impliqués dans le modèle de dynamique d'assemblage à l'équilibre.

Parmi les paramètres présents dans le tableau ci-dessus, deux d'entre eux peuvent être mesurés directement à partir de signaux expérimentaux :  $\Delta t_{on}$  et  $n$ . En faisant l'hypothèse que nous travaillons à l'équilibre, la représentation théorique de ces deux grandeurs peut être décrite comme suit.

On suppose qu'à un instant  $t$  (par exemple au début du film de protéines), il y a  $n$  sous-unités fluorescentes sur un site d'ARN. Au cours de l'intervalle de temps  $\Delta t$  suivant, trois événements peuvent avoir lieu :

- Une nouvelle sous-unité fluorescente arrive (marche montante) sur l'ARN avec une constante d'association  $k_{on}$  et donc une probabilité  $k_{on}\Delta t$ , en considérant un intervalle de temps suffisamment petit.

- Une sous-unité déjà accrochée disparaît (marche descendante), que ce soit en se décrochant ou par mécanisme photophysique (photoblanchiment), avec une constante de temps  $k_{off}$ . Or, comme l'ARN possède  $n$  dimères marqués, la probabilité qu'une marche descendante se produise vaut  $nk_{off}\Delta t$ .

- Il ne se passe rien et la probabilité correspondante est la complémentaire des deux événements précédents, soit :  $1 - k_{on}\Delta t - nk_{off}\Delta t$ .

Ainsi, la probabilité d'avoir  $n$  sous-unités fluorescentes à l'instant  $t + \Delta t$  s'exprime selon la relation suivante :

$$p(n, t + \Delta t) = k_{on}\Delta t p(n - 1, t) + (n + 1)k_{off}\Delta t p(n + 1, t) + (1 - k_{on}\Delta t - nk_{off}\Delta t)p(n, t)$$

**Or, comme nous avons fait l'hypothèse de travailler à l'équilibre, les probabilités demeurent constantes au cours du temps.** L'égalité  $p(n, t + \Delta t) = p(n, t) = p(n)$  est donc vérifiée et la relation s'écrit désormais :

$$k_{on}p(n-1) + (n+1)k_{off}p(n+1) - (k_{on} + nk_{off})p(n) = 0$$

On décide de diviser l'égalité précédente par le paramètre  $k_{off}$  :

$$\frac{k_{on}}{k_{off}}p(n-1) + (n+1)p(n+1) - \left(\frac{k_{on}}{k_{off}} + n\right)p(n) = 0$$

En réarrangeant l'expression précédente et en posant  $\mu = \frac{k_{on}}{k_{off}}$ , on obtient :

$$(n+1)p(n+1) - \mu p(n) = np(n) - \mu p(n-1)$$

Lorsque l'on se place dans le cas où il n'y a aucune sous-unité marquée présente sur l'ARN, comme la probabilité  $p(-1)$  est nulle, il vient que  $p(1) = \mu p(0)$ . **Par récurrence on obtient l'égalité  $np(n) = \mu p(n-1)$  qui est valable pour tout  $n$ , ce qui nous permet d'écrire :**

$$p(n) = \frac{\mu}{n}p(n-1) = \frac{\mu^2}{n(n-1)}p(n-2) = \dots = \frac{\mu^n}{n!}p(0)$$

Enfin, comme la condition de normalisation impose  $\sum_n p(n) = 1$ , cela conduit à  $p(0) = e^{-\mu}$  et donc :

$$\boxed{p(n) = \frac{\mu^n}{n!}e^{-\mu}} \quad (4.1)$$

**La distribution des valeurs de  $n$  doit donc suivre une loi de Poisson dont la moyenne vaut :  $\langle n \rangle = \mu = \frac{k_{on}}{k_{off}}$ .** De plus, la constante de temps des marches

descendantes s'exprime selon la relation suivante :  $k_{off} = k_u + k_b$  avec  $k_u = \frac{1}{\tau_u}$  et  $k_b = \frac{1}{\tau_b}$  où  $\tau_u$  correspond au temps de résidence moyen des protéines et  $\tau_b$  est la durée moyenne avant photoblanchiment.

En ce qui concerne la dynamique des événements d'accroche  $\Delta t_{on}$ , nous pouvons commencer par supposer qu'une sous-unité fluorescente vient de se lier à un site d'ARN. À partir de cet événement, on découpe le temps en une succession d'intervalles temporels notés  $\Delta t$ . De cette façon, la probabilité qu'une nouvelle sous-unité s'accroche durant le  $(i+1)^{eme}$  intervalle s'écrit selon le produit des probabilités qu'une sous-unité ne s'accroche pas sur  $i$  intervalles et de la probabilité qu'un dimère finisse par se lier entre  $t = i\Delta t$  et  $t + \Delta t$ . On obtient l'expression suivante :

$$p(t)\Delta t = (1 - k_{on}\Delta t)\dots(1 - k_{on}\Delta t)k_{on}\Delta t = \left(1 - \frac{k_{on}t}{i}\right)^i k_{on}\Delta t \approx k_{on}e^{-k_{on}t}\Delta t$$



pour laquelle on utilise le fait que :  $e^{-x} = \lim_{i \rightarrow \infty} \left(1 - \frac{x}{i}\right)^i$ . Ainsi l'expression finale est :

$$\boxed{p(\Delta t_{on}) = k_{on} e^{-k_{on} \Delta t_{on}}} \quad (4.2)$$

**La distribution des événements d'accroche  $\Delta t_{on}$  doit donc suivre une loi exponentielle dont la moyenne vaut :  $\langle \Delta t_{on} \rangle = \frac{1}{k_{on}}$ .**

#### 4.2.2 Génération des jeux de données simulées

La simulation de signaux caractéristiques d'une dynamique d'assemblage de capsides à l'équilibre se base donc sur les deux lois démontrées dans la section précédente. Afin de générer des jeux de données simulées, trois fonctions Python ont été créées et sont brièvement détaillées ci-dessous :

##### - $I_{Ground-truth}$ :

Le but de cette fonction est de générer au fur et à mesure les signaux en marches d'escalier vrais, non bruités qui correspondent à ce que l'on appelle la *Ground-truth* (vérité fondamentale). Parmi les différentes variables d'entrée, il est important d'en retenir cinq :  $\langle n \rangle$ ,  $k_{on}$ ,  $Nb_{frames}$ ,  $h$  et  $\sigma_h$ . Les trois derniers paramètres correspondent respectivement au nombre d'images contenues dans les signaux, à la valeur moyenne de l'intensité émise par un fluorophore et à son écart-type.

Au début de la fonction, le nombre initial de sous-unités marquées  $n_0$  présentes sur l'ARN est tiré au sort parmi une distribution de Poisson de moyenne  $\langle n \rangle$ . Pour chaque sous-unité présente à la première image, une valeur d'intensité va être choisie parmi une distribution gaussienne centrée sur  $h$  et d'écart-type  $\sigma_h$  afin de connaître l'intensité initiale  $I_0$ . Ces intensités d'émission sont stockées dans un tableau. Par la suite, la position temporelle de la première marche est déterminée en tirant au sort cette valeur parmi une distribution exponentielle de constante de temps  $\tau = \frac{1}{k_{on} + nk_{off}}$ . Ici, la distribution exponentielle tient compte de la probabilité qu'une nouvelle sous-unité arrive sur l'ARN et de celle d'en voir une quitter la capsid parmi les  $n$  déjà présentes. Pour la génération de données, on raisonne en nombre d'images et on fait en sorte que l'écart entre deux marches consécutives soit supérieur ou égal à deux images pour que Ruptures puisse les détecter.

Une fois la position de la première marche connue, l'intensité est donc constante de la première image jusqu'à celle précédent la première rupture. À ce moment, la nature de la marche est choisie dans la fonction *Marche*. En sortie de cette dernière, le nombre de sous-unités marquées  $n$  est mis à jour, tout comme le tableau contenant les valeurs d'intensité d'émission de chaque fluorophore. Enfin, la nouvelle constante de temps  $\tau$  est calculée et nous pouvons désormais déterminer la position de la prochaine marche. Ce

processus est répété jusqu'à ce qu'on atteigne la dernière image du signal.

- Marche :

L'objectif ici est de déterminer la nature de la marche à venir. Pour ce faire, on utilise notamment le paramètre de seuillage suivant :  $step_{th} = \frac{k_{on}}{k_{on} + nk_{off}}$ . Ce seuil est comparé en début de fonction à un nombre réel compris entre 0 et 1 qui est tiré au sort aléatoirement par la fonction *random* du module Python qui porte le même nom. Si ce réel est strictement inférieur à  $step_{th}$ , la marche sera montante, sinon elle sera descendante. Il faut donc comprendre que plus  $n$  est important, plus le seuil favorise l'apparition d'une marche descendante. En effet, pour  $n = 0$  on a  $step_{th} = 1$ , ce qui signifie que la marche sera forcément montante. Au contraire, plus  $n$  est grand, plus  $step_{th}$  tend vers 0 ce qui correspond à une probabilité d'obtenir une marche descendante de plus en plus importante.

Selon le résultat de la comparaison, plusieurs paramètres seront modifiés. Si la marche est montante,  $n$  augmente d'une sous-unité, une nouvelle valeur d'intensité est tirée au sort avant d'être comptabilisée dans le signal en cours de création. En revanche, si la marche est descendante, la capsidie simulée perd une sous-unité marquée et une intensité d'émission est retirée du tableau correspondant et du signal qui est en train d'être généré. Dans les deux cas,  $step_{th}$  est modifié.

- Inoisy :

Cette fonction ajoute un bruit gaussien aux signaux en marches d'escalier obtenus en sortie de  $I_{Ground-truth}$ . En effet, à chaque image du signal de référence, une certaine valeur est ajoutée ou retranchée à l'intensité correspondante. La valeur en question est choisie parmi une distribution gaussienne centrée en 0 et d'écart-type  $\sigma_{noise}$  qui est calculé à partir de la relation suivante :

$$\sigma_{noise} = \sqrt{\sigma_{fond}^2 + (\sqrt{n}\sigma_{fluo})^2}$$

où  $\sigma_{fond}$  décrit l'écart-type du bruit de fond, sans fluorophore, et  $\sigma_{fluo}$  correspond au bruit de fluorescence. Ce dernier dépend directement du nombre de fluorophores considérés. Ainsi,  $\sigma_{noise}$  est calculé à chaque image et pour tous les signaux générés dans un jeu de données.

Ces trois fonctions permettent de générer chacun des 1600 signaux composant une série de données. Ils respectent donc tous les conditions fixées en début de simulation par les variables  $n$ ,  $k_{on}$ ,  $h$ ,  $\sigma_{noise}$ . Ensuite, chaque jeu de données est utilisé afin de comparer les performances de quickPBSA et Ruptures. Pour ce dernier, les cinquante premiers signaux simulés vont former les données d'entraînement afin de trouver la pénalité la plus

adaptée. Avant de passer aux résultats, je présente dans la section suivante les critères qui me permettent de comparer les ajustements réalisés par Ruptures et quickPBSA.

### 4.2.3 Présentation des métriques utilisées en *Machine Learning*

Une fois que les algorithmes Ruptures et quickPBSA ont estimé les positions des marches pour chaque signal, l'évaluation de leurs performances peut débuter. Afin d'évaluer la précision avec laquelle les variations soudaines d'intensité sont repérées, j'ai pu m'appuyer sur trois métriques d'analyse utilisées en *Machine Learning* :

- Le *Recall* ou Rappel calcule le pourcentage des marches existantes qui ont bien été repérées par l'algorithme.

- La *Precision* ou Précision renvoie le pourcentage de marches détectées par l'algorithme qui existent.

- Le  $F_{score}$  fait la moyenne harmonique entre le Rappel et la Précision.

Pour calculer ces métriques, il faut comparer l'estimation faite par chaque algorithme avec la réalité. Selon qu'une marche est correctement repérée, manquante ou rajoutée, l'événement correspondant est comptabilisé dans l'une des quatre catégories décrites ci-dessous :

- **Les vrais positifs (TP)** : il s'agit des marches correctement détectées.

- **Les faux positifs (FP)** : ce sont les marches repérées qui n'existent pas.

- **Les faux négatifs (FN)** : ils correspondent aux marches existantes non repérées.

- **Les vrais négatifs (TN)** : ce sont les autres événements où, à raison, l'algorithme n'a placé aucune marche.

Ces différents événements sont donc comptabilisés pour chaque signal étudié en comparant la liste des positions réelles des marches connues via la *Ground-truth* et la liste prédite par l'un des deux algorithmes. Ensuite, les valeurs de TP, FP et FN vont être sommées sur l'ensemble des signaux d'une même série de données avant d'être utilisées afin de calculer les trois métriques présentées plus haut, dont les expressions sont données ci-dessous :

$$Recall = \frac{TP}{TP + FN} \quad Precision = \frac{TP}{TP + FP} \quad F_{score} = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

Pour Ruptures, ces métriques seront calculées sur trois sous-jeux de données compris dans la même série de 1600 signaux. On distingue les 50 premiers signaux qui correspondent aux **données d'entraînement**, les 50 suivants qui font partie des **données de validation** et les 1500 derniers qui forment les **données test**. La série de validation a pour but de travailler sur un petit jeu de données afin de montrer que les résultats obtenus avec les données d'entraînement sont retrouvés pour des signaux inconnus. Enfin, que ce

soit pour Ruptures et quickPBSA, les performances de détection des marches sont données en autorisant une erreur de plus ou moins une image sur les positions des ruptures.

### 4.3 Analyse des performances des algorithmes de détection ...

Pour mener à bien la comparaison entre Ruptures et quickPBSA, trois jeux de données contenant 1600 signaux chacun ont été simulés à partir de différents paramètres clefs dont les valeurs sont reportées dans le Tableau 4.3. Ces dernières proviennent toutes de mesures expérimentales mise à part la valeur moyenne du bruit de fond  $I_{fond}$  qui a été volontairement augmentée afin que les signaux bruités n'aient pas de valeurs négatives.

Série	$Nb_{frames}$	$t_{ech}$ (s)	$\langle n \rangle$	$k_{on}$ ( $s^{-1}$ )	$h$	$\sigma_h$	$I_{fond}$	$\sigma_{fond}$	$\sigma_{fluo}$
D1									0,005
D2	1600	0,375	3	0,015	0,100	0,045	0,150	0,013	0,025
D3									0,045

**TABLEAU 4.3 – Paramètres de simulation de signaux mimant la dynamique d’assemblage de sous-unités marquées.**  $Nb_{frames}$  indique le nombre d’images par signal,  $t_{ech}$  la durée entre deux images consécutives,  $\langle n \rangle$  le nombre moyen de sous-unités marquées,  $k_{on}$  la constante d’association des sous-unités marquées,  $h$  la hauteur moyenne d’une marche,  $\sigma_h$  l’écart-type de cette dernière,  $I_{fond}$  la valeur moyenne du bruit de fond,  $\sigma_{bkg}$  et  $\sigma_{fond}$  les écart-types des bruits de fond et de fluorescence. De  $h$  à  $\sigma_{fond}$ , les valeurs sont données en unités arbitraires (u.a).

La seule différence réside dans le niveau de bruit de fluorescence qui augmente progressivement lorsque l’on passe des jeux de données D1 à D3. La valeur choisie pour la série D2 est celle qui a été mesurée expérimentalement. Ainsi, les signaux obtenus sont plus ou moins bruités (Figure 4.5) et les variations d’intensité suivent la modélisation de la dynamique d’assemblage et la méthodologie de génération des données décrites respectivement aux sections 4.2.1 et 4.2.2. Cela signifie que dans un premier temps, la photophysique des fluorophores simulés est simplifiée puisque seule le photoblanchiment est pris en compte. Grâce aux résultats obtenus dans ces conditions, il est possible d’avoir un premier avis sur les performances de chaque algorithme. Dans un second temps, j’utilise les trois mêmes jeux de données mais pour lesquels des clignotements ont été rajoutés. **J’ai considéré que 20 % des fluorophores clignotent au moins une fois et que la durée moyenne et le nombre moyen de passages en OFF valent respectivement 25 secondes et 2,2.** La fonction utilisée pour rajouter les clignotements est détaillée en annexe F.

Systématiquement, Ruptures a été entraîné sur chaque jeu de données, avec et sans clignotements. Les pénalités associées sont données dans les sections suivantes. Pour quickPBSA, les prédictions ont été faites à partir d’un seuil pouvant aller de 0,01 à 0,03 u.a selon le jeu de données. Les valeurs précises sont données par le Tableau F.2 de l’annexe F. Enfin, la Figure 4.5 montre des exemples d’ajustements en marches d’escalier issus de Ruptures et quickPBSA.

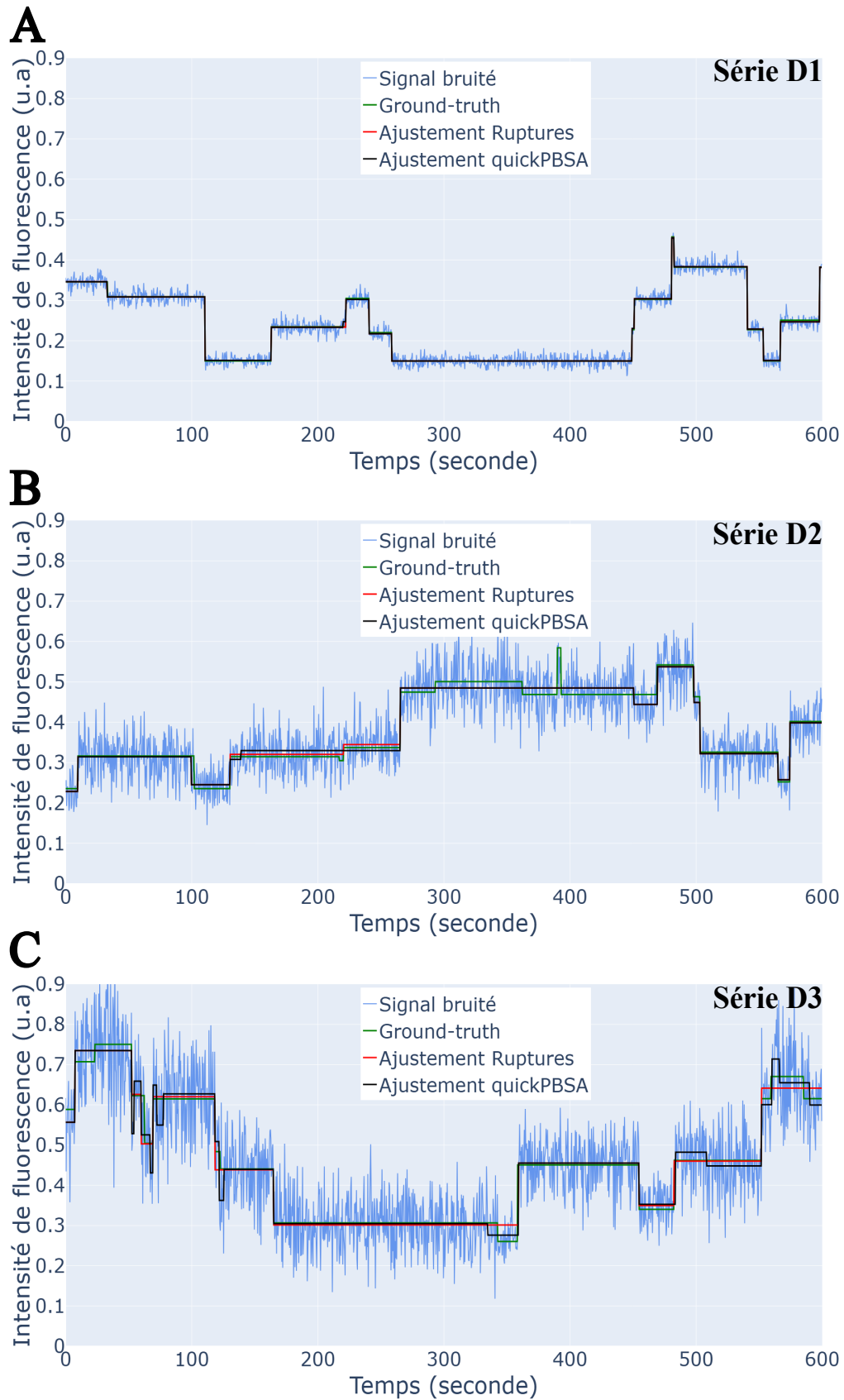
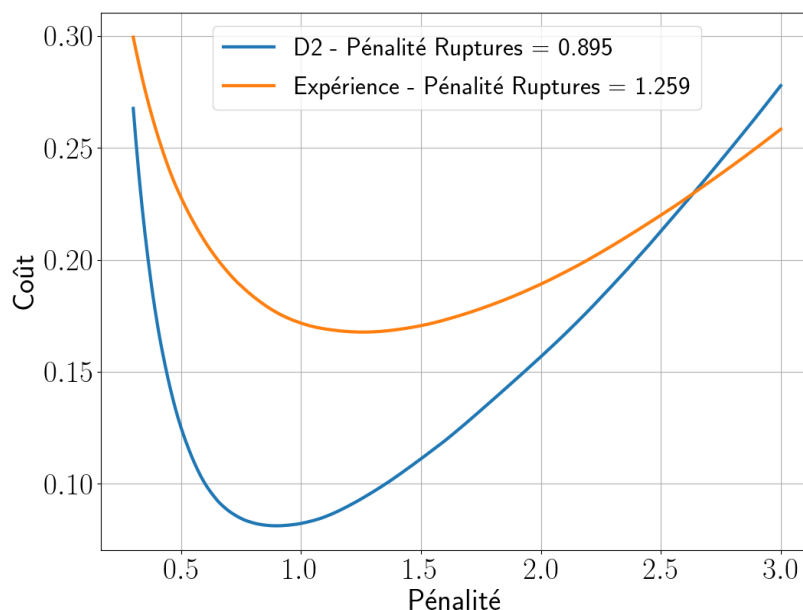


FIGURE 4.5 – Illustration des trois séries de données simulées. A-B-C) Signaux issus respectivement des séries D1, D2 et D3. Les données bruitées sont en bleu, la *Ground-truth* en vert, les prédictions faites par Ruptures en rouge et celles de quickPBSA en noir.

### 4.3.1 ... pour des caractéristiques photophysiques simples

#### 4.3.1.1 Ruptures : choix de la pénalité adaptée

L'algorithme de détection Ruptures a donc été entraîné sur les 50 premiers signaux de chaque jeu de données simulées. À partir des prédictions obtenues pour les différentes pénalités testées, une fonction de coût a pu être tracée comme celle représentée en bleu sur la Figure 4.6 pour la série D2. Dans ce cas, le coût minimal est légèrement inférieur à 0,10 et la pénalité correspondante est de 0,895. Ce résultat est directement comparé à la fonction de coût issue d'un entraînement effectué avec des données acquises en expérience. Il est important de préciser que les conditions expérimentales sont très proches des paramètres utilisés pour générer la série D2. En comparant les deux courbes, nous pouvons voir que la pénalité expérimentale est plus élevée mais aussi que la fonction de coût associée semble moins piquée que celle de la simulation. En théorie, la marge de manoeuvre dans le choix de la pénalité est donc plus faible avec les données simulées puisque l'écart entre les positions réelles des marches et les prédictions de l'algorithme augmente plus rapidement dès que l'on s'éloigne du minimum de la fonction.



**FIGURE 4.6** – Comparaison d'une fonction de coût sur des données expérimentales et simulées. En bleu, la fonction de coût obtenue après avoir entraîné Ruptures sur les 50 premiers signaux simulés du jeu de données D2. En orange, la fonction de coût obtenue après avoir entraîné Ruptures sur 50 signaux choisis aléatoirement parmi les 819 collectés au cours d'une expérience d'assemblage réalisée à l'équilibre. Les paramètres d'acquisition sont les suivants : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

La différence entre les deux fonctions de coût peut s'expliquer par une mauvaise estimation d'un paramètre de simulation mais aussi par la difficulté à représenter du mieux possible l'hétérogénéité des données expérimentales. Cependant, parmi les trois jeux de

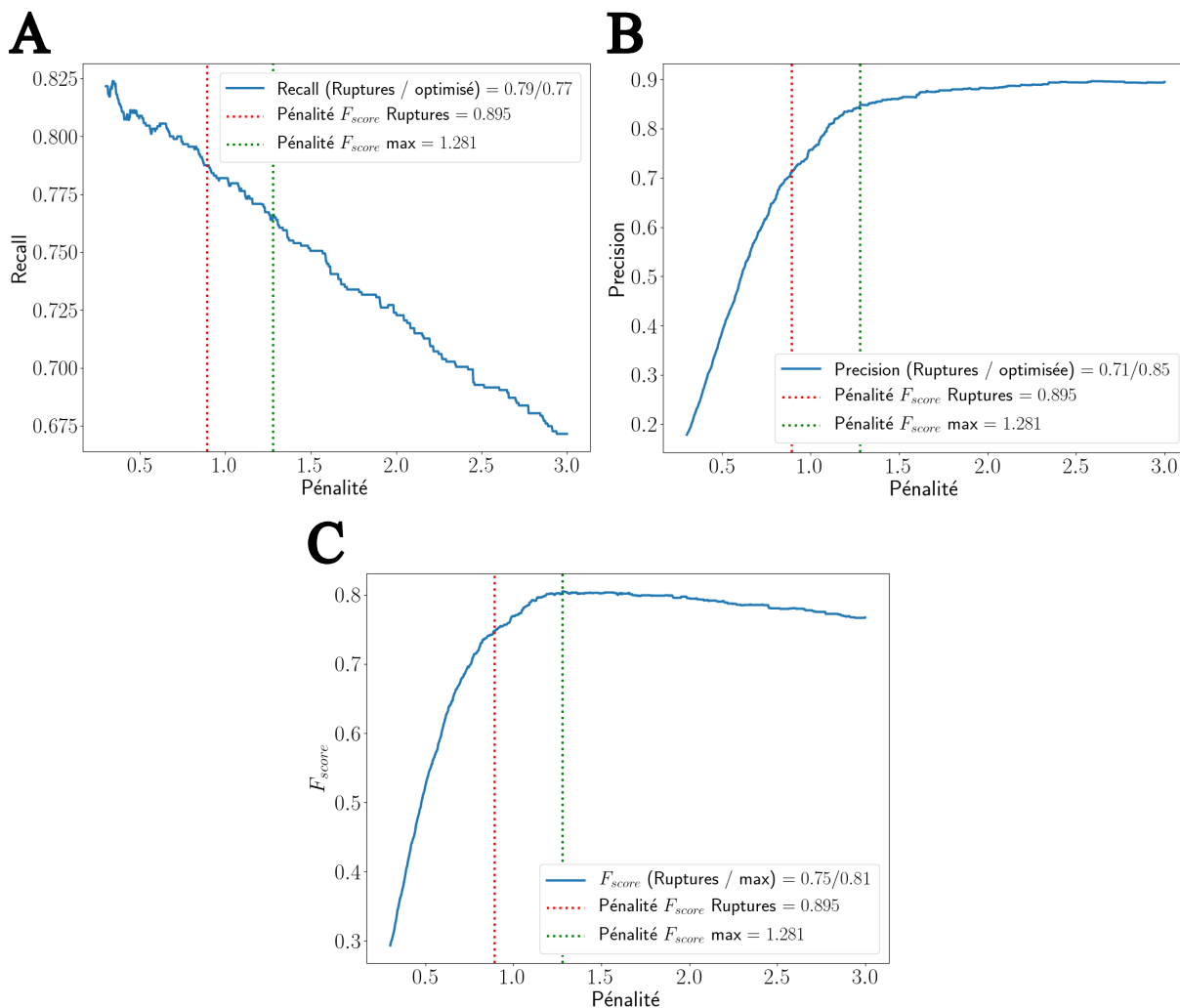
données simulées, la série D2 est celle qui se rapproche le mieux de l'expérience. En effet, en terme de coût minimal la série D1 se situe aux alentours de 0,002 contre 0,50 pour la série D3. De plus, le Tableau 4.4 montre que les pénalités estimées par Ruptures, pour minimiser le coût, sont encore plus faibles que celle trouvée pour D2. En comparant les trois valeurs, nous pouvons remarquer qu'il ne semble pas y avoir d'évolution cohérente de la pénalité optimale avec l'augmentation du bruit de fluorescence. Cependant, les pénalités sont très proches les unes des autres et comme le montre la fonction de coût de la Figure 4.6, elles appartiennent toutes à la zone où l'écart entre réalité et prédiction évolue très peu.

	D1	D2	D3
Pénalité Ruptures	0,819	0,895	0,735
Pénalité $F_{score}$	1,22	1,28	1,70

**TABLEAU 4.4 – Données d'entraînement : recherche de la meilleure pénalité.** La "Pénalité Ruptures" correspond à la valeur qui permet de réduire la fonction de coût d'après l'algorithme alors que la "Pénalité  $F_{score}$ " est celle qui maximise le  $F_{score}$  pour chaque jeu de données.

En parallèle, la dernière ligne du Tableau 4.4 montre une autre valeur de pénalité pour chaque jeu de données. Ici, pour chaque pénalité testée, les paramètres TP, FP et FN ont été sommés sur l'ensemble des signaux d'entraînement afin de calculer le *Recall*, la *Precision* et d'en déduire le  $F_{score}$ . La Figure 4.7 montre l'évolution des trois métriques en fonction de la pénalité pour la série D2.

Sur la Figure 4.7.C, nous pouvons voir que le  $F_{score}$  est maximal pour une pénalité supérieure à celle qui minimise le coût. Nous retrouvons la même tendance pour les jeux de données D1 et D3 comme le montre la dernière ligne du Tableau 4.4. Cette observation peut surprendre sachant que le but de Ruptures est de repérer du mieux possible les variations d'intensité présentes dans un signal. Cependant, l'explication de cette différence trouve son origine dans la façon dont Ruptures fonctionne. En effet, en cherchant à minimiser la fonction de coût, l'algorithme fait en sorte d'avoir l'écart le plus faible entre réalité et fonction d'ajustement. En choisissant une pénalité plus faible, Ruptures va diminuer son seuil de détection afin de repérer plus de marches. De cette façon, l'algorithme va manquer moins de marches réelles mais potentiellement en observer plus qui n'existent pas. C'est exactement ce que l'on retrouve lorsque l'on s'intéresse à l'évolution des deux autres métriques utiles au calcul du  $F_{score}$ . Quand la pénalité diminue, le *Recall* croît alors que la *Precision* décroît. En d'autres termes, pour une pénalité décroissante, le nombre de faux négatifs diminue alors que les faux positifs sont de plus en plus nombreux. Le fait que la *Precision* varie plus rapidement que le *Recall* sur la même plage de pénalité explique l'écart entre le  $F_{score}$  maximal, qui vaut 80,5 % ici, et celui obtenu avec la pénalité donnée par Ruptures, à savoir 74,8 %.



**FIGURE 4.7** – Représentation des métriques d’analyse pour le jeu de données D2. A) *Recall*. B) *Precision*. C)  $F_{score}$ . Pour le calcul du *Recall* et de la *Precision*, les paramètres TP, FP et FN ont été sommés sur l’ensemble des signaux d’entraînement et ce pour chaque pénalité. En pointillés rouges : valeur de pénalité qui minimise le coût. En pointillés verts : valeur de pénalité qui maximise le  $F_{score}$ .

La comparaison au niveau des performances de détection a été menée sur les trois séries, que ce soit pour les données d’entraînement et de validation (50 signaux de chaque). Dans le Tableau 4.5, nous pouvons voir que les observations faites pour la série D2 sont valables quels que soient le jeu de données considéré et les signaux étudiés. **En effet, le  $F_{score}$  est toujours moins bon si nous utilisons la pénalité qui minimise la fonction de coût, et donc, choisie par Ruptures.** Les valeurs de *Recall* et *Precision* suivent également le raisonnement énoncé plus haut. Par ailleurs, hormis pour la série D1, nous pouvons remarquer que le  $F_{score}$  est plus faible avec les données de validation, et ce quelle que soit la pénalité choisie. Cette différence peut s’expliquer par le fait que les positions des marches ne sont pas connues lors de l’étape de validation et donc que Ruptures ne possède pas de point de comparaison. Enfin, plus globalement, les performances de détection diminuent pour un bruit de fluorescence croissant. **Ce comportement était**



attendu : l'algorithme a plus de mal à détecter précisément des ruptures dans la valeur moyenne pour des signaux très bruités.

	D1		D2		D3	
	Ent.	Val.	Ent.	Val.	Ent.	Val.
$Recall_{Rupt}$	95,0 %	97,0 %	78,8 %	75,9 %	53,7 %	51,2 %
$Recall_{opt}$	94,7 %	96,5 %	76,6 %	72,9 %	47,8 %	43,9 %
$Precision_{Rupt}$	94,9 %	96,0 %	71,3 %	64,1 %	39,3 %	38,7 %
$Precision_{opt}$	97,8 %	98,2 %	84,8 %	75,4 %	66,6 %	59,2 %
$F_{score,Rupt}$	95,0 %	96,5 %	74,8 %	69,5 %	45,4 %	44,1 %
$F_{score,opt}$	96,2 %	97,3 %	80,5 %	74,1 %	55,6 %	50,4 %

**TABLEAU 4.5** – Données d'entraînement et de validation : comparaison des métriques selon la pénalité choisie. "Ent." correspond aux données d'entraînement et "Val." aux données de validation. Les métriques qui portent en indice la mention "Rupt" ont été calculées à partir des prédictions faites avec la pénalité qui minimise le coût. En revanche, si elles sont suivies par "opt", cela signifie que la pénalité utilisée maximise le  $F_{score}$ .

#### 4.3.1.2 Capacité des algorithmes à bien détecter les marches

Les phases d'entraînement et de validation pour Ruptures nous ont permis de déterminer la pénalité pour laquelle les meilleures performances sont obtenues. Nous avons donc choisi la pénalité maximisant le  $F_{score}$  afin de travailler sur les données tests composées de 1500 signaux. En ce qui concerne quickPBSA, le seuil utilisé dépend du jeu de données considéré. Le Tableau 4.6 résume les performances de détection des deux algorithmes pour les séries D1 à D3.

	D1		D2		D3	
	quickPBSA	Ruptures	quickPBSA	Ruptures	quickPBSA	Ruptures
$Recall$	91,4 %	95,3 %	69,7 %	71,8 %	48,2 %	42,9 %
$Precision$	92,7 %	98,2 %	68,6 %	81,5 %	49,5 %	62,6 %
$F_{score}$	92,0 %	96,8 %	69,1 %	76,4 %	48,8 %	50,9 %

**TABLEAU 4.6** – Données tests : comparaison des métriques obtenues avec quickPBSA et Ruptures. Pour Ruptures, les métriques ont été calculées à partir des prédictions réalisées avec la pénalité qui maximise le  $F_{score}$ .

Comme observé sur les données d'entraînement et de validation, plus le bruit de fluorescence augmente, plus les performances des deux algorithmes diminuent. **De plus, pour les trois séries de données, Ruptures montre un meilleur  $F_{score}$  que quickPBSA : les estimations faites par le premier sont plus proches de la *Ground-truth* que celles du second.** En regardant les résultats plus en détails, Ruptures détecte un nombre de faux positifs plus faible que quickPBSA, ce qui se traduit par une meilleure  $Precision$ . En revanche, il semblerait que pour un bruit de fluorescence croissant, le nombre de faux

négatifs repérés par Ruptures devient plus important que celui de quickPBSA : le *Recall* du second finit par être meilleur. En d'autres termes, Ruptures manque plus de marches.

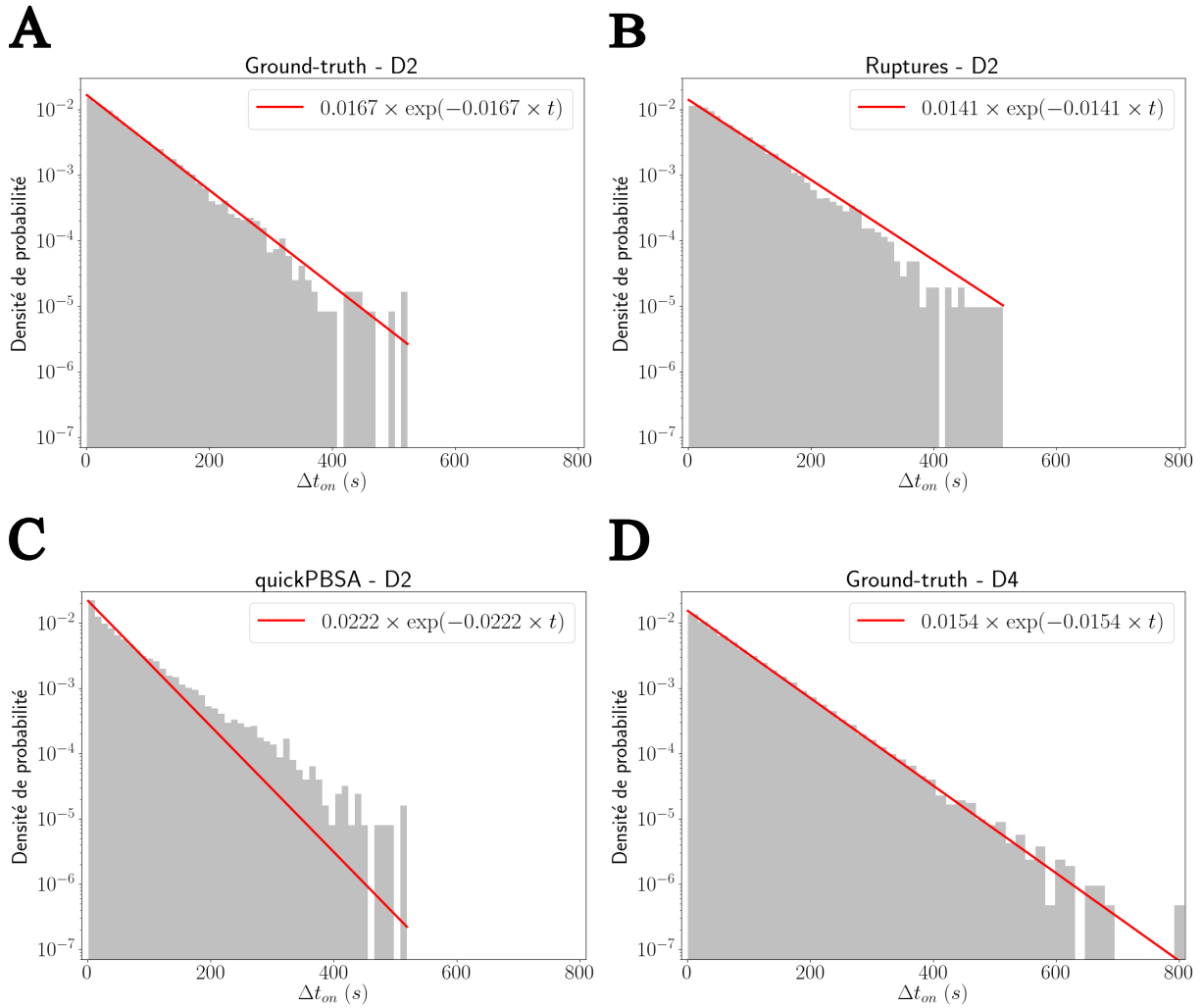
### 4.3.1.3 Influence sur le calcul des paramètres d'assemblage

Après avoir comparé les deux algorithmes de détection sur leur faculté à se rapprocher du mieux possible de la *Ground-truth*, nous pouvons évaluer l'influence de ces estimations sur le calcul des paramètres d'assemblage.

#### - Constante d'association :

Afin de calculer la constante d'association, je relève simplement l'écart temporel  $\Delta t_{on}$  entre deux marches montantes consécutives que ce soit pour la *Ground-truth* ou les ajustements faits par Ruptures et quickPBSA. Ces valeurs sont collectées sur l'ensemble des 1500 signaux tests et la distribution des  $\Delta t_{on}$  est tracée (Figure 4.8). À partir de cette distribution, un ajustement exponentiel est réalisé et normalement, il devrait correspondre à l'allure de l'histogramme obtenu. La constante de temps de l'exponentielle nous donne directement la valeur de  $k_{on}$ . Pour les simulations, nous avons choisi une constante d'association de  $0,015 \text{ s}^{-1}$  qui semble typique des expériences menées à l'équilibre.

Sur la Figure 4.8.A-B-C, trois distributions ont été représentées pour la série D2. Elles correspondent respectivement aux constantes d'association obtenues avec la *Ground-truth*, Ruptures et quickPBSA dont les valeurs issues de l'ajustement sont les suivantes :  $0,0167 - 0,0141 - 0,0221 \text{ s}^{-1}$ . Dans un premier temps, il est intéressant de remarquer que la valeur obtenue avec la *Ground-truth* est supposée être celle qui est la plus proche de la valeur théorique de  $0,015 \text{ s}^{-1}$  vu que nous connaissons les positions des marches. Or, dans le cas de la série D2, c'est Ruptures qui semble donner la meilleure estimation de  $k_{on}$ . Pour trouver l'origine de ce décalage entre le paramètre de simulation et le résultat issu de la *Ground-truth*, il faut se pencher sur la statistique des valeurs de  $\Delta t_{on}$ . En effet, sur 1600 images, il est possible de détecter en moyenne 20 marches par signal. En considérant une hypothèse simple selon laquelle, une marche sur deux est montante, on obtient environ 10 valeurs de  $\Delta t_{on}$  pour chaque signal. Ce nombre de tirages parmi une loi exponentielle de constante de temps  $0,015 \text{ s}^{-1}$  semble trop faible pour être totalement représentatif de la distribution associée, notamment aux  $\Delta t_{on}$  longs. Pour remédier à ce problème, j'ai généré un nouveau jeu de données, appelé D4, avec un bruit de fluorescence égal à celui de D2 mais où chaque signal possède cette fois-ci 16000 images. La Figure 4.8.D montre la distribution des  $\Delta t_{on}$  pour cette quatrième série. Dans ce cas, la constante de temps est de  $0,0153 \text{ s}^{-1}$ , ce qui est bien plus proche de la valeur théorique. **En multipliant la durée des signaux par 10, plus d'une centaine de valeurs de  $\Delta t_{on}$  sont collectées par signal, ce qui permet d'être plus représentatif de la condition imposée en simulation mais également d'affiner la distribution aux  $\Delta t_{on}$  longs.**



**FIGURE 4.8** – **Distributions des temps entre deux marches montantes consécutives.** Les trois premiers histogrammes concernent la série D2 et les résultats obtenus avec les données tests de A) la *Ground-truth*, B) Ruptures et C) quickPBSA. En D) on regarde de nouveau la *Ground-truth* mais pour la série D4. Les courbes rouges correspondent aux ajustements réalisés avec une loi exponentielle.

Ainsi, pour les signaux générés avec 1600 images, la valeur attendue pour la constante d'association est bien celle issue de la *Ground-truth*, à savoir  $0,0167 \text{ s}^{-1}$  pour D2. En comparant les deux algorithmes de détection, nous voyons que Ruptures renvoie une distribution plus proche d'un ajustement exponentiel que quickPBSA. Pour aller plus loin, la valeur de  $k_{on}$  est surestimée pour quickPBSA alors qu'elle est sous-estimée pour Ruptures. Ce comportement illustre tout à fait les différences observées au niveau des métriques d'analyse. En effet, pour quickPBSA, le *Recall* est plus élevé que la *Precision*, ce qui signifie que le nombre de faux positifs est majoritaire. **En d'autres termes, plus de fausses marches sont repérées par l'algorithme, dont certaines d'entre elles peuvent être montantes, ce qui contribue à collecter de faibles valeurs de  $\Delta t_{on}$  : la constante d'association est donc plus grande.** En ce qui concerne Ruptures, la *Precision* étant plus importante que le *Recall*, le raisonnement est inversé.

	D1 (s <sup>-1</sup> )		D2 (s <sup>-1</sup> )		D3 (s <sup>-1</sup> )		D4 (s <sup>-1</sup> )	
	Fit	Moy.	Fit	Moy.	Fit	Moy.	Fit	Moy.
<i>Ground-truth</i>	<b>0,0174</b>	<b>0,0170</b>	<b>0,0167</b>	<b>0,0167</b>	<b>0,0168</b>	<b>0,0166</b>	<b>0,0154</b>	<b>0,0150</b>
Ruptures	0,0167	0,0165	0,0141	0,0149	0,0112	0,0120	0,0119	0,0127
quickPBSA	0,0201	0,0171	0,0222	0,0175	0,0188	0,0165	0,0180	0,0150
Valeur attendue	0,015							

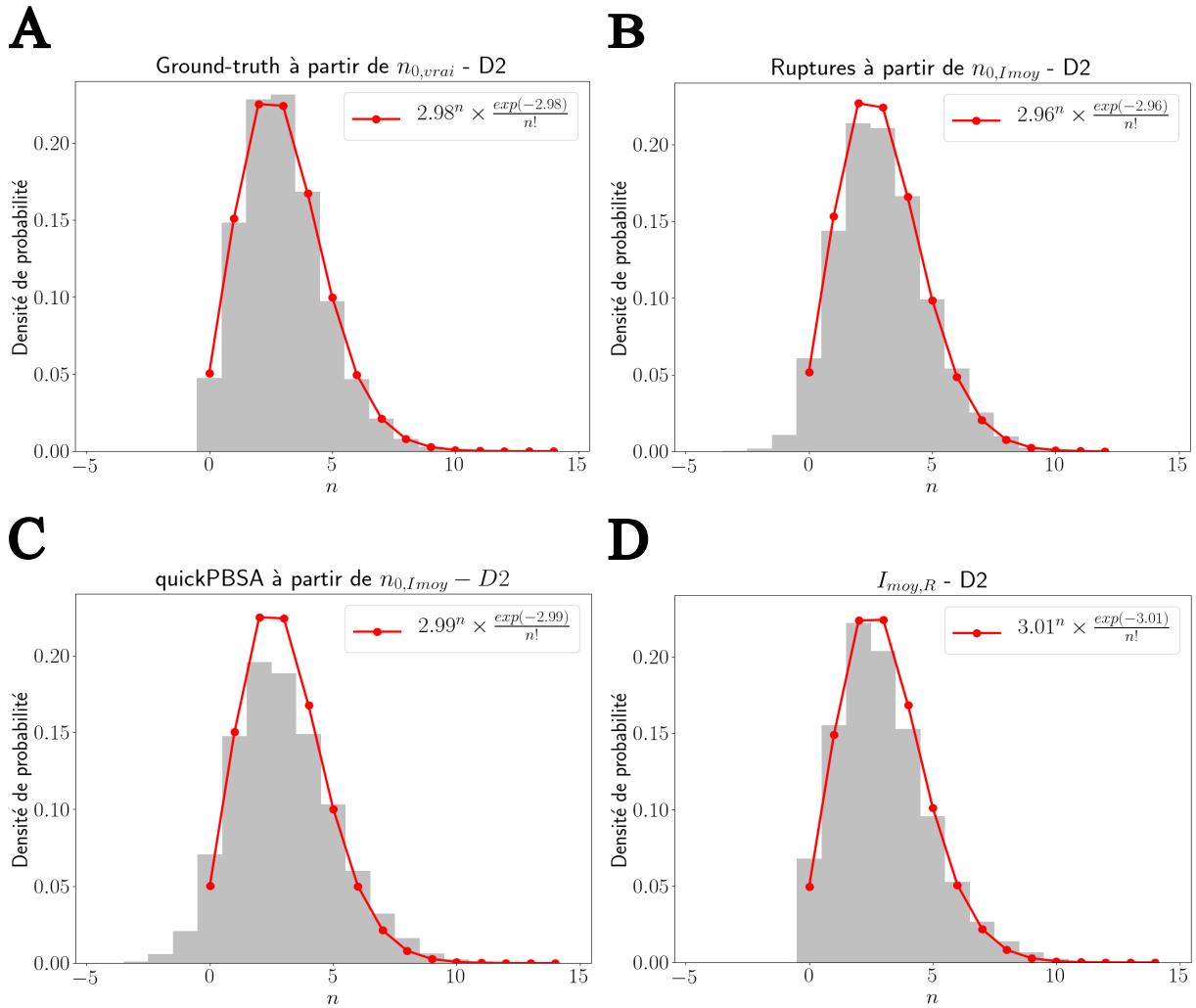
**TABLEAU 4.7 – Données tests : estimation des constantes d’association  $k_{on}$ .** Les constantes d’association ont été calculées sur les 1500 signaux tests des séries D1 à D4 à partir de la *Ground-truth*, Ruptures et quickPBSA. Les valeurs issues des colonnes "Fit" donnent les  $k_{on}$  estimées avec la constante de temps de l’ajustement exponentiel. Pour les colonnes "Moy.", la valeur moyenne des  $\Delta t_{on}$  a été calculée.

Pour finir, le Tableau 4.7 regroupe les valeurs de  $k_{on}$  estimées avec Ruptures et quickPBSA ou pour la *Ground-truth*, et ce pour chacun des quatre jeux de données. De manière générale, pour Ruptures et la *Ground-truth*, l’estimation faite par l’ajustement et la valeur moyenne des  $\Delta t_{on}$  sont équivalentes. Par ailleurs, Ruptures sous-estime la valeur de la constante d’association attendue pour les quatre séries étudiées. Au contraire, avec l’ajustement exponentiel, quickPBSA surestime tout le temps  $k_{on}$  mais est très proche de la *Ground-truth* si on calcule manuellement la valeur moyenne. **En d’autres termes, quickPBSA a plus de difficultés que Ruptures à traduire la distribution exponentielle obtenue avec la *Ground-truth*.** Dans tous les cas, cette analyse nous permet d’avoir un regard critique sur les estimations faites par les deux algorithmes. **Pour la série D2, l’ajustement exponentiel de Ruptures sous-estime  $k_{on}$  de 16 % alors que celui de quickPBSA la surestime de 32 %.**

#### - Nombre moyen de sous-unités marquées :

Dans un premier temps, comme avec  $k_{on}$ , je ne m’intéresse qu’aux estimations de  $\langle n \rangle$  faites pour la série D2. Ici, les estimations ont été réalisées selon différentes façons. Pour la *Ground-truth*, les valeurs de  $n$  sont connues à chaque image et la distribution représentée en Figure 4.9.A est donc celle qui fait office de référence. Dans le titre de cette dernière, je note  $n_{0,vrai}$  le nombre attendu de sous-unités marquées présents en début de signal. En ce qui concerne Ruptures et quickPBSA, le nombre initial de dimères marqués est estimé en prenant la partie entière du rapport entre l’intensité initiale  $I_0$  et la hauteur moyenne d’une marche  $h$  : on note cette valeur  $n_{0,I_{moy}}$ . Ensuite l’évolution de  $n$  au cours du temps suit les ajustements réalisés par les deux algorithmes : une marche montante correspond au gain d’une sous-unité alors qu’un événement descendant implique la perte d’un dimère marqué. Cette méthode d’estimation à partir de Ruptures et quickPBSA est illustrée en Figure 4.9.B-C. Les trois distributions ainsi obtenues sont toutes comparées à une loi de Poisson comme le montrent les courbes rouges. Par ailleurs, pour la *Ground-truth*, Ruptures et quickPBSA, une valeur de  $n$  a été collectée pour chacune des 1600 images

sur les 1500 signaux test, ce qui fait  $2,4 \times 10^6$  valeurs.



**FIGURE 4.9** – Distributions du nombre de sous-unités marquées par ARN pour la série **D2**. Les histogrammes A,B et C ont été obtenus respectivement avec les données tests de la *Ground-truth*, Ruptures et quickPBSA : à chaque rupture repérée, une sous-unité est retirée ou ajoutée au nombre précédent en fonction de la nature de la marche. En D) la méthode  $I_{moy,R}$  a été utilisée à partir de l’ajustement provenant de Ruptures. Les courbes rouges correspondent aux ajustements réalisés avec une loi de Poisson.

En comparant les allures des trois premiers histogrammes, nous pouvons dire qu’elles sont globalement similaires. De plus, les estimations de  $\langle n \rangle$  sont très proches de la valeur imposée en simulation qui est de 3 sous-unités. **Cependant, comme pour la constante d’association, la distribution issue de quickPBSA est la plus éloignée de l’ajustement théorique tracé en rouge, alors qu’avec Ruptures la différence est moins marquée.** L’origine de ce décalage, qui n’est pas observé avec la *Ground-truth*, peut être lié à la présence de valeurs négatives dans les distributions des Figures 4.9.B-C. Ces valeurs négatives illustrent parfaitement l’existence de faux-positifs au sein des prédictions faites par chacun des deux algorithmes utilisés. **En effet, si Ruptures ou quickPBSA posi-**

**tionnent deux marches descendantes alors que le nombre de sous-unités vaut 1,  $n$  deviendra négatif à la suite de ces deux événements.** Le nombre de valeurs négatives est relativement faible en comparaison aux  $2,4 \times 10^6$  estimations collectées sur toute une série de données tests. Par exemple pour D2, elles représentent respectivement 2,8 % et 1,3 % de la statistique totale avec quickPBSA et Ruptures. Néanmoins, nous avons voulu voir si une autre méthode de calcul de  $n$  pouvait nous permettre d'éviter ces erreurs, tout en étant proche de la valeur imposée en simulation.

Pour ce faire, une seconde méthode d'estimation de  $\langle n \rangle$  a été utilisée à partir des prédictions faites par Ruptures d'où la notation  $I_{moy,R}$ . Dans ce cas, l'intensité de l'ajustement est divisée à chaque image par  $h$  puis la valeur entière la plus proche est conservée. Nous obtenons donc également  $2,4 \times 10^6$  valeurs de  $n$  et l'histogramme qui illustre cette méthode est représenté à la Figure 4.9.D. **En utilisant la hauteur moyenne d'une marche  $h$ , aucune valeur négative ne peut être obtenue. Par ailleurs, nous pouvons remarquer que l'ajustement par une loi de Poisson fonctionne très bien et que  $\langle n \rangle$  reste très proche de la valeur théorique.** La méthode  $I_{moy,R}$  semble donc tout à fait adaptée à ce que nous souhaitons mesurer. Dans le Tableau 4.8, la comparaison de ces différentes méthodes d'estimation est élargie aux séries D1 et D3. En quelques mots, les valeurs de  $\langle n \rangle$  sont toujours très proches de la *Ground-truth* ou de la valeur théorique, que ce soit pour Ruptures, quickPBSA ou  $I_{moy,R}$ . En effet, malgré un bruit de fluorescence plus important et des performances de détection plus faibles pour la série D3, l'estimation de  $\langle n \rangle$  est toujours très bonne. Enfin, pour les trois séries, quickPBSA et Ruptures renvoient des valeurs négatives pour  $n$ .

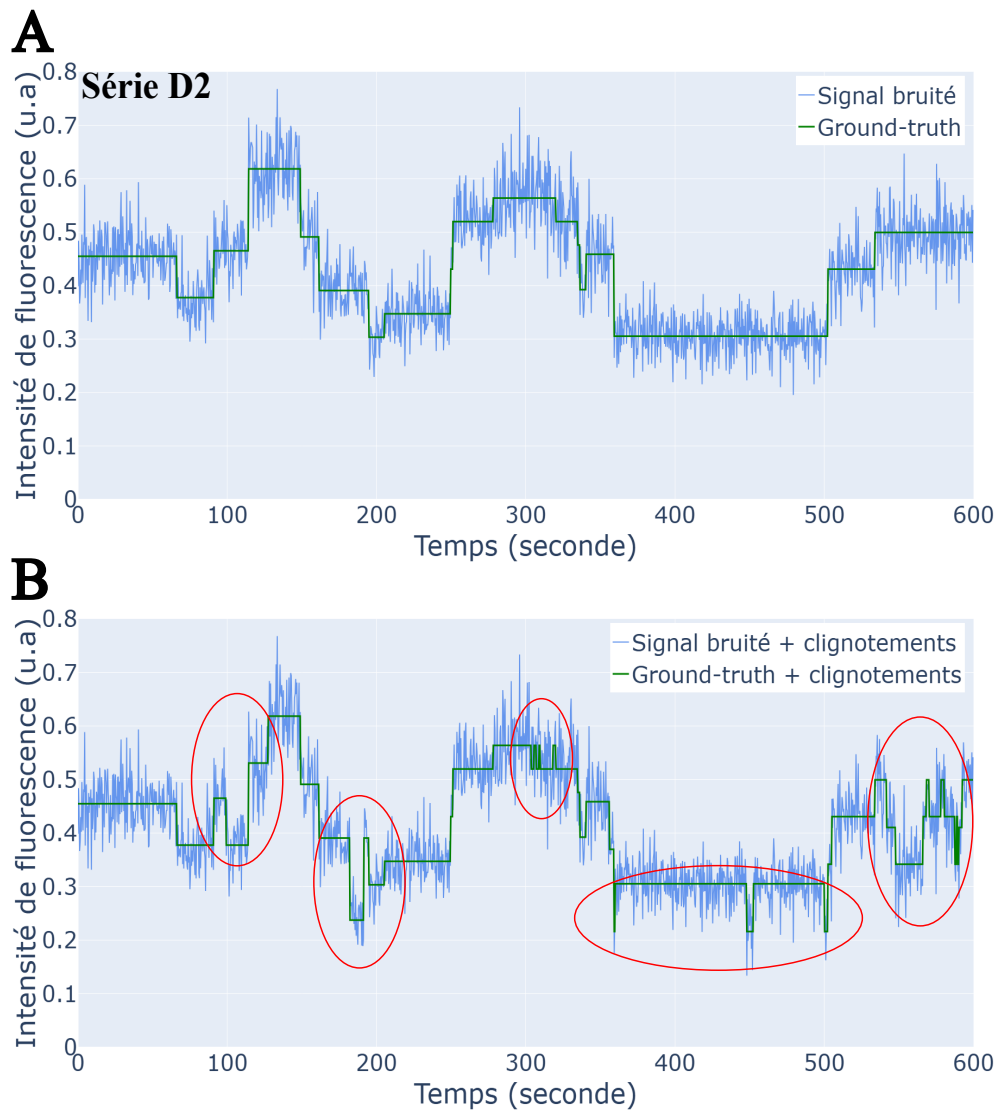
	D1			D2			D3		
	$n_{min}$	$n_{max}$	$\langle n \rangle$	$n_{min}$	$n_{max}$	$\langle n \rangle$	$n_{min}$	$n_{max}$	$\langle n \rangle$
<b><i>Ground-truth</i></b>	<b>0</b>	<b>12</b>	<b>3,01</b>	<b>0</b>	<b>14</b>	<b>2,98</b>	<b>0</b>	<b>12</b>	<b>3,01</b>
Ruptures	-2	13	3,00	-3	12	2,96	-3	11	3,01
quickPBSA	-3	13	3,03	-4	14	2,99	-5	12	3,03
$I_{moy,R}$	0	13	3,01	0	14	3,01	0	13	3,05
Valeur attendue	$\langle n \rangle = 3$								

**TABLEAU 4.8 – Données tests : estimation du nombre moyen de sous-unités marquées  $\langle n \rangle$ .** Les valeurs de  $\langle n \rangle$  ont été obtenues à partir des données tests de la *Ground-truth*, Ruptures, quickPBSA et  $I_{moy,R}$  pour les séries D1 à D3. Les colonnes notées  $n_{min}$  et  $n_{max}$  correspondent respectivement aux valeurs minimales et maximales de  $n$  pour chaque distribution.

### 4.3.2 ... pour un fluorophore qui présente des clignotements

Dans cette dernière section, je montre comment les résultats décrits précédemment pour la *Ground-truth*, Ruptures et quickPBSA sont modifiés en présence de clignotements. Pour rappel, au cours de la génération de données, les séries D1 à D3 ont été dupliquées

afin de modifier la photostabilité des fluorophores. Les variables initiales de simulation ont été conservées mais trois nouveaux paramètres ont été ajoutés. **En effet, on considère qu'environ 20 % des fluorophores clignotent, que la durée moyenne de passages en OFF vaut 25 secondes et que le nombre moyen de passage en OFF est égal à 2,2.** Ces valeurs ont été estimées expérimentalement sur des données de photophysique. La fonction permettant de rajouter les clignotements aux signaux simulés est détaillée à l'annexe F. La Figure 4.10 montre un exemple de signal appartenant à la série D2 avec et sans clignotements.



**FIGURE 4.10** – Illustration des clignotements sur la série D2. Signal de la série D2 A) sans clignotements et B) avec. La *Ground-truth* est tracée en vert, le signal bruité en bleu et les clignotements sont entourés en rouge.

Avec ces changements, Ruptures et quickPBSA ont donc dû effectuer de nouvelles prédictions sur la totalité des signaux étudiés. Pour l'algorithme développé par Hummert et al. (2021), le seuil de détection des marches dépend encore une fois du jeu de données

considéré (voir annexe le Tableau F.2). En ce qui concerne Ruptures, ce dernier a été de nouveau entraîné sur chacune des séries de données ce qui signifie que les pénalités maximisant le  $F_{score}$  ne sont plus les mêmes. Je ne reviens pas en détails sur l'évolution de ces pénalités pour chaque jeu de données mais ce qui est important de retenir est que ces dernières sont toutes plus faibles que précédemment. Pour la série D2, la pénalité maximisant le  $F_{score}$  est de 1,28 sans clignotements, elle vaut désormais 1,06. Cette évolution est cohérente avec la définition donnée à la section 4.1.2 : **comme seul le nombre de marches a augmenté, la pénalité est plus faible**. Le Tableau F.1 regroupe les pénalités obtenues avec et sans clignotements pour les trois séries.

En prenant en compte les nouvelles valeurs de pénalités, il est possible d'évaluer les performances de Ruptures et quickPBSA. Les valeurs de *Recall*, *Precision* et  $F_{score}$  sont regroupées dans le Tableau 4.9 pour les séries D1 à D3. **Nous pouvons voir que quel que soit le bruit de fluorescence considéré, les performances globales de Ruptures sont toujours meilleures que celles de quickPBSA. De plus, dans tous les cas, le  $F_{score}$  est plus faible avec clignotements**. Ce résultat était attendu puisqu'en rajoutant de nombreux événements, dont certains ne durent que quelques images, quickPBSA et Ruptures sont plus susceptibles de se tromper. Par ailleurs, il est important de remarquer que la *Precision* n'évolue presque pas pour les deux algorithmes et que c'est surtout le *Recall* qui diminue. En d'autres termes, les algorithmes n'inventent pas plus de marches mais en manquent plus.

	D1		D2		D3	
	quickPBSA	Ruptures	quickPBSA	Ruptures	quickPBSA	Ruptures
<i>Recall</i>	91,4 %	95,3 %	69,7 %	71,8 %	48,2 %	42,9 %
<i>Recall<sub>cl</sub></i>	89,1 %	93,5 %	61,3 %	61,8 %	37,9 %	34,7 %
<i>Precision</i>	92,7 %	98,2 %	68,6 %	81,5 %	49,5 %	62,6 %
<i>Precision<sub>cl</sub></i>	88,6 %	98,1 %	67,5 %	80,7 %	51,7 %	60,3 %
$F_{score}$	92,0 %	96,8 %	69,1 %	76,4 %	48,8 %	50,9 %
$F_{score,cl}$	88,8 %	95,7 %	64,3 %	70,0 %	43,8 %	44,1 %

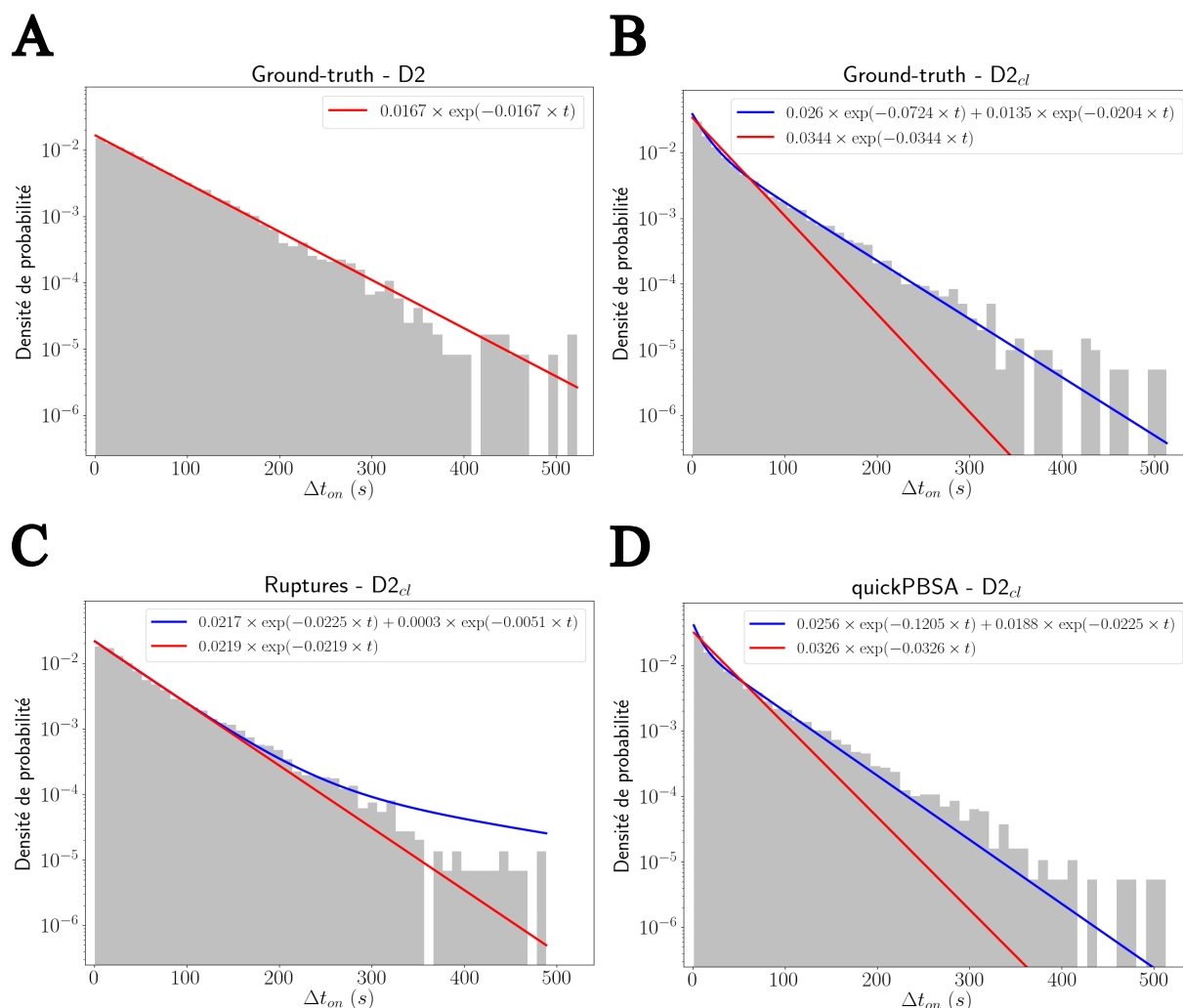
TABLEAU 4.9 – Données tests : métriques caractérisant quickPBSA et Ruptures avec et sans clignotements. Pour Ruptures, les métriques ont été calculées à partir des prédictions réalisées avec la pénalité qui maximise le  $F_{score}$ . Les résultats en rouge correspondent aux données avec clignotements et sont indexés par "cl".

Il est désormais intéressant de regarder à quel point la présence de clignotements peut affecter l'estimation des paramètres d'assemblage d'intérêt tels que  $\langle n \rangle$  et  $k_{on}$ .

Pour commencer, si l'on raisonne à partir de la *Ground-truth*, il est logique de s'attendre à une augmentation de la constante d'association et donc à une diminution de la valeur moyenne de  $\Delta t_{on}$ . En effet, sachant que plusieurs passages en OFF sont rajoutés à chaque signal, la durée entre deux événements montants



consécutifs ne peut que diminuer. De plus, une simple loi exponentielle ne semble plus être le meilleur modèle pour décrire toute la population des  $\Delta t_{on}$  comme cela est le cas sans clignotements (Figure 4.11.A). Il est vrai qu'en regardant la courbe rouge de la Figure 4.11.B, nous pouvons remarquer que l'ajustement exponentiel est bien adapté pour les petites valeurs de temps, mais beaucoup moins dès que l'on dépasse une cinquantaine de secondes. **Cependant, en faisant évoluer le modèle vers une double exponentielle représentée en bleu, la distribution est mieux décrite.**



**FIGURE 4.11** – Influence des clignotements sur la distribution des  $\Delta t_{on}$ . Les quatre histogrammes concernent les données tests de la série D2 et les résultats obtenus sans clignotements pour A) la *Ground-truth* et avec clignotements pour B) la *Ground-truth*, C) Ruptures et D) quickPBSA. Les courbes rouges et bleues correspondent aux ajustements réalisés respectivement avec une exponentielle et une double exponentielle.

En réalité, ce changement est cohérent avec la façon dont les données ont été générées. En effet, les descriptions faites à la section 4.2.2 et en annexe F, mettent en évidence que deux lois mono-exponentielles de constantes de temps  $0,040 \text{ s}^{-1}$  et  $0,015 \text{ s}^{-1}$  sont utilisées afin de choisir respectivement la durée des clignotements et la position des marches. Pour

la série D2, on trouve  $0,072 \text{ s}^{-1}$  et  $0,020 \text{ s}^{-1}$ , ce qui est très proche des deux paramètres de simulation. La légère surestimation peut être une nouvelle fois associée à une statistique trop faible, et donc à des signaux trop courts. En réalisant la même étude avec la série de données D4, pour laquelle les signaux comptent 16 000 images, les deux constantes de temps diminuent légèrement et sont plus proches des valeurs utilisées en simulation :  $0,057 \text{ s}^{-1}$  et  $0,017 \text{ s}^{-1}$ .

En regardant cette fois-ci les Figures 4.11.C et D, nous pouvons nous faire une idée sur la proximité des résultats obtenus par les deux algorithmes de détection avec ceux de la *Ground-truth*. Pour Ruptures, nous voyons que l'allure de la distribution est plus proche d'un ajustement avec une simple exponentielle plutôt qu'avec une double exponentielle. Même si la constante d'association est plus élevée par rapport à la situation sans clignotements ( $0,022 \text{ s}^{-1}$  contre  $0,014 \text{ s}^{-1}$ ), **Ruptures ne permet pas de distinguer clairement la population associée au mécanisme photophysique de celle caractéristique des instants d'arrivée des protéines.** En revanche, quickPBSA, renvoie une distribution très proche de l'ajustement en double exponentielle et donc des résultats obtenus par la *Ground-truth*. La même analyse a été réalisée pour les séries D1 et D3 et les valeurs de constante d'association sont regroupées dans le Tableau 4.10.

	D1 ( $\text{s}^{-1}$ )			D2 ( $\text{s}^{-1}$ )			D3 ( $\text{s}^{-1}$ )		
	Fit-1	Fit-2	Moy.	Fit-1	Fit-2	Moy.	Fit-1	Fit-2	Moy.
<i>GT</i>	/	<b>0,0174</b>	<b>0,0170</b>	/	<b>0,0167</b>	<b>0,0167</b>	/	<b>0,0168</b>	<b>0,0166</b>
<i>GT<sub>cl</sub></i>	<b>0,0676</b>	<b>0,0189</b>	<b>0,0267</b>	<b>0,0724</b>	<b>0,0204</b>	<b>0,0262</b>	<b>0,0623</b>	<b>0,0175</b>	<b>0,0263</b>
R	/	0,0167	0,0165	/	0,0141	0,0149	/	0,0112	0,0120
<i>R<sub>cl</sub></i>	<b>0,0645</b>	<b>0,0185</b>	<b>0,0253</b>	<b>0,0225</b>	<b>0,0051</b>	<b>0,0200</b>	<b>0,0602</b>	<b>0,0144</b>	<b>0,0157</b>
qB	/	0,0201	0,0171	/	0,0222	0,0175	/	0,0188	0,0165
<i>qB<sub>cl</sub></i>	<b>0,0844</b>	<b>0,0200</b>	<b>0,0267</b>	<b>0,121</b>	<b>0,0225</b>	<b>0,0238</b>	<b>0,0530</b>	<b>0,0166</b>	<b>0,0194</b>
Simulation	$k_{on} = 0,015 \text{ s}^{-1}$								

**TABLEAU 4.10 – Données tests : comparaison des  $k_{on}$  avec et sans clignotements.** Les valeurs de constante d'association ont été obtenues à partir des données tests présentant ou non des clignotements pour la *Ground-truth* (GT), quickPBSA (qB) et Ruptures (R) sur les séries D1 à D3. Dans les colonnes "Moy.", la valeur moyenne des  $\Delta t_{on}$  a été calculée afin d'obtenir  $k_{on}$ . Les résultats en rouge correspondent aux données avec clignotements et sont indexés par "cl". Dans leur cas, une double exponentielle a été utilisée, ce qui explique la présence de deux constantes de temps : "Fit-1" et "Fit-2".

Si nous nous penchons sur ce tableau, il est difficile de tirer des conclusions aussi claires que celles obtenues dans le cas où les clignotements sont absents, mais quelques tendances peuvent être retenues. **Premièrement, lorsque  $k_{on}$  est estimée à partir de la valeur moyenne des  $\Delta t_{on}$ , la constante d'association est toujours plus grande en présence de clignotements et cette fois-ci, les deux algorithmes de détection renvoient une valeur inférieure à celle de la *Ground-truth*.** En ce qui

concerne la caractérisation des distributions par une double exponentielle, il ne semble pas y avoir un algorithme qui soit plus proche que l'autre de la *Ground-truth*, et ce quelle que soit la série de données considérée. En effet, Ruptures a tendance à trouver des valeurs inférieures aux constantes de temps de la *Ground-truth* quand l'ajustement en double exponentielle fonctionne (D1 - D3), mais la série D2 nous montre que ce n'est pas tout le temps le cas. Avec quickPBSA, la comparaison avec la *Ground-truth* dépend du niveau de bruit de fluorescence : les constantes de temps sont sur-estimées pour les séries D1 et D2, alors qu'elles sont sous-estimées pour D3.

Ainsi, pour avoir une première indication de l'effet des clignotements, le plus simple est de comparer les valeurs moyennes estimées avec les deux algorithmes par rapport à celle renvoyée par la *Ground-truth*, sans clignotements. **Dans ce cas, on remarque une surestimation d'environ 20 % pour Ruptures contre près de 43 % pour quickPBSA.**

Pour conclure cette section, l'influence des clignotements a aussi été étudiée sur l'estimation du nombre moyen de sous-unités marquées. Toutes les valeurs concernant cette analyse sont regroupées dans le Tableau 4.11. **Comme attendu, la valeur de  $\langle n \rangle$  estimée par chaque méthode d'estimation diminue en présence de clignotements et ce pour toutes les séries de données étudiées.** Mis à part ce changement, les conclusions tirées de l'analyse sans clignotements sont toujours valables ici. En effet, Ruptures et quickPBSA renvoient toujours des estimations négatives lorsque nous utilisons la nature des marches détectées pour calculer  $n$ . **Ainsi,  $I_{moy,R}$  reste encore une fois la meilleure méthode d'estimation pour suivre l'évolution du nombre de sous-unités marquées au cours du temps.**

	D1			D2			D3		
	$n_{min}$	$n_{max}$	$\langle n \rangle$	$n_{min}$	$n_{max}$	$\langle n \rangle$	$n_{min}$	$n_{max}$	$\langle n \rangle$
<b>GT</b>	<b>0</b>	<b>12</b>	<b>3,01</b>	<b>0</b>	<b>14</b>	<b>2,97</b>	<b>0</b>	<b>12</b>	<b>3,01</b>
<b>GT<sub>cl</sub></b>	<b>0</b>	<b>12</b>	<b>2,93</b>	<b>0</b>	<b>14</b>	<b>2,91</b>	<b>0</b>	<b>12</b>	<b>2,93</b>
R	-2	13	3,00	-3	12	2,96	-3	11	3,01
R <sub>cl</sub>	-3	13	2,92	-4	12	2,92	-3	11	2,97
qB	-3	13	3,04	-5	14	2,99	-5	12	3,04
qB <sub>cl</sub>	-5	16	2,92	-7	15	2,92	-5	12	2,97
$I_{moy,R}$	0	13	3,01	0	14	3,01	0	13	3,05
$I_{moy,R,cl}$	0	13	2,93	0	14	2,93	0	15	2,97
Simulation	$\langle n \rangle = 3$								

**TABLEAU 4.11 – Données tests : comparaison des  $\langle n \rangle$  avec et sans clignotements.** Les valeurs de  $\langle n \rangle$  ont été obtenues à partir des données tests présentant ou non des clignotements pour la *Ground-truth* (GT), quickPBSA (qB), Ruptures (R) et  $I_{moy,R}$  sur les séries D1 à D3. Les colonnes notées  $n_{min}$  et  $n_{max}$  caractérisent respectivement les valeurs minimales et maximales de  $n$  pour chaque distribution. Les résultats en rouge correspondent aux données avec clignotements et sont indexés par "cl".

## 4.4 Conclusion

Pour conclure, la complexité des données de fluorescence auxquelles nous sommes intéressés rend presque impossible une analyse manuelle de ces dernières. En effet, les variations d'intensité collectées sont tellement hétérogènes selon leur nature ou leur dynamique que l'utilisation d'algorithmes de détection tels que Ruptures et quickPBSA est obligatoire. Ces derniers ont des modes de fonctionnement différents mais se basent tous les deux sur le même principe. Ils utilisent un paramètre (pénalité pour Ruptures et seuil pour quickPBSA) pour repérer des variations soudaines d'intensité au sein de la valeur moyenne d'un signal temporel. Cependant, au vue de l'hétérogénéité des données expérimentales que nous pouvons enregistrer, j'ai dû évaluer les performances de ces deux algorithmes sur des signaux de fluorescence simulés. La génération des différents jeux de données étudiés suit un modèle de dynamique d'assemblage sur lequel nous nous basons afin d'analyser les signaux expérimentaux. Les paramètres de simulation ont été choisis dans le but de coller le plus possible à l'expérience (série D2) mais également pour évaluer l'influence du niveau de bruit dans un signal sur les performances de Ruptures et quickPBSA (série D1 à D3).

Les études menées avec et sans clignotements nous ont permis d'en apprendre davantage sur les performances des deux algorithmes de détection. De manière générale, Ruptures semble plus cohérent avec la *Ground-truth* que quickPBSA, ce qui se matérialise notamment par un meilleur  $F_{score}$ . **De plus, sans clignotements, les distributions de  $\Delta t_{on}$  et  $n$  obtenues par quickPBSA sont plus éloignées des ajustements théoriques caractérisés respectivement par une loi exponentielle et une loi de Poisson.** Avec Ruptures, un certain décalage est observé, mais il est moins marqué. En revanche, en ce qui concerne le nombre moyen de sous-unités  $\langle n \rangle$ , les deux algorithmes souffrent du même biais. En effet, en suivant l'évolution du nombre de sous-unités marquées au travers de la nature des marches (montantes ou descendantes), Ruptures et quickPBSA ont tendance à faire apparaître des valeurs négatives de  $n$ , ce qui semble être directement lié aux performances de détection. Pour pallier ce problème, j'ai donc utilisé la méthode  $I_{moy,R}$  sur les ajustements de Ruptures, qui estime  $n$  à chaque image en prenant la valeur entière du rapport  $\frac{I_{fit}}{h}$  où  $I_{fit}$  est l'intensité de l'ajustement en marches d'escalier. **Les résultats obtenus à partir de la méthode  $I_{moy,R}$  donnent une valeur de  $\langle n \rangle$  très proche de celle issue de la *Ground-truth* et aucune valeur négative n'est obtenue.**

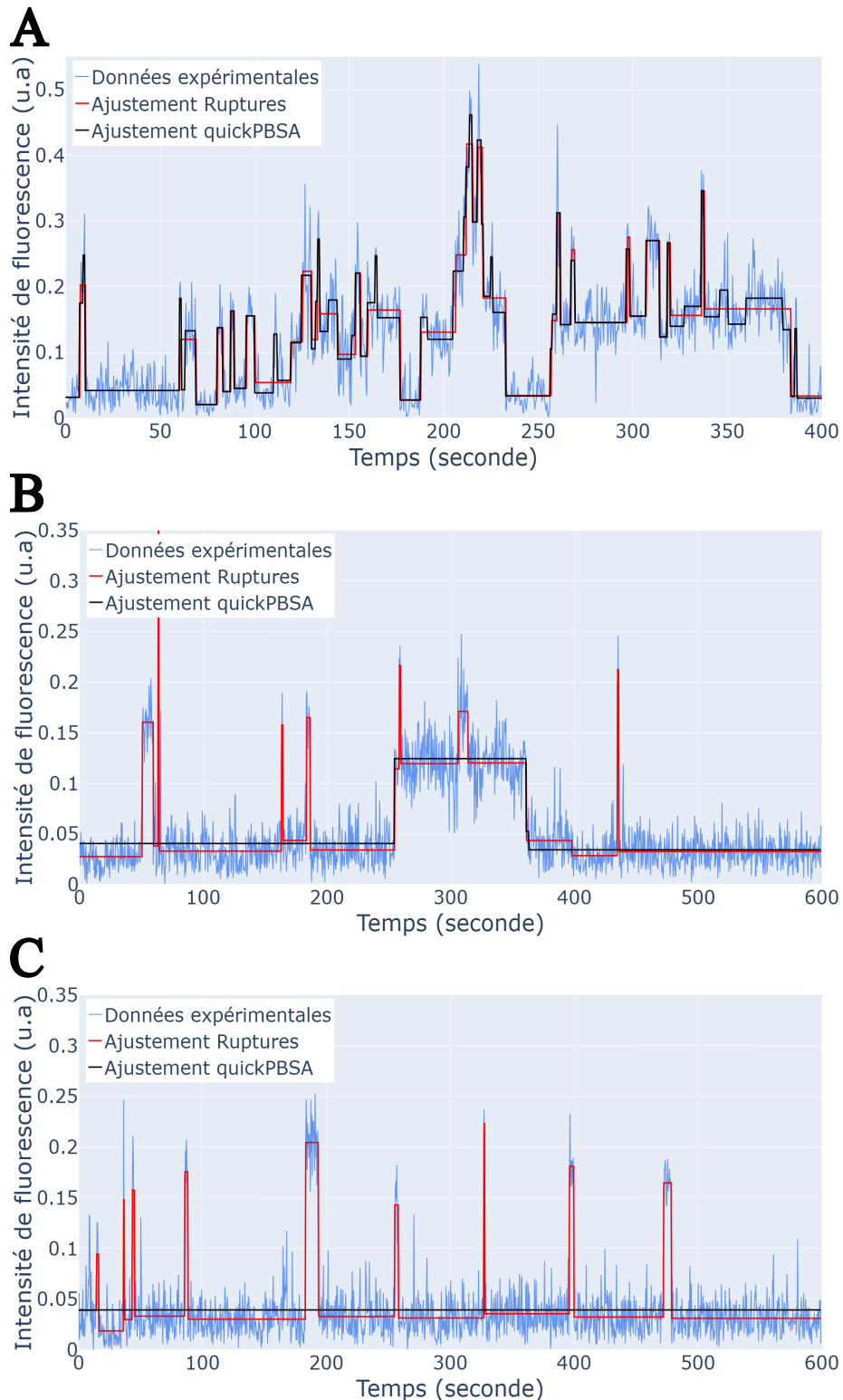
En présence de clignotements, les conclusions concernant l'estimation de  $\langle n \rangle$  restent les mêmes, malgré une baisse logique du nombre moyen de sous-unités marquées. Pour la constante d'association, les différentes valeurs calculées sont logiquement plus élevées en ajoutant des passages en OFF si l'on calcule manuellement la valeur moyenne des  $\Delta t_{on}$ .

Cependant, l'allure en double exponentielle, qui caractérise la façon dont les signaux sont simulés avec clignotements, est difficilement exploitable à partir des distributions renvoyées par les deux algorithmes. En effet, il n'y a pas vraiment de tendance claire qui se dégage si l'on compare les différentes séries de données. Ainsi, pour avoir une première estimation de l'influence des clignotements, le plus simple est de comparer les valeurs moyennes de  $\Delta t_{on}$  calculées manuellement. **Par exemple, pour la série D2, Ruptures renvoie une constante d'association avec clignotements plus grande que celle de la *Ground-truth* sans clignotements : la surestimation vaut environ 20 %, contre 43 % pour quickPBSA.** Cette correction est évidemment valable uniquement pour les paramètres de clignotements utilisés à savoir : environ 20 % de fluorophores qui clignotent, un nombre moyen de passages en OFF de 2,2 et une durée moyenne de passages en OFF de 25 secondes.

**Globalement, Ruptures semble donc donner de meilleurs résultats que quickPBSA sur des signaux simulés.** Afin de compléter leur comparaison, les deux algorithmes ont été testés sur des données expérimentales proches des conditions de dynamique d'assemblage à l'équilibre. Pour ce faire, Ruptures a été entraîné à partir de signaux annotés manuellement. En ce qui concerne quickPBSA, le seuil utilisé pour la série D2 avec clignotements a été choisi puisque ce jeu de données se rapproche le plus des signaux expérimentaux. Sur la Figure 4.12, je montre des signaux de fluorescence provenant de trois sites d'ARN différents. Dans le premier cas, les deux algorithmes repèrent de nombreuses marches et comme attendu, quickPBSA semble en détecter plus. En revanche, pour les deux autres cas où les variations d'intensité sont moins nombreuses, quickPBSA ne semble pas fonctionner. Pour ces mêmes signaux, les différents seuils sélectionnés entre 0,01 et 0,05 ne permettent pas de détecter de marches. **Il semblerait donc que Ruptures, avec une même pénalité, puisse mieux s'adapter à différents types de signaux expérimentaux que quickPBSA.** Ces deux raisons nous ont poussés à préférer Ruptures pour l'analyse des expériences de dynamique d'assemblage à l'équilibre. Il faut rappeler que quickPBSA a été développé afin de caractériser des marches descendantes, typiques du photoblanchiment. Il est donc assez logique que cet algorithme fonctionne moins bien sur nos données d'assemblage. Enfin, un dernier point positif en faveur de Ruptures est que nous sommes régulièrement en contact avec Charles Truong, l'une des personnes ayant participé à son développement. Nous avons pu échanger sur différents points afin de savoir si Ruptures est adapté à nos besoins et s'il est possible de l'améliorer.

Pour finir, comme je l'ai expliqué à la section 4.1.2, la pénalité renvoyée initialement par Ruptures est celle qui minimise la fonction de coût. Cependant, afin d'optimiser au mieux les performances de Ruptures, nous avons choisi de travailler avec la pénalité qui maximise le  $F_{score}$ . Le décalage entre ces deux valeurs semble dépendre des paramètres de

simulation utilisés. En effet, en présence de clignotements, la pénalité maximisant le  $F_{score}$  est multipliée par un facteur 1,75 pour la série D2 contre 2,25 pour D3. Ainsi, en partant de la pénalité de 1,259 qui minimise la fonction de coût expérimentale de la Figure 4.6, il est possible d'avoir une estimation de la pénalité théorique maximisant le  $F_{score}$ . En utilisant le facteur multiplicatif de la série D2, on trouve 2,20. **Cette valeur est très proche de la pénalité que j'utilise pour analyser les signaux expérimentaux typiques d'une dynamique d'assemblage à l'équilibre à savoir 2,47. Il est donc cohérent de conserver cette pénalité.**



**FIGURE 4.12** – Application de Ruptures et quickPBSA aux données expérimentales. A-B-C) Évolution de l'intensité de fluorescence issue de sous-unités marquées sur trois sites d'ARN différents. Les données expérimentales sont en bleu alors que les ajustements provenant de Ruptures et quickPBSA sont respectivement en rouge et noir. Ici, seules des sous-unités non marquées ont été incubées avec l'ARN pendant 15 minutes, puis 50 nM en protéines, dont 10 % sont marquées, ont été injectés. Les paramètres d'acquisition sont les suivants : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 5,4 mW.

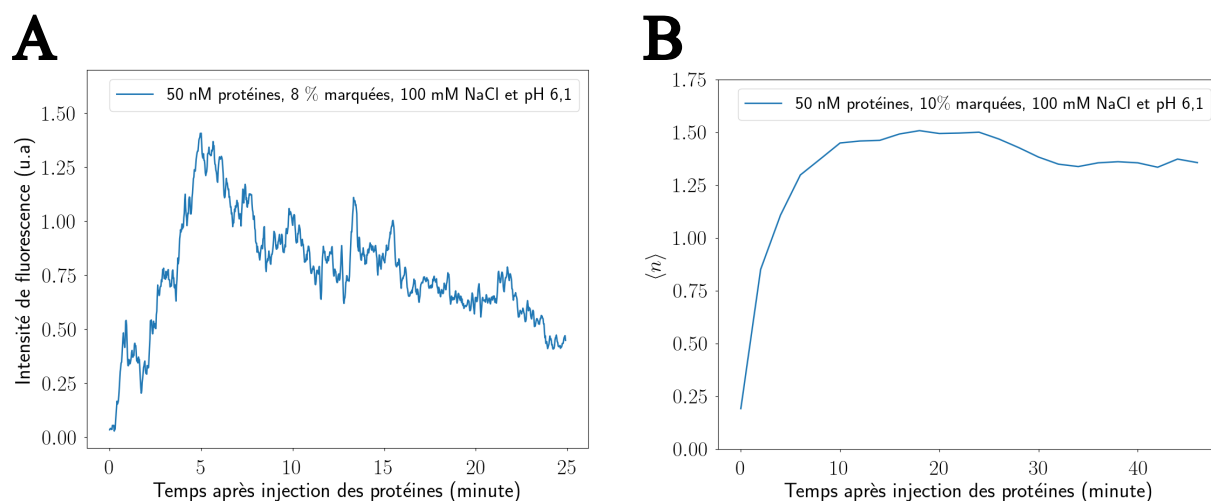
## Chapitre - 5

---

# Dynamique d'auto-assemblage du CCMV à l'équilibre et hors équilibre

---

Ce cinquième chapitre regroupe les principales analyses issues de l'étude du mécanisme d'auto-assemblage du CCMV. Pour commencer, je présente les résultats des assemblages étudiés en photométrie de masse. Ensuite, je décris les conditions expérimentales générales qui ont été utilisées en microscopie de fluorescence avant de détailler les différents contrôles réalisés vis-à-vis des interactions non spécifiques. Enfin, dans les parties quatre et cinq de ce chapitre, je reviens sur les deux principales expériences menées en TIRFM, qui portent respectivement sur l'étude de la dynamique d'assemblage à l'équilibre, puis hors équilibre (Figures 5.1.A et B).



**FIGURE 5.1** – Assemblage de capsides de CCMV vu sous deux conditions différentes. A) Suivi hors équilibre des premières étapes de l'assemblage de capsides pleines. B) Identification de l'état d'équilibre du mécanisme d'auto-assemblage du CCMV. Les deux expériences ont été menées en microscopie TIRF.



## 5.1 Caractérisation de nucléocapsides en photométrie de masse

Pour rendre compte de l'état d'assemblage de capsides virales, les techniques expérimentales se basant sur des mesures d'ensemble, telles que le SAXS, nécessitent généralement des concentrations initiales en protéines assez élevées, pouvant aller jusqu'à plusieurs dizaines de micromolaires. En effet, il est important d'obtenir suffisamment d'objets ayant une taille proche de celles de capsides pour réaliser une mesure convenable et collecter des informations précises. En revanche, en travaillant à l'échelle de la molécule unique, les concentrations utilisées deviennent assez limitantes. À la section 2.3.2.1, je présente la FCS en indiquant que les concentrations en objets marqués peuvent varier entre le picomolaire et le nanomolaire. En ce qui concerne le TIRFM, la limite maximale en espèce fluorescente ne doit pas excéder une trentaine de nanomolaires pour que la considération de "molécule unique" tienne toujours. Par ailleurs, à la section 3.2 je montre que la chimie de surface est une source de limitation supplémentaire de la concentration en sous-unités marquées puisqu'il est difficile de réaliser des expériences au-delà de 10 nM. Enfin, le taux de marquage utilisé ne peut pas être trop faible si l'on souhaite avoir une interprétation représentative du comportement de la totalité de la population en protéines.

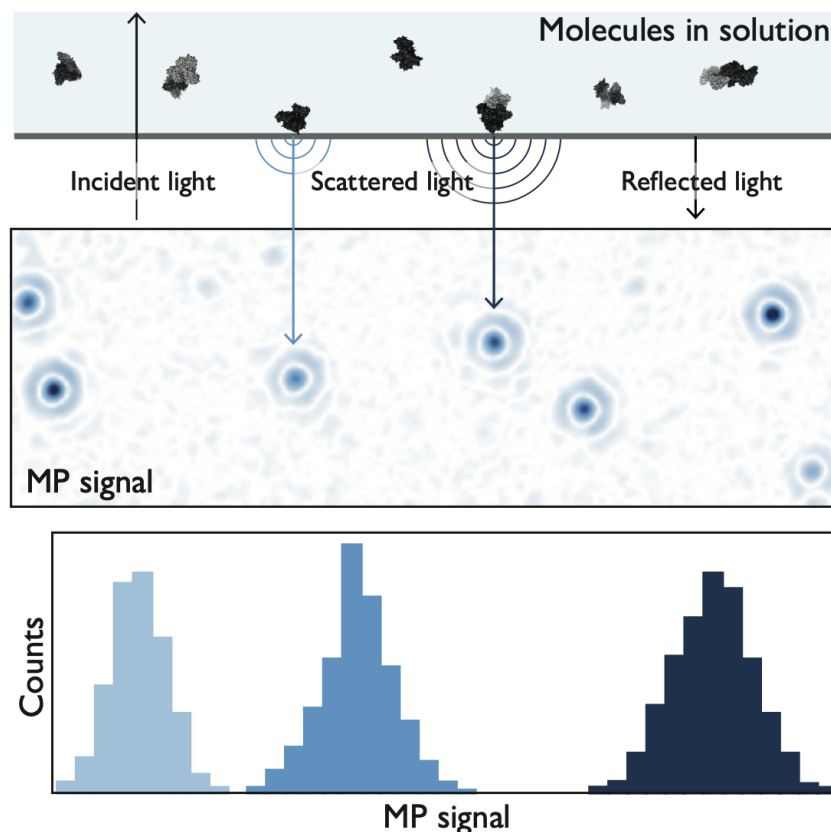
En prenant en compte ces différentes contraintes, les concentrations totales en protéines que nous utilisons en microscopie de fluorescence ne peuvent pas dépasser quelques centaines de nanomolaires. Il est donc important de caractériser la taille des objets obtenus dans ces conditions. **Pour ce faire, nous avons utilisé une autre technique expérimentale qui fonctionne à l'échelle de la molécule unique : la photométrie de masse.** Dans cette partie, je commence par présenter le principe de cette technique avant de détailler brièvement les conditions expérimentales utilisées. Les résultats expérimentaux sont décrits dans un second temps. Je précise que ces expériences ont été réalisées avec l'aide de Bertrand Raynal, ingénieur de recherche à l'Institut Pasteur, qui travaille, entre autres, avec l'appareil *TwoMP* de la compagnie *Refeyn*.

### 5.1.1 Réaliser une expérience en photométrie de masse

#### 5.1.1.1 Description de la technique

La photométrie de masse est une technique de molécule unique dont le principe découle de ce que je décris à la section 2.3.2.2 concernant l'iSCAT. En effet, comme le montre la Figure 5.2, le signal mesuré provient de l'interférence entre la lumière réfléchiée par la lamelle de verre et la lumière diffusée par les particules d'intérêt. Ces interférences créent donc un motif de diffraction pour chaque particule détectée, dont le contraste est notamment proportionnel au volume de cette dernière. Ainsi, durant une expérience réalisée en photométrie de masse, une goutte contenant les objets d'intérêt est déposée sur une lamelle de microscope et, à chaque fois qu'une particule rentre en contact avec la

surface de verre, l'appareil enregistre le contraste mesuré pour cette dernière. Il est donc possible de distinguer des objets de taille différente au sein d'un même échantillon et de tracer les histogrammes en contraste correspondants (Figure 5.2).



**FIGURE 5.2** – Principe de la photométrie de masse. On note *incident light* pour la lumière incidente, *scattered light* pour la lumière diffusée par la particule, *reflected light* pour la lumière réfléchiée par la surface et *MP signal* pour le signal, *i.e.* contraste, acquis en photométrie de masse. Cette figure est adaptée du site de la compagnie *Refeyn*.

De plus, la relation linéaire entre le contraste des interférences et le volume des objets détectés fait que le signal acquis en photométrie de masse est lui aussi proportionnel à leur masse molaire, dans la mesure où l'indice de réfraction est similaire pour les différentes particules considérées. Des résultats expérimentaux illustrant cette correspondance sont donnés par Young et al. (2018). En effet, les auteurs ont observé que le contraste des particules qu'ils ont étudiées évolue linéairement avec leur masse molaire. Cela signifie que pour caractériser des objets de masse inconnue avec cette technique, il est donc obligatoire de calibrer l'appareil à partir d'objets de masse connue. Le résultat de cette étape réalisée à partir de la BSA est illustré par la Figure G.1 à l'annexe G.

### 5.1.1.2 Conditions expérimentales utilisées

Les expériences de photométrie de masse doivent remplir deux objectifs :

- 1 - Caractériser les objets obtenus grâce à leur masse, et ce dans des conditions de concentration proches de celles utilisées en expérience de microscopie de fluorescence.
- 2 - Évaluer l'influence du marquage sur la taille des particules formées.

En effet, cette technique expérimentale est très bien adaptée à nos besoins puisqu'en plus de fonctionner sans marquage, elle travaille, à l'échelle de la molécule unique, sur une gamme de masses molaires assez large. Par exemple, le modèle *TwoMP* permet de repérer des objets de masse allant de 30 kDa à près de 5 MDa, ce qui est parfaitement adapté aux dimères de CCMV (40 kDa) et aux capsides complètes (4,6 MDa).

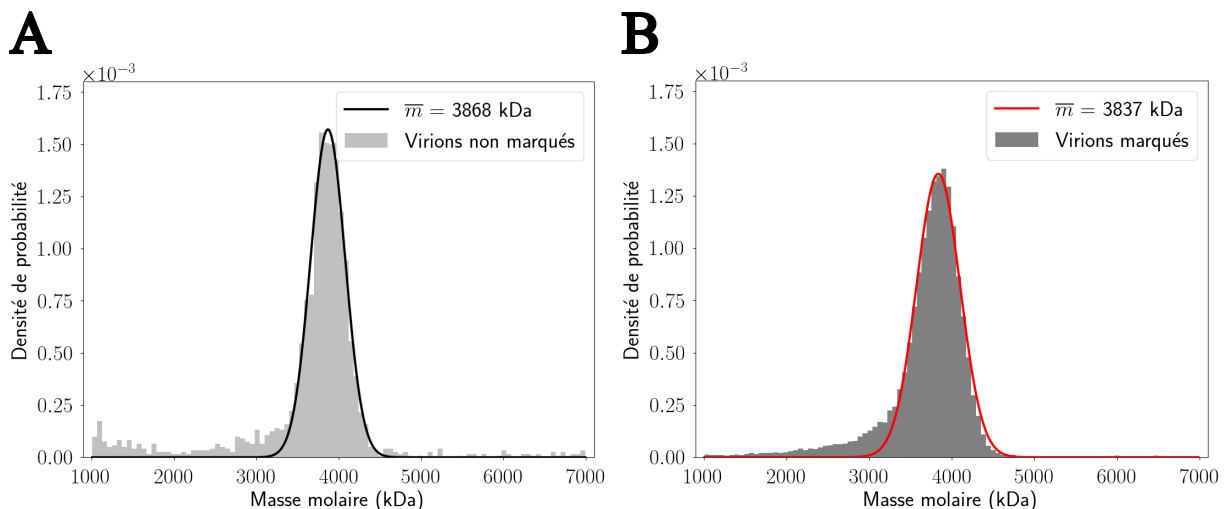
**Dans le but de remplir ces deux objectifs, les sous-unités et les ARN ont été mélangés dans un même tube pour cinq concentrations totales en protéines : 50 nM - 250 nM - 300 nM - 500 nM et 1  $\mu$ M. Dans la majorité des cas, 5 nM en ARN simple brin (sans hybridation) ont été utilisés sauf pour 300 nM en protéines où la concentration en matériel génétique est de 1 nM. Par ailleurs, toutes les dilutions ont eu lieu dans le tampon B' (50 mM MES, 100 mM NaCl à pH 6,1) et le taux de marquage des protéines est de 0 ou 14,2 %. Nous avons donc étudié 10 conditions différentes qui ont toutes été préparées la veille des tests en photométrie de masse. Tous les assemblages se sont déroulés à 4°C, dans un volume de 400  $\mu$ L et pour une durée minimale d'incubation de 15 heures. Enfin, juste avant de caractériser les particules formées, j'ai dû diluer chaque échantillon par un facteur 5 directement sur l'appareil. Cela correspond donc à une concentration finale en ARN, *i.e.* concentration maximale en nucléocapsides, de 200 pM ou 1 nM. Cette nécessité de diluer ainsi que les paramètres d'acquisition sont détaillés à l'annexe G.**

Avant de présenter les résultats expérimentaux, il est intéressant de mettre en évidence les conditions expérimentales qui pourraient favoriser la formation de nucléocapsides complètes. Pour ce faire, je me base sur le critère énoncé par Cadena-Nava et al. (2019) qui stipule qu'un rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$  de 4,3 est nécessaire pour y arriver. En raisonnant à partir de 5 nM en ARN et d'une masse molaire d'environ 890 kDa pour 2774 nucléotides, on trouve une concentration massique d'environ  $4,5 \times 10^{-3}$  g/L en matériel génétique. Avec le rapport en masse de 4,3 cela correspond à environ  $1,9 \times 10^{-3}$  g/L en protéines soit une concentration molaire de près de 940 nM en prenant 20,3 kDa pour la masse molaire d'un monomère. **Cette valeur de 4,3 correspond donc à un rapport en concentration de 188, ce qui est légèrement supérieur au nombre de monomères d'une capsidite complète de CCMV qui est de 180.** En appliquant ce raisonnement aux assemblages réalisés avec 5 nM en ARN et 50 nM, 250 nM, 500 nM et 1  $\mu$ M en protéines, les rapports en masse sont les suivants : 0,2 - 1,1 - 2,3 - 4,6 (respectivement 10 - 50 - 100 et 200 en concentration). Pour 1 nM en matériel génétique et 300 nM en protéines, le rapport en masse est d'environ 6,8 (ou 300 en concentration). **Ainsi,**

deux conditions sur cinq semblent favorables pour former des nucléocapsides complètes.

### 5.1.1.3 Caractérisation de virions natifs

Les premières acquisitions en photométrie de masse nous ont permis de caractériser l'appareil avec des particules stables et connues que sont les dimères de protéines (voir Figure G.2) et les virions natifs (Figure 5.3). Les virions marqués ont été dilués dans le tampon de marquage (pH = 8,2), noté M, alors que pour les particules sans Atto647N, il s'agit du tampon de stockage à pH 4,8 qui est appelé Vir. La concentration finale en objets sur l'appareil est comprise entre 1 et 5 nM. Les histogrammes ont été obtenus à partir de deux acquisitions de deux minutes chacune, réalisées sur des zones différentes.



**FIGURE 5.3 – Photométrie de masse : caractérisation de virions natifs de CCMV.** Distributions des masses molaires de virions natifs A) non marqués et B) marqués. Dans chaque cas, un ajustement gaussien est tracé (courbe noire ou rouge), dont la valeur moyenne est notée  $\bar{m}$ . Les virions non marqués ont été dilués dans le tampon de stockage Vir alors que pour les espèces marquées, il s'agit du tampon de marquage M.

Nous pouvons voir que dans les deux cas, un pic situé autour de 3,9 MDa est observé alors que la masse molaire d'un virion natif est de 4,6 MDa. Deux raisons peuvent être avancées afin d'expliquer cette sous-estimation. La première est que, malgré l'ajout de l'uréase, la calibration effectuée en début de journée reste peu précise dans le cas d'objets de plusieurs mégadaltons. La seconde hypothèse a été avancée par Bertrand Raynal et concerne la taille du virus. En effet, avec 28 - 30 nm de diamètre, il semblerait que les protéines les plus éloignées de la surface contribuent moins que les autres au signal de photométrie de masse, ce qui pourrait entraîner une sous-estimation du contraste total issu d'un virion natif et donc de sa masse molaire. Toutefois, le plus important est que nous puissions caractériser des virions natifs, ce qui nous permet d'avoir une indication concernant la limite supérieure de la masse d'une capsidie assemblée *in vitro*.

Pour conclure, je souhaite préciser que la masse molaire estimée par l'appareil n'est pas celle effective en protéines. En effet, le signal diffusé par la particule provient à la fois des protéines de capsid et de l'ARN qui est encapsulé. **Or, comme la calibration a été effectuée uniquement sur les protéines, les valeurs de masse molaire renvoyées par le logiciel d'analyse correspondent à des masses apparentes en protéines, pour lesquelles une partie du signal collecté est issu de l'ARN.** Nous ne pouvons donc pas directement utiliser les masses molaires calculées à partir de la calibration pour retrouver le nombre de sous-unités par objet. En revanche, en se basant sur la caractérisation des virions natifs, cela ne change pas l'interprétation que nous pouvons avoir sur le fait qu'une capsid soit plus ou moins assemblée.

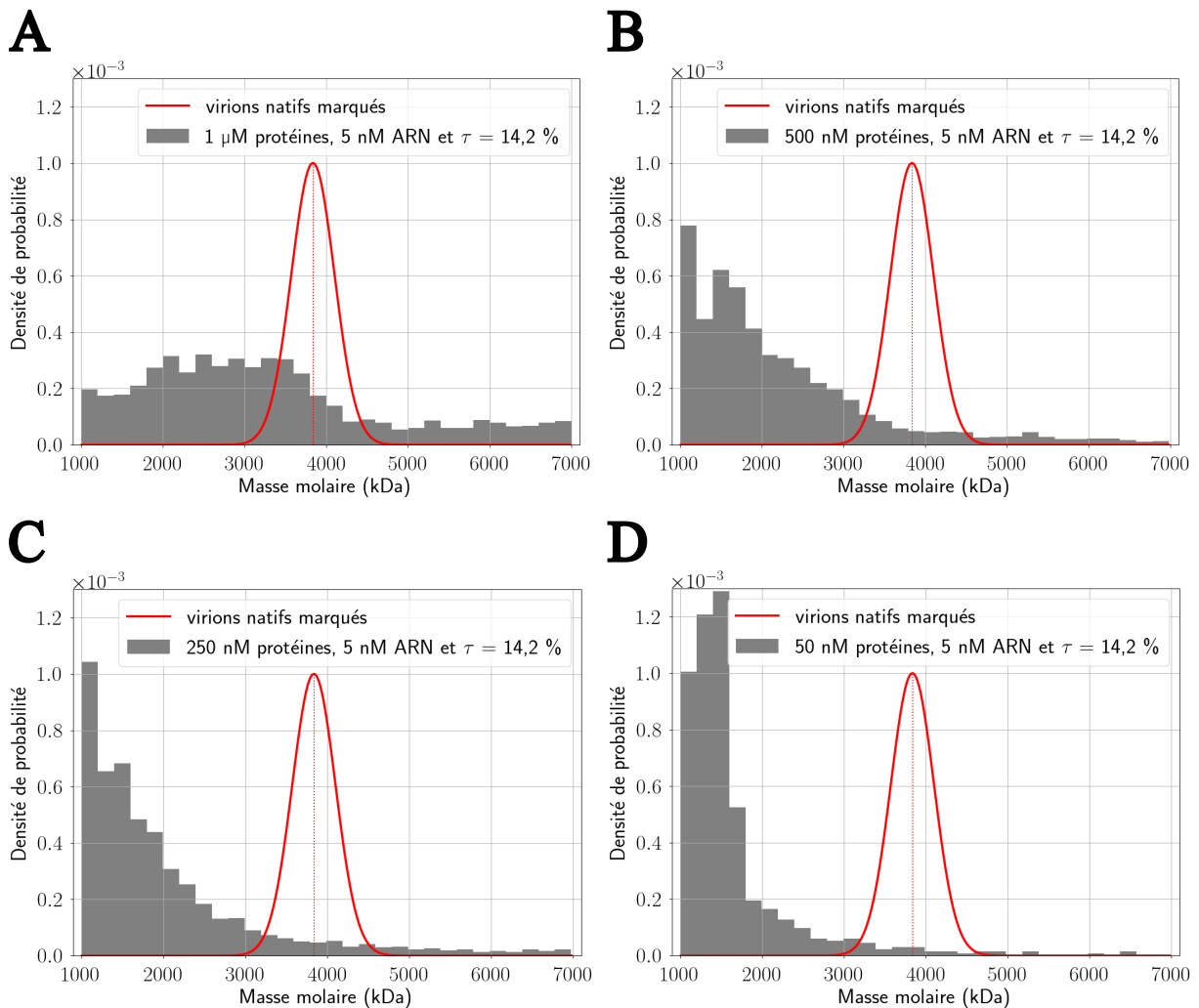
### 5.1.2 Suivi de la taille des objets en fonction de la concentration en sous-unités

Les différents résultats détaillés dans les trois sous-parties suivantes ont été obtenus dans le tampon B', à partir de deux acquisitions de deux minutes chacune et après avoir dilué les échantillons d'un facteur 5. Les deux films ont été enregistrés sur des zones différentes. Si des exceptions sont rencontrées, elles sont mentionnées au cas par cas.

#### 5.1.2.1 À partir de sous-unités marquées

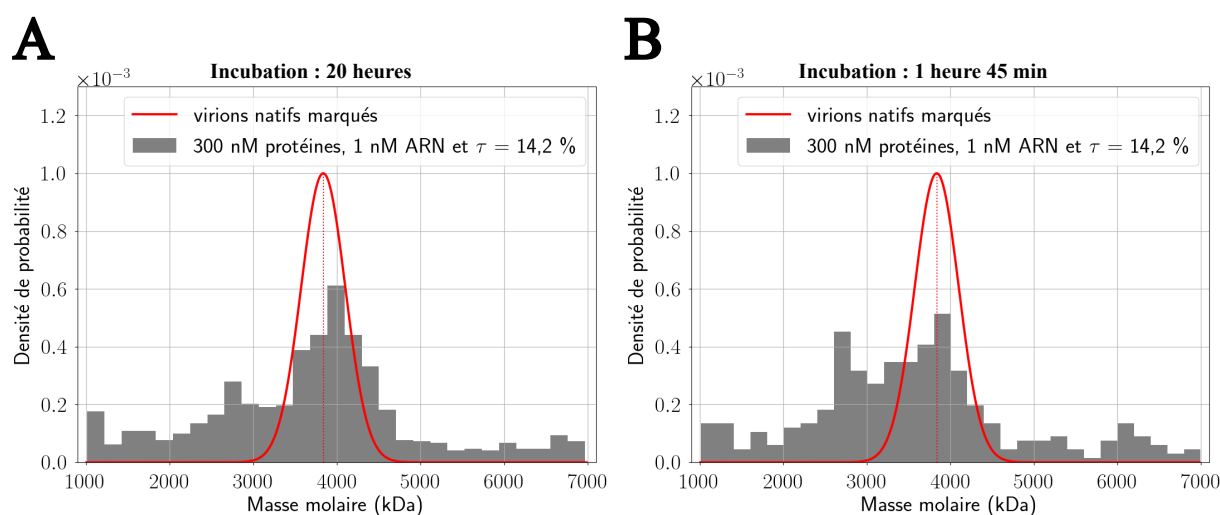
Pour commencer, nous nous intéressons au résultat des assemblages réalisés à partir de sous-unités dont 14,2 % sont marquées. La Figure 5.4 illustre les objets obtenus entre 1 MDa et 7 MDa à partir de 5 nM en ARN et pour des concentrations respectives en protéines de 1  $\mu$ M - 500 nM - 250 nM et 50 nM.

**Avec ces quatre histogrammes, il est clair que plus la concentration en protéines augmente, plus les objets formés ont une masse importante.** En effet, la population entre 1 et 1,5 MDa semble majoritaire pour une concentration en protéines de 50 nM (Figure 5.4.D). À l'inverse, pour 1  $\mu$ M en protéines (Figure 5.4.A), la distribution est plus importante entre 3 et 4 MDa que pour les trois autres concentrations. Je précise que les distributions sont volontairement limitées à 1 MDa pour les petits objets et 7 MDa pour les plus gros afin que les histogrammes soient plus clairs. Les particules de masse inférieure à 1 MDa peuvent représenter aussi bien des dimères libres en solution que des assemblages très préliminaires. Au maximum ces derniers sont formés par une vingtaine de dimères environ, ce qui ne nous intéresse pas forcément dans cette expérience. Au-delà de 7 MDa, les événements sont, en théorie, beaucoup plus rares et peu pertinents pour notre analyse, même si j'explique plus tard que dans certains cas, leur apparition peut provenir de la façon dont les assemblages ont été réalisés.



**FIGURE 5.4 – Photométrie de masse : caractérisation d’assemblages à partir de dimères marqués.** Distributions des masses molaires des objets obtenus à partir de 5 nM en ARN et A) 1  $\mu\text{M}$ , B) 500 nM, C) 250 nM et D) 50 nM en protéines. L’ajustement gaussien qui caractérise la distribution des virions marqués est tracé en rouge et les produits des assemblages sont représentés en gris foncé. Le taux de marquage des protéines est de 14,2 %.

Toujours sur la Figure 5.4, nous pouvons comparer la position de l’ajustement gaussien tracé en rouge, qui est caractéristique des virions marqués, avec les distributions obtenues pour les quatre premières conditions d’assemblage. **Dans tous les cas, nous ne voyons pas de population clairement identifiée dont la masse estimée est centrée sur les 3,9 MDa obtenus pour les virions natifs.** Il est vrai que la condition avec 1  $\mu\text{M}$  en protéines (Figure 5.4.A) est celle où des particules avec des masses élevées sont observées, mais la population majoritaire est plutôt centrée autour de 2,8 - 3 MDa. Ainsi, pour favoriser davantage la formation de capsides entières, une dernière condition a été testée à partir de 300 nM en protéines et seulement 1 nM en ARN au lieu de 5 nM précédemment. Sur la Figure 5.5.A, je compare la distribution de particules obtenues dans ces conditions avec l’ajustement gaussien qui caractérise les virions marqués.



**FIGURE 5.5 – Photométrie de masse : caractérisation de la population obtenue à partir de 300 nM en protéines dont 14,2 % sont marquées.** Distributions des masses molaires des particules obtenues à partir de 1 nM en ARN et 300 nM en protéines pour une durée d’incubation de A) 20h et B) 1h45. L’ajustement gaussien qui caractérise la distribution des virions marqués est tracé en rouge et les produits des assemblages sont représentés en gris foncé. Le taux de marquage des protéines est de 14,2 %.

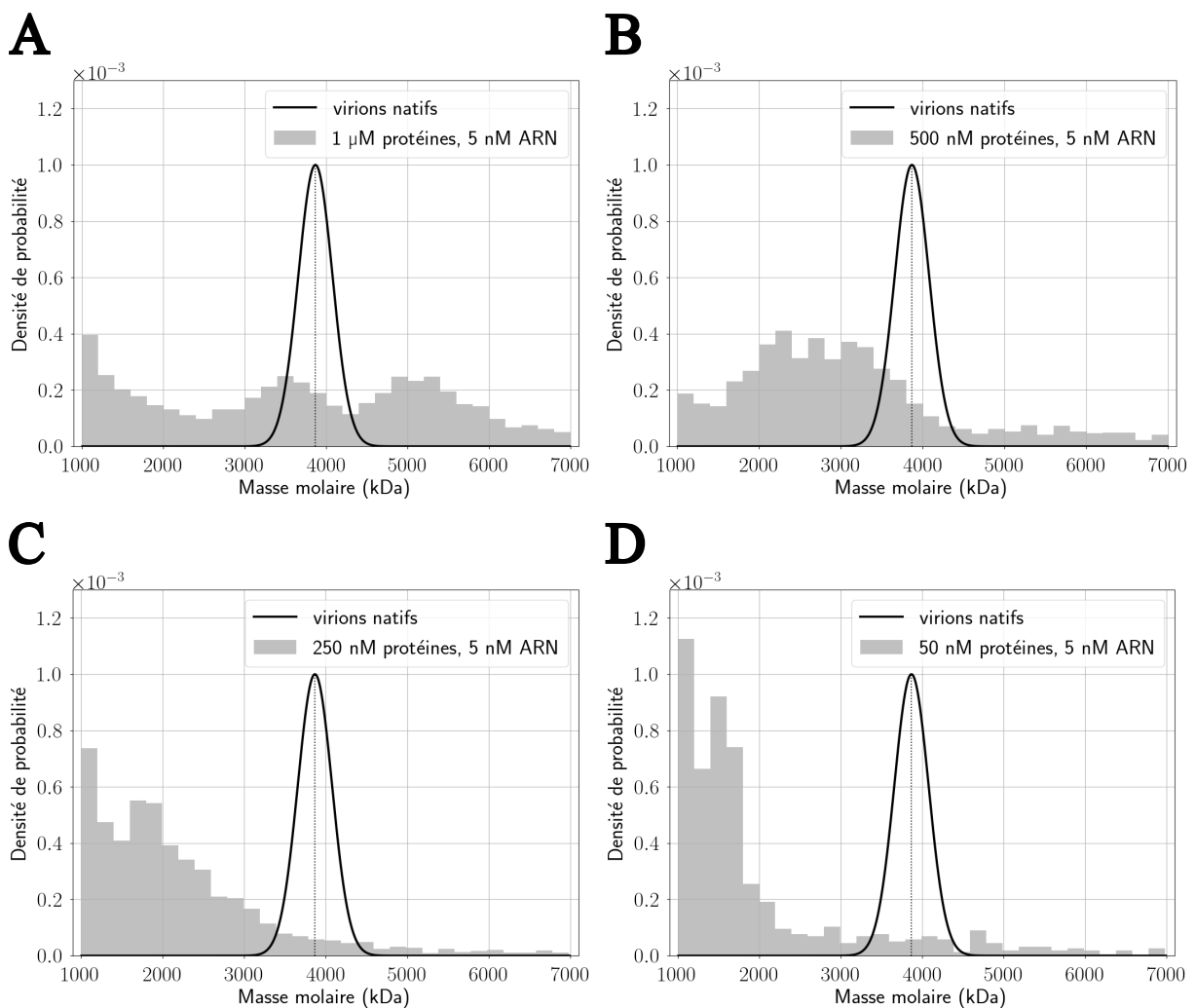
Avec cette nouvelle condition d’assemblage, nous pouvons observer que le pic de la distribution entre 1 MDa et 7 MDa se situe au niveau de celui des virions natifs. Cela confirme bien l’hypothèse émise à la section 5.1.1.2 : pour un assemblage de nucléocapsides, c’est le rapport en masse entre la concentration en protéines et celle en matériel génétique qui est important. **Avec un ratio de 6,8 pour la condition illustrée par la Figure 5.5, nous sommes au-dessus de la valeur de 4,3 proposée par Cadena-Nava et al. (2019) qui favorise la formation de capsides complètes. En travaillant avec 1  $\mu$ M en protéines et 5 nM en ARN, le rapport en masse est de 4,6 et la Figure 5.4.A nous montre bien que des objets de masse similaire aux virions natifs sont obtenus.** En revanche, à la différence de la Figure 5.5.A, la population de particules proches de nucléocapsides n’est pas majoritaire. Il se peut que le rapport limite de 4,3 soit légèrement sous-estimé dans notre cas, mais il est important de préciser que pour obtenir cette valeur, les auteurs ont effectué leurs expériences à pH 4,5 contre 6,1 pour nous. L’utilisation d’un milieu plus acide permet de favoriser l’assemblage de capsides de CCMV.

Par ailleurs, sur la Figure 5.5.B, une seconde distribution est représentée pour 300 nM en protéines, mais avec une durée d’incubation de 1h45 au lieu de 20h. Nous voyons donc que même pour une période d’incubation plus courte, le pic de la distribution des masses entre 1 et 7 MDa est proche de celui des virions. **Ainsi, même si l’histogramme de la Figure 5.5.A a été obtenu après la plus longue durée d’incubation (20 heures contre 15 à 19 heures pour les autres), des capsides de masses similaires aux virions natifs peuvent être formées bien plus rapidement.** Enfin, nous pouvons

remarquer que dans les deux cas, il existe des objets dont la masse estimée par l'appareil est supérieure à 4,5 MDa. Ces particules sont déjà visibles sur la Figure 5.4. Je reviens à la section 5.1.2.3 sur leur origine potentielle.

### 5.1.2.2 Comparaison avec une situation classique, sans marquage

Afin d'avoir une indication sur une potentielle perturbation du mécanisme d'auto-assemblage liée à la présence de l'Atto647N sur certaines protéines, les cinq conditions expérimentales décrites précédemment ont été utilisées pour des sous-unités non marquées. Sur la Figure 5.6, les distributions en masse des quatre assemblages réalisés à partir de 5 nM en ARN sont représentées.



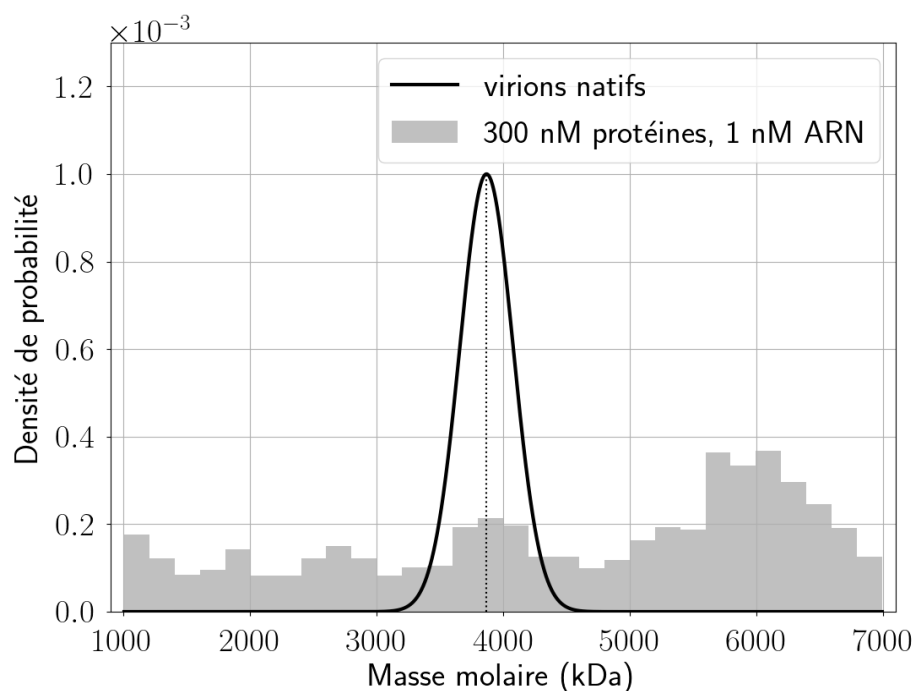
**FIGURE 5.6 – Photométrie de masse : caractérisation d’assemblages à partir de dimères non marqués.** Distributions des masses molaires des objets obtenus à partir de 5 nM en ARN et A) 1  $\mu$ M, B) 500 nM, C) 250 nM et D) 50 nM en protéines. L’ajustement gaussien qui caractérise la distribution des virions natifs est tracé en noir et les produits des assemblages sont représentés en gris clair.

Comme dans le cas où 14,2 % des protéines sont marquées, **nous retrouvons bien**



un décalage progressif vers les masses les plus élevées quand la concentration en sous-unités, et donc le rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$ , augmente. Néanmoins, il semblerait que par rapport à ce qu'il est possible d'observer à la Figure 5.4, les objets obtenus uniquement avec des dimères non marqués ont des masses plus élevées. Cette différence est surtout valable pour 500 nM et 1  $\mu\text{M}$  en protéines. En effet, sur la Figure 5.6.B, la distribution obtenue pour 500 nM en protéines est plus présente entre 3 et 4 MDa que celle à 250 nM. En comparaison, sur la Figure 5.4, nous voyons très peu de différences entre ces deux conditions d'assemblage lorsqu'une partie des sous-unités sont marquées.

De plus, comme en présence de protéines marquées, les distributions des quatre premières conditions d'assemblage sont comparées à la masse mesurée des virions natifs sans Atto647N. Ici, seule une concentration de 1  $\mu\text{M}$  en protéines permet d'observer une population dont la masse mesurée est proche du pic de virons. **En revanche, une deuxième population de particules est présente aux alentours de 5 - 5,2 MDa comme le montre le second pic de distribution visible sur la Figure 5.6.A.** Ces objets, plus gros que des capsides entières, existent également lorsqu'une partie des sous-unités sont marquées, mais la Figure 5.4 ne fait pas état de deux populations bien distinctes autour de 3,6 - 3,8 MDa puis 5 - 5,2 MDa.



**FIGURE 5.7** – Photométrie de masse : caractérisation de la population obtenue à partir de 300 nM en protéines sans marquage. Distribution des masses molaires des objets obtenus à partir de 1 nM en ARN et 300 nM en protéines. L'ajustement gaussien qui caractérise la distribution des virions natifs est tracé en noir et les produits des assemblages sont représentés en gris clair.

Enfin, en ce qui concerne le rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$  le plus élevé, à savoir à partir

de 300 nM en protéines et 1 nM en ARN, l'ajustement gaussien correspond parfaitement au premier pic de la distribution de particules assemblées (Figure 5.7). Dans ces conditions, on retrouve également deux populations bien distinctes, où chacune d'entre elles est encore plus décalée vers des masses élevées que pour 1  $\mu$ M en protéines et 5 nM en ARN. À la section 5.1.2.3, je propose une explication afin de justifier la présence de ces particules plus grosses qu'attendues.

Pour résumer, le Tableau 5.1 compare, en partie, la proportion d'objets ayant une masse comprise entre 3,5 et 4,3 MDa (appelés nucléocapsides) pour les cinq conditions d'assemblage, avec et sans marquage. La population totale considérée dans chaque cas concerne les particules de masse supérieure à 1 MDa. **Alors que la proportion de nucléocapsides augmente avec le rapport en masse quand une partie des sous-unités sont marquées, ce n'est pas vraiment le cas quand le marquage est absent.** En effet, une augmentation cohérente est obtenue de 50 à 500 nM en protéines mais pour les deux rapports en masse les plus élevés, une légère diminution est observée. Dans la section suivante, j'explique que cette baisse se traduit par une forte augmentation du pourcentage d'objets de masse supérieure à 4,3 MDa, qui est donné dans la colonne « super-capsides » du Tableau 5.1. Ces observations sont également valables lorsque l'on regarde l'influence de l'Atto647N : de 50 à 500 nM en protéines le pourcentage de capsides proches de virions est plus grand en absence de marquage, mais la tendance s'inverse pour des rapports en masse de 4,6 et 6,8.

	Rapport en masse	Proportion de nucléocapsides	Proportion de super-capsides	Statistique étudiée
Virions marqués	/	76 %	3,8 %	31214
$C_p = 300$ nM, $C_A = 1$ nM	6,8	32 %	36 %	1119
$C_p = 1$ $\mu$ M, $C_A = 5$ nM	4,6	13 %	31 %	3231
$C_p = 500$ nM, $C_A = 5$ nM	2,3	4,1 %	9,6 %	3397
$C_p = 250$ nM, $C_A = 5$ nM	1,1	3,8 %	8,2 %	2372
$C_p = 50$ nM, $C_A = 5$ nM	0,2	1,9 %	2,1 %	672
Virions	/	75 %	7,3 %	2074
$C_p = 300$ nM, $C_A = 1$ nM	6,8	13 %	62 %	1683
$C_p = 1$ $\mu$ M, $C_A = 5$ nM	4,6	11 %	54 %	3844
$C_p = 500$ nM, $C_A = 5$ nM	2,3	13 %	19 %	3664
$C_p = 250$ nM, $C_A = 5$ nM	1,1	4,5 %	8,9 %	3463
$C_p = 50$ nM, $C_A = 5$ nM	0,2	4,6 %	10 %	808

**TABLEAU 5.1 – Photométrie de masse : particules similaires aux virions natifs avec et sans marquage.**  $C_p$  indique la concentration en protéines et  $C_A$  celle en ARN. La proportion de nucléocapsides correspond aux objets de masse comprise entre 3,5 et 4,3 MDa parmi le nombre d'objets de masse supérieure à 1 MDa (noté "Statistique étudiée"). La proportion de super-capsides, concerne cette fois-ci les objets de masse supérieure à 4,3 MDa. Les résultats en rouge précisent que l'on travaille avec des protéines marquées à hauteur de 14,2 %.

Ainsi, le fait que certains dimères soient marqués n'empêche pas la formation d'objets similaires à des nucléocapsides. Toutefois, en comparant les particules obtenues pour un taux de marquage nul, il semblerait que la présence de l'Atto647N entraîne une différence dans la population d'objets formés. À la Figure G.3 de l'annexe G, je montre qu'en comparant les distributions deux à deux pour chaque rapport de masse  $\frac{\text{protéines}}{\text{ARN}}$ , un certain décalage est observé.

### 5.1.2.3 Justification des objets de masse plus élevée qu'attendue

Dans les deux sections précédentes, j'ai pu présenter les résultats des assemblages réalisés avec et sans marquage pour cinq rapports en masse  $\frac{\text{protéines}}{\text{ARN}}$ . En plus d'une certaine différence dans la proportion d'objets proches de virions natifs selon la présence ou non d'Atto647N, c'est l'apparition de particules plus grosses qu'attendues qui a retenu notre attention. Ces objets ont été retrouvés pour toutes les conditions expérimentales mais c'est surtout avec les deux rapports en masse les plus importants et en absence de marquage qu'une population bien distincte est visible.

Dans le Tableau 5.1, je donne la proportion prise par ces particules de plus de 4,3 MDa parmi toutes celles détectées à partir de 1 MDa dans la colonne « super-capsides ». **Nous voyons clairement que cette population apparaît de manière importante pour des rapports en masse de 4,6 et 6,8 en faveur des protéines et en absence de marquage.** En effet, en se basant seulement sur des assemblages effectués à partir de sous-unités non marquées, elle représente respectivement 54 et 62 % des objets avec une masse supérieure à 1 MDa, contre 19 % pour un rapport en masse de 2,3. En ce qui concerne les virions natifs, nous pouvons voir que très peu de ces particules sont détectées. Par ailleurs, avec marquage, on obtient respectivement 31 et 36 % pour 1  $\mu\text{M}$  et 300 nM en protéines, ce qui n'est pas non plus négligeable. **De manière générale, la proportion d'objets plus gros qu'attendus est plus faible en présence de l'Atto647N quel que soit le rapport en masse considéré.** Ainsi, à 300 nM et 1  $\mu\text{M}$  en protéines sans Atto647N, la prépondérance de tels objets peut expliquer pourquoi le pourcentage de capsides de masse comprise entre 3,5 et 4,3 MDa est plus important avec marquage (Tableau 5.1).

La formation de ces « super-capsides » n'est pas sans rappeler l'étude menée par Garmann et al. (2014) dans laquelle les auteurs ont montré l'importance d'un assemblage en deux étapes afin d'obtenir des capsides entièrement formées et sans défauts. Ce mécanisme est déjà détaillé au sein de la section 2.1.4.2 et le terme « deux étapes » signifie que les conditions de salinité et de pH doivent être modifiées successivement et non simultanément. **À l'inverse, un assemblage où la force ionique est diminuée en même temps que l'acidité devient plus marquée résulte dans l'apparition de nombreux agrégats.** Cette observation a été faite après avoir dialysé sous-unités et ARN

selon un rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$  de 6, pendant six heures, à 4°C et dans un tampon à pH 4,5 et force ionique de 0,1 M.

En revanche, les auteurs ont également montré qu'un mécanisme en une étape menant à des capsides proches de virions natifs est possible en utilisant un tampon avec un pH de 6 tout en conservant les autres paramètres d'assemblage énoncés juste avant. Par ailleurs, en plus de ces capsides entièrement assemblées, ils ont également observé une population d'objets plus gros pouvant s'apparenter à un doublet de capsides. **Selon eux, ces structures sont piégées cinétiquement et une façon de ne pas les obtenir tout en conservant un mécanisme en une étape serait de diminuer la concentration en protéines.** Dans notre cas, les conditions d'assemblage utilisées sont très similaires à celles de Garmann et al. (2014). En l'état, nous ne pouvons pas assurer avec certitude que les « super-capsides » obtenues correspondent à des doublets, mais leur proportion diminue bien avec le rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$ . Ces explications nous permettraient donc de justifier l'observation de cette population d'objets plus massiques.

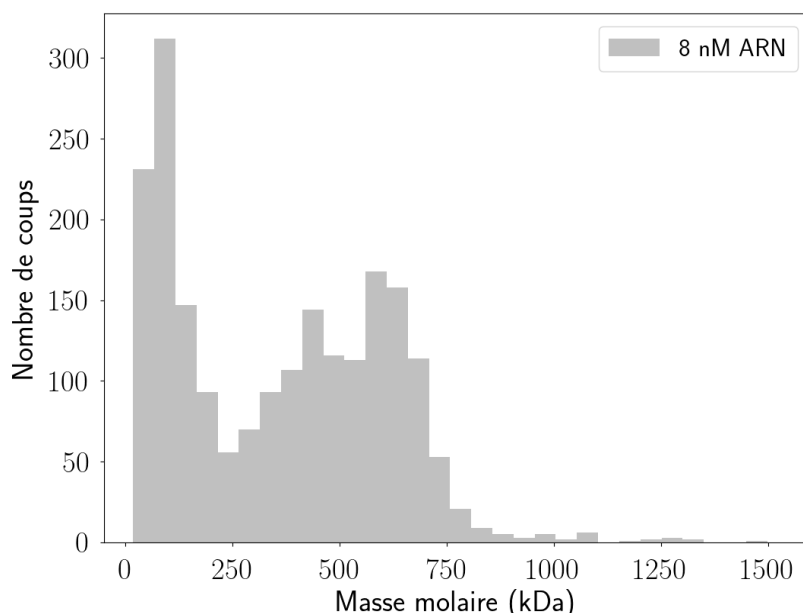
**En ce qui concerne les conditions expérimentales qui mettent en jeu des sous-unités marquées, ces « super-capsides » sont moins présentes. Une hypothèse ici serait que le marquage participe à limiter la formation de ces structures piégées cinétiquement.** Néanmoins, il est difficile de comprendre pourquoi les 14,2 % de dimères qui sont marqués peuvent empêcher la formation de tels objets entre sous-unités non marquées. En effet, pour une concentration totale de 300 nM, le rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$  reste supérieur à 4,3 même en excluant les dimères marqués. Pour apporter une réponse plus définitive sur ce point, il serait intéressant de refaire cette expérience en photométrie de masse et de la coupler à une observation en microscopie électronique. Cette dernière nous permettrait de visualiser les objets obtenus pour chacune des conditions testées.

Pour finir, je précise que les résultats obtenus par Garmann et al. (2014) proviennent de l'assemblage entre des sous-unités de CCMV non marquées et l'ARN de BMV. Dans notre cas, nous utilisons bien l'un des ARN du CCMV mais cette différence ne modifie pas l'interprétation qui vient d'être faite. En effet, je rappelle que Comas-Garcia et al. (2012) ont non seulement montré que des capsides entières et sans défauts peuvent être obtenues à partir de dimères de CCMV autour de l'ARN de BMV, mais surtout que l'assemblage est même plus efficace qu'avec le matériel génétique analogue aux sous-unités.

### 5.1.3 Conclusion

En conclusion, les expériences réalisées en photométrie de masse nous ont permis d'observer qu'en présence ou non de sous-unités marquées, la proportion d'objets ayant une masse proche de virions natifs augmente globalement avec le rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$ . **Il est donc possible de former des particules semblables à des capsides entières**

lorsqu'une partie des sous-unités sont marquées. Néanmoins, elles occupent une proportion plus faible parmi les objets de masse supérieure à 1 MDa que dans le cas où l'Atto647N est absent, sauf quand le rapport en masse devient supérieur à 4,3. De la même manière, le marquage semble réduire la possibilité de former les structures piégées cinétiquement mises en évidence par Garmann et al. (2014) dans le cadre d'un mécanisme d'assemblage en une étape. Par ailleurs, les objets de masse plus élevée que des virions natifs, assimilés à des doublets de capsides par les auteurs pour une solution à pH 6, disparaissent petit à petit quand le rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$  diminue. Il serait peut-être intéressant d'effectuer de nouvelles expériences dans ces conditions afin de confirmer l'hypothèse faite sur l'influence du marquage que ce soit sur la formation de capsides supposées entières ou de structures plus massives. L'utilisation d'un taux de marquage différent nous permettrait également de voir si les différences observées avec un assemblage effectué sans Atto647N sont cohérentes avec la concentration en sous-unités fluorescentes utilisée.



**FIGURE 5.8 – Photométrie de masse : caractérisation de l'ARN utilisé en expérience d'assemblage.** La distribution a été obtenue pour une concentration sur l'appareil de 2 nM dans le tampon B' et à partir d'une seule acquisition de deux minutes.

La photométrie de masse nous a permis de travailler à l'échelle de la molécule unique et donc d'avoir accès à l'hétérogénéité des objets formés pour chacune des conditions testées. Néanmoins, il faut préciser que tous ces résultats ont été obtenus après avoir dilué environ cinq fois chacun des échantillons utilisés. Cette étape est loin d'être anodine pour des objets partiellement assemblés et a pu modifier une partie de l'allure des distributions juste avant que l'acquisition soit réalisée. C'est pourquoi il sera intéressant d'utiliser le mode *Buffer free* qui est disponible sur l'appareil *TwoMP* afin d'éviter la dilution (voir annexe G). De plus, je rappelle que les masses calculées

par le logiciel d’analyse sont des masses apparentes en protéines puisque le signal collecté provient de l’ARN et des sous-unités mais que seule une calibration sur ces dernières a été effectuée. J’ai essayé d’observer la distribution associée à l’ARN en photométrie de masse, mais il est très difficile de ressortir une masse en particulier comme le montre la Figure 5.8. En effet, une très grande plage de masses molaires est détectée par l’appareil et elles sont, pour la grande majorité, inférieures aux 900 kDa attendus pour cet ARN simple brin de 2774 nucléotides.

Enfin, la masse du CCMV fait partie de la limite haute des échantillons correctement caractérisés avec l’appareil *TwoMP* que nous avons utilisé. Pour obtenir une meilleure précision, il pourrait être utile de réaliser de nouveaux tests sur le modèle *SamuxMP* de *Refeyn* qui est optimisé pour l’analyse de virus adéno-associés. En effet, une des forces de cet appareil est de pouvoir différencier des virus avec et sans matériel génétique mais également partiellement remplis. Il serait donc parfaitement adapté à nos expériences.

**Grâce aux différentes observations faites en photométrie de masse, nous avons un point de comparaison sur lequel nous appuyer afin d’interpréter nos résultats obtenus en microscopie de fluorescence. Les informations collectées quant à la taille des objets assemblés selon les conditions expérimentales vont nous servir à avoir un regard critique sur les valeurs de  $\langle n \rangle$  estimées.**

## 5.2 Présentation des conditions expérimentales en microscopie de fluorescence

Après avoir détaillé les différentes composantes essentielles du projet de thèse et les résultats obtenus en photométrie de masse, il est temps de présenter les expériences menées en microscopie de fluorescence sur le mécanisme d’auto-assemblage du CCMV. Il est notamment question de discuter des valeurs de  $\langle n \rangle$  et  $k_{on}$  mesurées expérimentalement à partir de signaux typiques de la dynamique d’assemblage à l’équilibre tels que celui illustré en Figure 3.50.

CATÉGORIE	NOM	COMPOSITION	pH
Chimie de surface	Tampon H1	10 mM HEPES	7,4
	Tampon H2	10 mM HEPES, 150 mM NaCl	7,4
Purification des protéines	Tampon Dim	50 mM HEPES, 500 mM NaCl, 1 mM EDTA (pH = 8), 1 mM DTT et 0,5 mM PMSF	7,5
Préparation d’échantillons	Tampon B	100 mM MES, 10 mM NaCl	6,1
	Tampon B’	50 mM MES, 100 mM NaCl	6,1
	Tampon B’’	50 mM MES, 200 mM NaCl	6,1

**TABLEAU 5.2 – Composition des solutions tampons essentielles aux expériences de molécule unique.**

Pour commencer, je rappelle les paramètres qu'il est important d'avoir en tête au moment d'aborder les sections 5.3 à 5.5. Dans le Tableau 5.2, sont regroupées les principales solutions tampons utilisées au cours des expériences que je décris par la suite. Le tampon H1 ou H2 permet de préparer la chimie de surface alors que la solution notée Dim intervient au cours de la purification de protéines. Le protocole de purification associé est décrit à l'annexe A. Enfin, les différents échantillons d'intérêt tels que les protéines, l'ARN, les marqueurs fluorescents biotinylés ou la NeutrAvidine sont préparés au sein des tampons B, B' ou B''.

En début de chaque expérience de microscopie TIRF, les lamelles de microscope sont fonctionnalisées à l'aide d'une chimie de surface de type PLL-g-PEG et PLL-biotine. L'optimisation de cette dernière est détaillée à la section 3.2. Pour rappel, une bonne limitation des accroches non spécifiques en dimères marqués est atteinte lorsque 0,5 à 1,5 g/L en PLL-g-PEG et  $10^{-4}$  g/L en PLL-biotine sont incubées pendant 30 minutes sur une lamelle de verre. Pour les expériences décrites à la section 5.5, la chimie de surface a été préparée dans le tampon H1, alors que pour celles des sections 5.3 et 5.4 il s'agit de H2. Par la suite, la NeutrAvidine puis les ARN hybridés sont injectés sur les surfaces fonctionnalisées. Les concentrations de ces échantillons dépendent principalement de la couverture en biotine. Pour la NeutrAvidine, les concentrations peuvent valoir 20 - 80 ou 400 nM alors que pour l'ARN, la concentration varie de 50 pM à près de 1 nM. L'important à la fin est d'observer entre 1000 et 2000 sites potentiels d'assemblage sur le champ de vue dans le but d'avoir une bonne statistique de travail.

L'hybridation de l'ARN se fait avec un oligo-biotinylé pour que celui-ci soit fixe sur la lamelle mais également avec un oligo-Atto565 afin de le repérer. La détection des sites potentiels d'assemblage se produit au minimum juste après l'injection de l'ARN et à d'autres moments ultérieurs en fonction de l'expérience réalisée. À la suite du premier film vert, les dimères marqués par un Atto647N sont injectés pour que le mécanisme d'auto-assemblage puisse débiter. La durée d'incubation des sous-unités, dépend de l'objectif de l'expérience, tout comme les conditions d'illumination choisies pour le TIRFM. Pour finir, chaque acquisition enregistrée au cours d'une expérience suit le protocole de traitement détaillé à la section 3.4. Je rappelle que cette étape est nécessaire afin de corriger une éventuelle dérive mécanique mais surtout pour traiter les images (réduction du bruit, détection des particules fluorescentes).

### 5.3 Suivons-nous réellement un assemblage de capsides ?

Pour suivre l'assemblage de capsides virales pleines en microscopie de fluorescence, notre but est d'enregistrer l'intensité de fluorescence émise par des sous-unités marquées au niveau d'un ARN qui possède lui-même un fluorophore. Un exemple de signal d'intérêt correspond notamment à celui observé précédemment en Figure 3.50. Cependant, rien ne

nous assure dans ce cas que nous observons bien des accroches de dimères marqués sur un ARN. En effet, ces variations d'intensité peuvent être issues d'interactions non spécifiques entre les différents éléments nécessaires pour observer l'assemblage de capsides virales sur une lamelle de verre en microscopie de fluorescence. Pour éclaircir ce point, je présente deux exemples d'interactions non spécifiques : entre les sous-unités et la NeutrAvidine puis entre les Atto647N-NHS-Ester et l'ARN accroché à la surface.

### 5.3.1 Les sous-unités marquées interagissent peu avec la NeutrAvidine

Lorsque j'ai détaillé les différentes étapes de caractérisation et d'optimisation de la chimie de surface à la section 3.2, il a été question du rôle non négligeable du couple biotine - NeutrAvidine vis-à-vis des interactions non spécifiques des dimères marqués. En effet, la concentration massique en PLL-biotine a dû passer de 1 g/L à  $10^{-4}$  g/L afin de minimiser les adsorptions non souhaitées en sous-unités fluorescentes. Cependant, ces dernières étant toujours présentes, il est intéressant de regarder si les dimères ont plus tendance à se fixer au couple biotine - NeutrAvidine avec ou sans ARN. **Pour ce faire, le pourcentage d'accroches spécifiques de type ARN - sous-unités marquées a été comparé avec celui caractérisant les interactions non spécifiques entre ces mêmes sous-unités et la NeutrAvidine.** En absence d'ARN hybridés, ce sont des Atto565-biotinylés qui ont été injectés afin de repérer la NeutrAvidine. Le protocole expérimental est décrit ci-dessous :

**1** - Injection des ARN hybridés ou des Atto565-biotinylés, incubation de 2 minutes et rinçage dans le tampon B'.

**2** - Acquisition du film vert.

**3** - Injection des protéines dans le tampon B' : avec ARN, 50 nM en protéines avec un taux de marquage de 10 % mais aussi 200 et 400 nM dont 2,5 % sont marquées. En présence d'Atto565-biotinylés, seulement 50 nM en protéines marquées à hauteur de 10 % ont été injectées.

**4** - Incubation pendant 15 minutes puis second film vert.

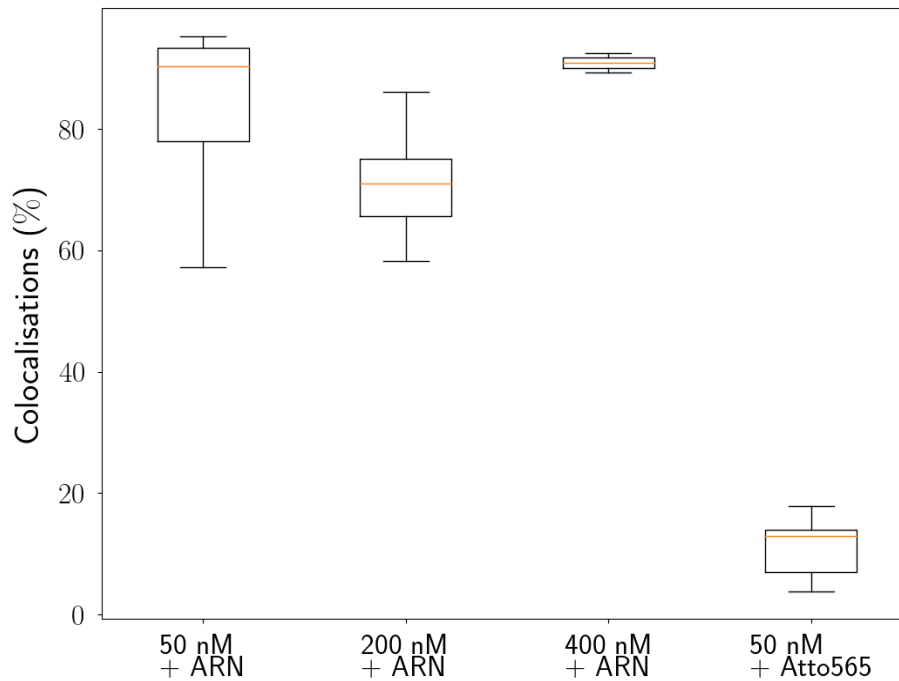
**5** - Acquisition rouge lancée dans la foulée : 1600 images, 100 ms de temps d'exposition et 375 ms de temps de cycle.

**6** - Troisième film vert.

Ici, toutes les étapes de préparation des échantillons et de rinçage ont eu lieu dans le tampon B'. En ce qui concerne l'analyse, les sites d'intérêt excités par le laser vert (Atto565-biotinylés et ARN) sont sélectionnés en colocalisant ceux présents juste avant et juste après l'acquisition dans la voie rouge. En choisissant de procéder ainsi, la statistique disponible diminue mais nous sommes sûrs de travailler à partir d'échantillons qui n'ont pas bougé durant le film de protéines. Dans un second temps, nous regardons si des sous-unités marquées viennent se fixer aux Atto565-biotinylés ou aux ARN. Les



signaux collectés sont fournis à Ruptures et une interaction spécifique ou non spécifique est comptabilisée si à au moins une reprise, l'intensité du signal en marches d'escalier est supérieure à la somme entre la valeur moyenne du fond et l'intensité moyenne d'émission d'un fluorophore  $h$ . Cette valeur nous permet d'accéder au pourcentage de colocalisations en la divisant par le nombre total de sites détectés dans la voie verte. Les différents pourcentages sont regroupés par condition expérimentale sur la Figure 5.9.



**FIGURE 5.9** – Comparaison des interactions spécifiques entre sous-unités marquées et ARN et non spécifiques entre sous-unités marquées et NeutrAvidine. Sur l'axe des abscisses, je donne les concentrations totales en protéines. Pour 50 nM, le taux de marquage est de 9,4 % contre seulement 2,5 % pour 200 et 400 nM.

Les quatre conditions expérimentales ont été analysées à partir de plusieurs cellules micro-fluidiques et même sur différentes journées d'expérience pour certaines. Par ailleurs, la valeur de  $h$  a toujours été estimée dans les mêmes conditions d'illumination. La Figure 5.9 montre que le pourcentage d'interactions spécifiques est 4 à 6 fois supérieur à celui caractérisant les accroches non spécifiques. **En d'autres termes, les sous-unités préfèrent interagir avec l'ARN hybridé plutôt qu'avec un Atto565-biotinylé.** Par ailleurs, avec cette expérience contrôle, nous ne sommes pas en mesure de dire si les dimères s'accrochent non spécifiquement à la NeutrAvidine ou à l'Atto565-biotinylé. **S'il s'agit plutôt de la NeutrAvidine, ce pourcentage devrait être encore plus faible du fait de l'encombrement engendré par la fixation d'un ARN à la surface.** En effet, les sous-unités marquées auraient plus de mal à accéder à la NeutrAvidine, sans interagir avec le matériel génétique en amont. Enfin, dans le tableau 5.3, des informations complémentaires sont précisées par rapport à la comparaison entre inter-

actions spécifiques et non spécifiques. Le pourcentage de colocalisations a été calculé en sommant, pour chaque condition expérimentale, le nombre de sites verts avec au moins une accroche en sous-unité puis en divisant cette valeur par le nombre total d'ARN ou d'Atto565-biotinylés.

Condition expérimentale	$N_{\text{jours}}$	$N_{\text{cellules}}$	$N_{\text{vert}} \pm \text{std}$	$N_{\text{rouge}} \pm \text{std}$	Colocalisations
50 nM + ARN	3	9	$913 \pm 246$	$1937 \pm 381$	80,9 %
200 nM + ARN	2	8	$868 \pm 202$	$2029 \pm 217$	70,9 %
400 nM + ARN	1	3	$813 \pm 234$	$2207 \pm 80$	91,0 %
50 nM + Atto565	2	7	$921 \pm 160$	$1236 \pm 314$	10,8 %

**TABLEAU 5.3 – Analyse quantitative des interactions spécifiques et non spécifiques impliquant les sous-unités marquées.**  $N_{\text{jours}}$  et  $N_{\text{cellules}}$  signifient respectivement "nombre de jours d'expérience" et "nombre de cellules utilisées pour l'analyse".  $N_{\text{vert}}$  indique le nombre moyen d'ARN ou d'Atto565-biotinylés colocalisés entre les deuxième et troisième films acquis dans la voie de détection verte.  $N_{\text{rouge}}$  correspond au nombre moyen de dimères détectés en début de film sur la totalité du champ de vue. Les deux valeurs sont données avec leur écart-type noté std.

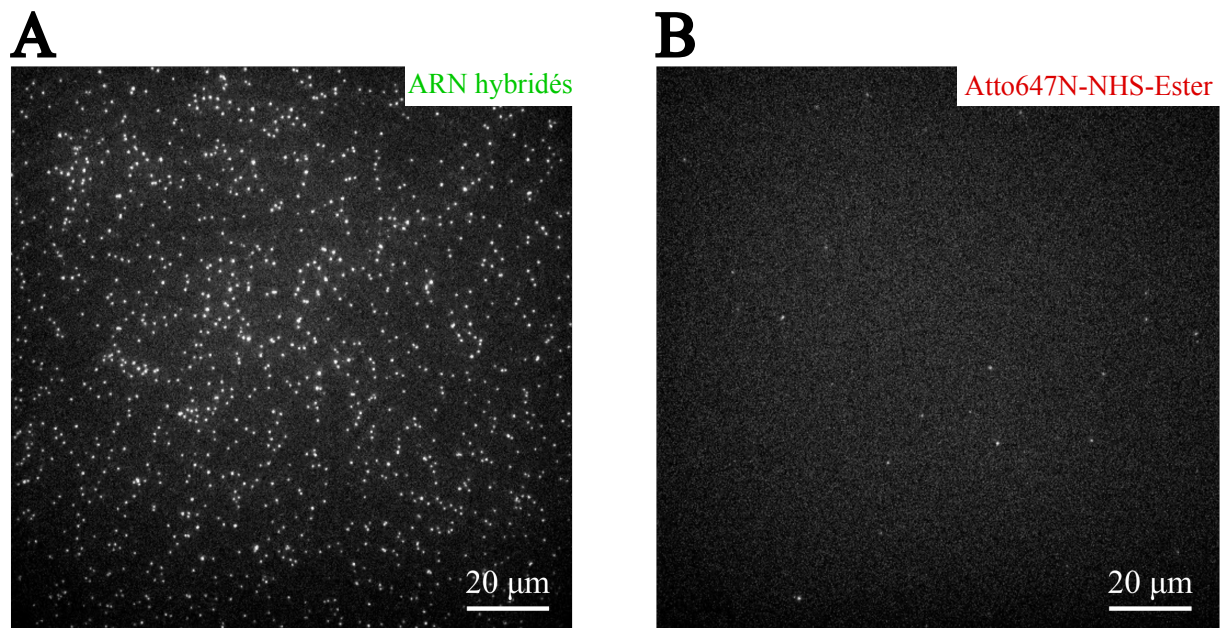
En plus des pourcentages de colocalisations entre sous-unités marquées et ARN ou Atto565-biotinylés, il est intéressant de regarder la statistique en points verts et rouges, notés  $N_{\text{vert}}$  et  $N_{\text{rouge}}$ . En effet, pour un même nombre moyen de particules marquées avec un Atto565, le nombre de dimères présents à la surface en début d'acquisition est plus important d'environ un facteur 2 lorsque l'ARN est présent. Par ailleurs, la statistique en sous-unités marquées est également deux fois supérieure à celle en ARN, mais cela peut s'expliquer par deux raisons. Premièrement, il est possible que certains ARN possèdent l'oligo-biotinylé mais pas celui avec l'Atto565 si l'hybridation est incomplète. Si cela se produit, nous ne pouvons pas voir où vont se fixer ces ARN et s'ils interagissent avec des protéines marquées. De plus, comme nous ne regardons que les sites colocalisés entre les films 2 et 3 réalisés dans la voie de détection verte, les dimères peuvent se fixer sur des ARN visibles au deuxième film mais pas durant le troisième. Dans ce cas, ces sites d'ARN ne font pas partie de la colocalisation effectuées entre les deux derniers films verts. Pour ces deux raisons, la statistique notée  $N_{\text{vert}}$  peut être légèrement sous-estimée.

**Dans tous les cas, voici ce qu'il faut retenir : les interactions non spécifiques entre les sous-unités marquées et la NeutrAvidine (ou les marqueurs fluorescents biotinylés) sont suffisamment faibles pour ne pas être un problème.**

### 5.3.2 Le marqueur Atto647N n'interagit pas avec l'ARN

Les expériences précédentes ne nous assurent pas que ce sont bien les dimères qui vont se fixer sur le matériel génétique grâce à leur bras N-terminal chargé positivement et non les fluorophores Atto647N. De plus, il est probable que certains marqueurs soient libres en solution malgré les différentes dialyses effectuées au cours de la purification des

protéines. Si c'est le cas, peut-être que ces fluorophores peuvent interagir avec l'ARN et donc biaiser l'interprétation faite précédemment. Pour s'assurer du contraire, nous avons adapté le protocole de la section précédente en remplaçant les 5 nM en protéines marquées par la même concentration en Atto647N-NHS-Ester, toujours dans le tampon B'. À part ce changement, toutes les étapes ont été conservées et ce sont quatre cellules micro-fluidiques qui ont été étudiées dans les mêmes conditions. La couverture en ARN au début du deuxième film vert et celle en Atto647N-NHS-Ester à la première image du film rouge sont illustrées respectivement sur les Figures 5.10.A et B.



**FIGURE 5.10 – Interactions non spécifiques entre ARN hybridés et Atto647N-NHS-Ester.** A) Première image du deuxième film vert obtenue 15 minutes après avoir injecté les ARN hybridés. B) Première image du film rouge qui a débuté suite à une incubation de 15 minutes de 5 nM en Atto647N-NHS-Ester. Les deux images correspondent au même champ de vue.

Sur l'image de gauche, nous pouvons observer un peu plus de 900 sites d'ARN hybridés, dont 735 sont encore présents sur le film 3. En revanche, sur l'image de droite, seulement une vingtaine de points sont détectés au début de l'acquisition enregistrée dans la voie rouge, et ce après une période d'incubation des Atto647N de 15 minutes. Lorsque l'on applique la même analyse que dans la section précédente, aucune accroche en fluorophores n'a été détectée sur les 735 sites d'ARN au cours des 1600 images du film rouge. Ainsi, sur cette première cellule, les Atto647N-NHS-Ester ne se fixent pas sur le matériel génétique.

Pour confirmer ce résultat, des mesures similaires ont été faites sur trois autres cellules micro-fluidiques, sauf que le film de 1600 images a été raccourci à 10 images. En effet, si les fluorophores ne s'accrochent pas aux ARN pendant 15 minutes, il y a peu de chance pour que cela intervienne après. Le Tableau 5.4 regroupe les résultats issus des quatre

cellules micro-fluidiques étudiées ici.

Numéro de cellule	$N_{\text{vert}}$	$N_{\text{rouge}}$	Colocalisations
1	735	22	0 %
2	213	11	0 %
3	694	14	0,3 %
4	721	35	0,1 %

**TABLEAU 5.4** – Analyse quantitative des interactions non spécifiques entre ARN et Atto647N-NHS-Ester.  $N_{\text{vert}}$  indique le nombre d'ARN colocalisés entre les deuxième et troisième films acquis dans la voie de détection verte.  $N_{\text{rouge}}$  correspond au nombre d'Atto64N-NHS-Ester détectés en début de film.

Ces résultats supplémentaires nous montrent bien que les Atto647N-NHS-Ester ne se fixent pas aux ARN. En d'autres termes, les interprétations tenues à la section précédente restent valables. Ce sont bien les interactions spécifiques entre sous-unités marquées et ARN qui ont été comptabilisées.

## 5.4 Analyse de la dynamique d'assemblage à l'équilibre

Dans cette quatrième partie, je m'intéresse aux résultats expérimentaux obtenus en microscopie de fluorescence concernant la dynamique d'assemblage à l'équilibre des sous-unités de CCMV. Pour commencer, je décris le protocole utilisé de manière générale ainsi que la façon dont les analyses sont menées. Dans un second temps, je me penche sur la validité du modèle mathématique détaillé à la section 4.2.1 en vérifiant que l'équilibre thermodynamique entre ARN et protéines de capsid est bien atteint, puis en insistant sur les limitations imposées par le photoblanchiment. Enfin, je termine en présentant l'influence de la concentration totale en sous-unités sur les estimations des paramètres d'assemblage que sont  $k_{on}$  et  $\langle n \rangle$ .

### 5.4.1 Du microscope à l'étude des signaux

La mise en place des expériences de dynamique d'assemblage, que ce soit dans le déroulement ou l'analyse, m'a pris un certain temps avant d'arriver à la version utilisée en fin de thèse. Dans cette section, je décris brièvement les principales étapes qu'il faut retenir.

#### 5.4.1.1 Description d'une expérience typique

La chimie de surface est préparée selon le protocole habituel (annexe C) puis la NeutrAvidine est injectée. À ce moment, les échantillons d'ARN et de protéines des solutions mères sont dilués dans le tampon B' aux concentrations souhaitées. Je rappelle qu'initialement le matériel génétique est conservé à 1  $\mu\text{M}$  dans la solution d'hybridation, alors

que les sous-unités sont stockées dans le tampon Dim. **Par ailleurs, les concentrations en protéines étudiées dans cette partie sont les suivantes : 10 et 50 nM pour un taux de marquage de 10 % mais aussi 200 et 400 nM dont 2,5 % sont marquées.** L'ARN est généralement concentré aux alentours de quelques centaines de picomolaires, mais l'objectif est surtout d'obtenir une statistique de 1000 à 2000 points. En considérant qu'au moment d'injecter les protéines, le matériel génétique n'est plus présent dans le volume de la cellule micro-fluidique, il est possible d'estimer grossièrement le rapport en concentration. En effet, à partir de 1500 ARN sur un champ de vue de  $137 \times 137 \mu\text{m}^2$ , on trouve près de  $4,0 \times 10^6$  points sur l'ensemble du canal de  $50 \text{ mm}^2$ . En utilisant la constante d'Avogadro, l'ensemble de ces sites d'ARN correspond environ à  $6,4 \times 10^{-18}$  mol ce qui équivaut à 0,64 pM pour un volume dans la cellule micro-fluidique de 10  $\mu\text{L}$ . Le rapport en concentration  $\frac{\text{protéines}}{\text{ARN}}$  avec 10 nM de protéines est donc de  $1,6 \times 10^4$ , ce qui largement supérieur à la valeur de 188 (4,3 pour un rapport en masse) proposé par Cadena-Nava et al. (2019). **Cela signifie que le ratio entre sous-unités et ARN ne va pas être un facteur limitant pour observer des nucléocapsides. Il s'agira plutôt des concentrations propres en protéines et surtout en ARN, qui peuvent potentiellement défavoriser le processus d'auto-assemblage.**

En ce qui concerne les acquisitions, le protocole mis en place suit le schéma détaillé à la section 5.3.1 pour les tests d'accroches non spécifiques. Je donne de nouveau les principales étapes à retenir ci-dessous :

- 1 - Injection de l'ARN puis incubation de 2 minutes avant rinçage avec 100 à 150  $\mu\text{L}$  de tampon B'.
- 2 - Acquisition du film vert.
- 3 - Injection des protéines dans le tampon B'.
- 4 - Incubation des protéines pendant 15 minutes.
- 5 - Après 14 minutes d'attente, un second film vert est enregistré.
- 6 - À  $t = 15$  minutes, l'acquisition rouge : 1600 images, 100 ms de temps d'exposition et 375 ms de temps de cycle.
- 7 - Pour finir, un troisième film vert est acquis.

Ces étapes sont répétées pour chaque cellule micro-fluidique utilisée au cours de la journée. Je précise que les échantillons d'ARN et de dimères sont préparés depuis leur solution stock respective pour chaque cellule. Un des objectifs est de ne pas travailler trop longtemps avec une solution diluée en ARN puisque cela peut favoriser la perte de l'hybridation. **Pour les dimères, cette précaution a pour but de conserver une concentration en espèce marquée et donc un nombre d'accroches à peu près stable tout au long de la journée.** En effet, les tests de chimie de surface ont mis en avant une diminution progressive du nombre d'adsorptions non spécifiques en sous-unités

marquées si la même solution est conservée toute la journée (3.2.4). Pour finir, au début ou à la fin de chaque journée d'expérience, une cellule est consacrée à la photophysique afin d'avoir une estimation de la hauteur moyenne d'une marche  $h$  ainsi que de la durée moyenne avant photoblanchiment  $\tau_b$ . Ces informations sont obtenues en injectant des Atto647N-biotinylés sur la surface, puis en réalisant la même acquisition dans la voie de détection rouge qu'avec les dimères.

#### 5.4.1.2 Protocole d'analyse mis en place

À la suite des premières étapes de traitement communes à toutes les expériences faites en microscopie TIRF, les deuxième et troisième films d'ARN, sont utilisés conjointement afin de détecter les sites présents durant la totalité du film de protéines. **Les différentes positions enregistrées en sortie de colocalisation nous permettent de savoir à quelles coordonnées du champ de vue regarder pour suivre des interactions entre ARN et dimères.** Ensuite, les intensités de fluorescence provenant des sous-unités marquées sont collectées sur les 1600 images de l'acquisition avant que l'algorithme de détection Ruptures soit appliqué. Une fois les signaux en marches d'escalier obtenus, la mesure des paramètres d'assemblage peut débuter :

- L'information temporelle  $\Delta t_{on}$  est mesurée entre deux événements montants consécutifs sur chaque signal.

- Pour avoir accès à  $n$ , **qui correspond au nombre de sous-unités / dimères marqué(e)s présent(e)s sur l'ARN**, je soustrais la valeur moyenne du bruit de fond à l'intensité de fluorescence du signal de Ruptures détectée à un instant  $t$  : j'obtiens donc une estimation de l'intensité réellement émise par les dimères marqués accrochés à l'ARN. Ensuite, je récupère la partie entière du rapport entre cette valeur d'intensité et la hauteur moyenne d'une marche  $h$ . Cela me permet donc d'obtenir une estimation de  $n$  pour un instant donné  $t$ , et ces étapes peuvent être appliquées à toutes les images de chaque signal. Je rappelle que l'annexe E détaille le protocole d'analyse des expériences de photophysique permettant d'estimer  $h$ .

Une fois que toutes les valeurs de  $\Delta t_{on}$  et  $n$  sont collectées, les distributions sont tracées sur chaque journée d'expérience et comparées respectivement à une loi exponentielle et de Poisson. **Pour finir, je précise que les valeurs ces informations sont mesurées sur les sites d'ARN, colocalisés entre les deuxième et troisième film, qui possèdent au moins une accroche en sous-unités marquées.** Pour satisfaire ce critère, l'analyse détaillée à la section 5.3.1 est appliquée : il faut qu'au moins une fois, l'intensité du signal en marches d'escalier soit supérieure à la somme entre la valeur moyenne du bruit de fond et  $h$ . En procédant ainsi, la statistique étudiée est réduite mais on limite par la même occasion la possibilité de mesurer un nombre de sous-unités marquées de 0. Ces valeurs peuvent être majoritaires lorsque les particules fluorescentes détectées dans la voie verte

sont des oligo-Atto565, avec lesquels les dimères ont peu d'interactions (annexe D). Nous favorisons donc une analyse centrée sur les sites d'ARN réellement accrochés à la lamelle.

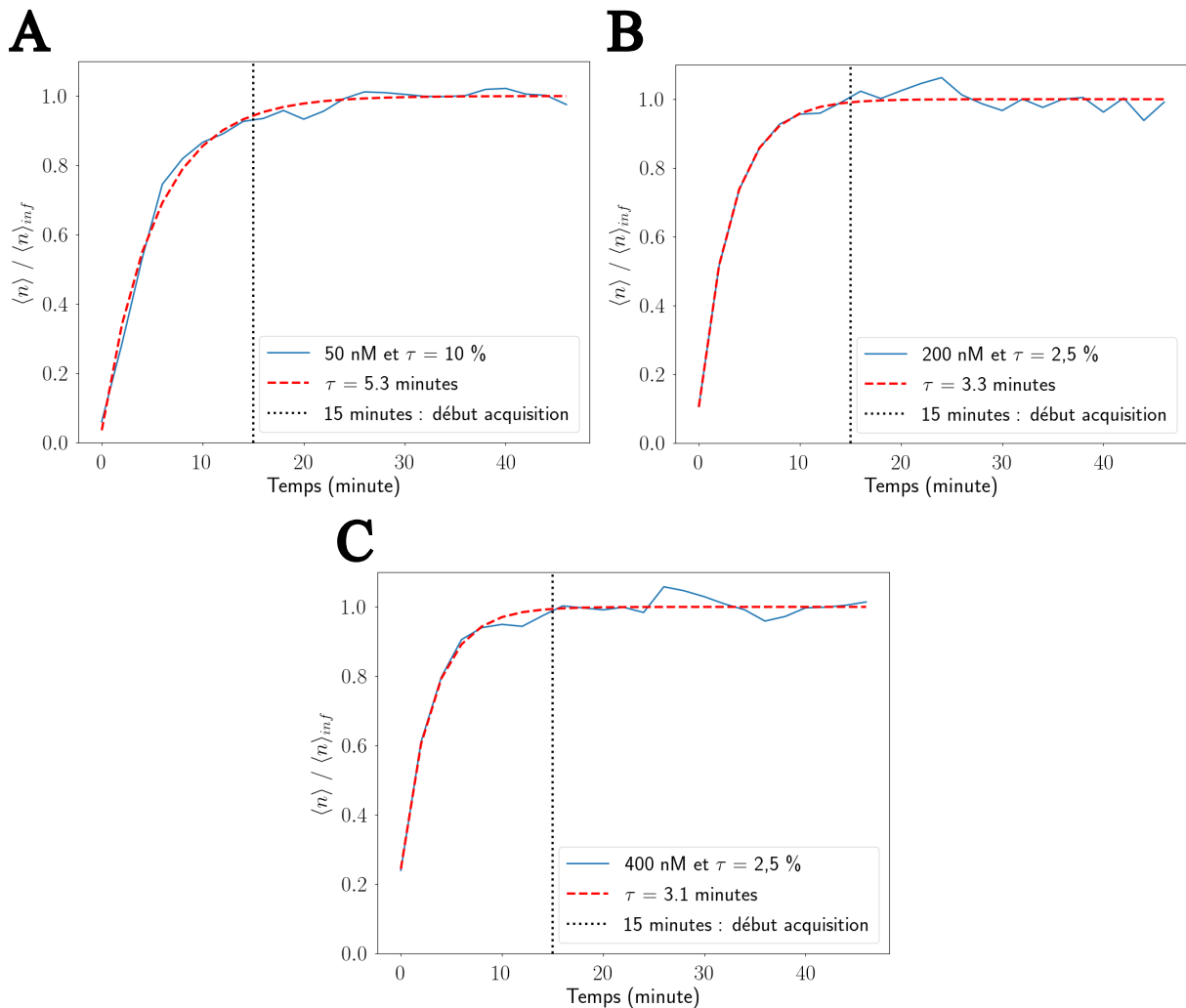
### 5.4.2 Vérification expérimentale de la validité du modèle

À la section 4.2.1, je précise que le modèle mathématique sur lequel reposent les expressions de  $p(n)$  et  $p(\Delta t_{on})$  s'applique uniquement à l'équilibre. Il est donc important de prouver que c'est bien le cas.

#### 5.4.2.1 Preuve de l'équilibre thermodynamique

En commençant l'acquisition quinze minutes après l'injection des protéines, nous nous attendons à ce que le nombre moyen de sous-unités fixées à l'ARN n'évolue plus et que cet état d'équilibre soit maintenu au travers d'échanges de dimères. Or, comme seulement une partie des dimères sont marqués, la taille totale des particules observées est estimée de manière indirecte à partir de deux paramètres : le nombre de sous-unités marquées  $n$  mesurées à un instant  $t$  et le taux de marquage considéré. Ainsi, avec nos expériences, l'équilibre se traduit par un nombre moyen de dimères marqués  $\langle n \rangle$  par ARN constant au cours du temps. Pour nous en assurer, des expériences de contrôle ont été réalisées, au cours desquelles une image toutes les deux minutes a été enregistrée dès l'injection des dimères. Cette modification du protocole, qui remplace les quinze minutes d'attente et le film de 1600 images, nous permet de limiter au maximum l'effet potentiel du photoblanchiment dans l'estimation de  $\langle n \rangle$ .

Sur la Figure 5.11, je montre des résultats typiques de l'évolution de  $\langle n \rangle$  au cours du temps pour une concentration totale en protéines de 50 - 200 - 400 nM et un taux de marquage respectif de 10 - 2,5 et 2,5 %. Dans les trois cas, nous pouvons voir que  $\langle n \rangle$  se stabilise vers une valeur constante au bout d'un certain temps. Pour 200 et 400 nM, le passage vers le régime permanent est atteint avant quinze minutes après injection alors que pour 50 nM, nous sommes à la toute fin du régime transitoire. Les constantes de temps  $\tau$  ont toutes été obtenues à partir d'un ajustement exponentiel qui se base sur l'expression suivante :  $\langle n \rangle(t) = A \times \left(1 - e^{-\frac{t}{\tau}}\right) + B$ .



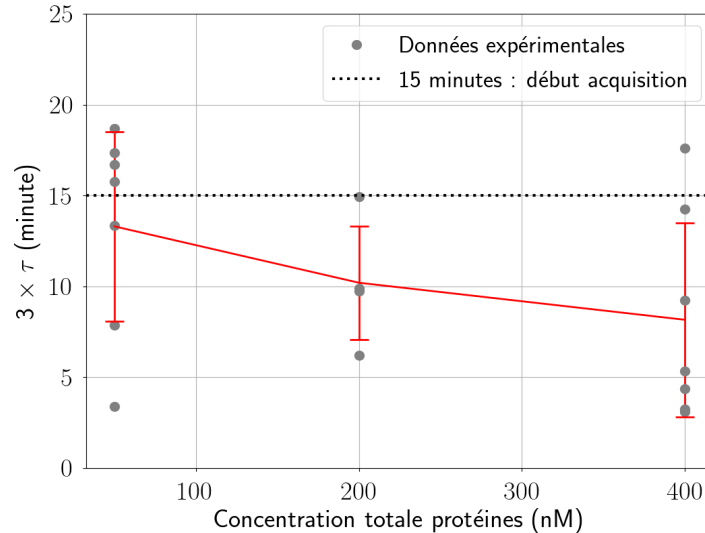
**FIGURE 5.11 – Dynamique à l'équilibre : observation expérimentale de l'atteinte de l'équilibre thermodynamique.** Évolution du nombre moyen de sous-unités marquées  $\langle n \rangle$  par ARN pour une concentration totale en protéines de A) 50, B) 200 et C) 400 nM. Le paramètre  $\langle n \rangle_{inf}$  indique la valeur limite du modèle exponentiel. Les taux de marquage respectifs sont de 10 - 2,5 et 2,5 %. Les données expérimentales sont en bleu, les ajustements exponentiels en pointillés rouges et les constantes de temps sont notées  $\tau$ . Les pointillés noirs à 15 minutes indiquent le début des acquisitions pour les expériences à l'équilibre.

Pour résumer l'ensemble des expériences menées selon ce protocole, la Figure 5.12 montre l'évolution du temps de réponse, noté  $3\tau$ , en fonction de la concentration en protéines. Ce paramètre correspond à la durée pour laquelle  $\langle n \rangle$  est à 95 % de sa valeur d'équilibre. Malgré une certaine dispersion observée pour chaque condition, le temps de réponse diminue légèrement quand la concentration en sous-unités augmente. Dans les trois cas, la valeur moyenne de  $3\tau$  est inférieure à quinze minutes, même si pour 50 nM en protéines il semblerait que nous soyons à la limite entre la fin du régime transitoire et le début du régime permanent.

**En somme, ces expériences nous ont montré que l'équilibre thermody-**



namique est atteint après 15 minutes, lorsque l'acquisition réalisée dans la voie de détection rouge débute. Cette observation confirme que durant le film de dimères, la taille des capsides partiellement ou complètement assemblées n'évolue plus dans le temps.



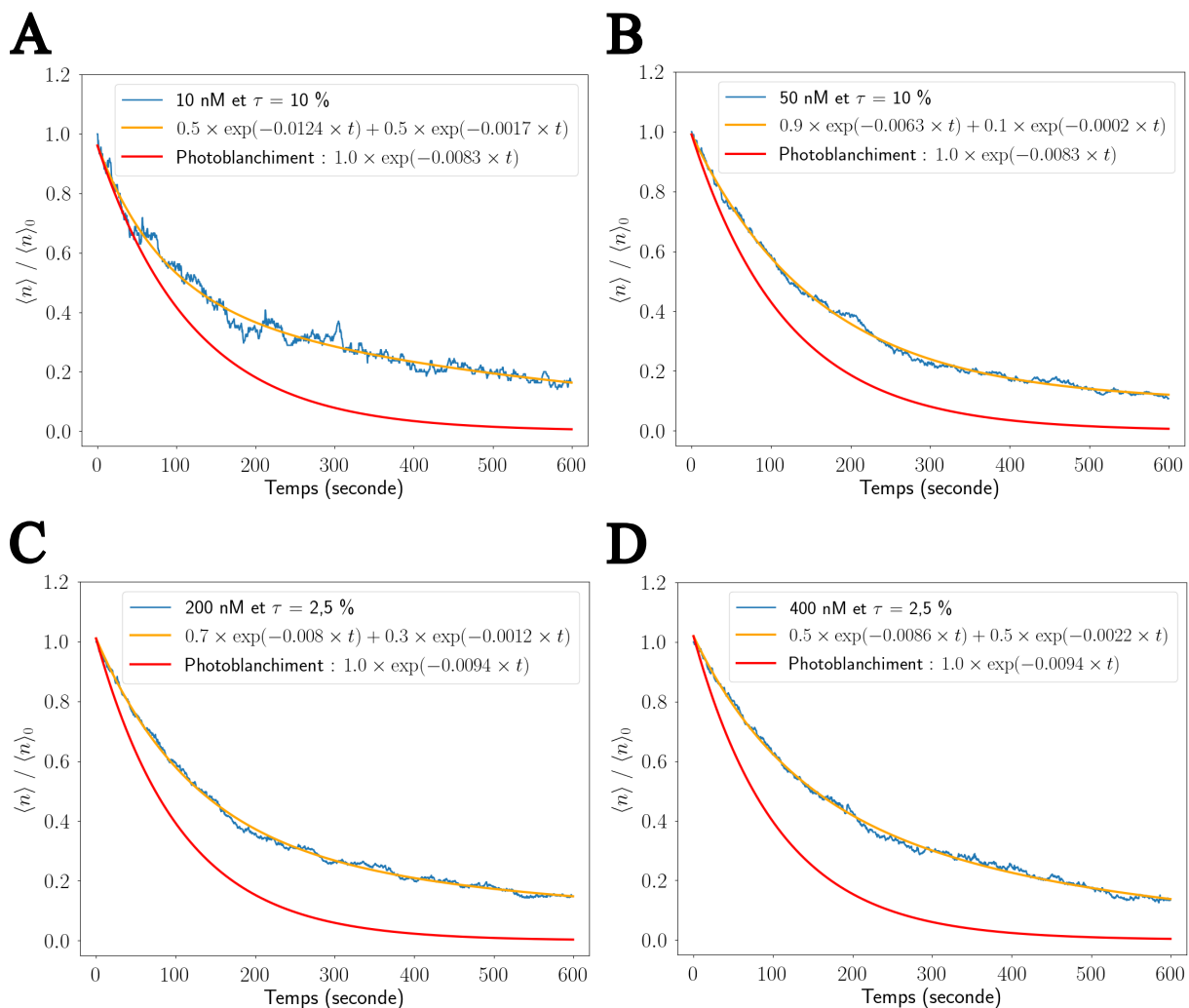
**FIGURE 5.12** – Dynamique à l'équilibre : influence de la concentration totale en protéines sur l'apparition de l'équilibre thermodynamique. Les traits verticaux rouges représentent l'écart-type autour des valeurs moyennes de  $3\tau$ . Les pointillés noirs indiquent le début des acquisitions dans la voie de détection rouge.

#### 5.4.2.2 Influence du photoblanchiment sur l'estimation de $\langle n \rangle$

En plus d'être une preuve de l'atteinte de l'équilibre thermodynamique, les expériences précédentes montrent également que nous pouvons bien suivre l'évolution de  $\langle n \rangle$  au cours du temps. La Figure 5.13 met elle aussi en valeur cette aptitude de notre système en conditions réelles de suivi de la dynamique d'assemblage. Ici, les acquisitions débutent 15 minutes après l'injection des dimères et durent dix minutes avec une image prise toutes les 375 ms.

**Quelle que soit la concentration en protéines, le nombre moyen de sous-unités marquées diminue clairement au cours du temps.** Cette évolution va à l'encontre de ce que nous nous attendions à observer en travaillant à l'équilibre thermodynamique, mais elle s'explique simplement par la présence du photoblanchiment. En effet, à partir du moment où le cycle d'illumination débute, les marqueurs fluorescents peuvent photoblanchir à tout moment, ce que l'expression de la constante  $k_{off}$  (section 4.2.1) prend bien en compte. Par ailleurs, sur la Figure 5.13, nous pouvons voir que la diminution de  $\langle n \rangle$  est plus lente que l'évolution caractéristique du photoblanchiment. Ce décalage peut en partie s'expliquer par une probable surestimation de la constante de photoblanchiment  $k_b$ . En effet, les ajustements exponentiels tracés en rouge sont obtenus à partir d'un contrôle photophysique réalisé à partir d'Atto647N-biotinylés directement

fixés à la surface. Or, durant une expérience d'assemblage, les sous-unités vont s'accrocher à différents endroits de l'ARN et donc être plus ou moins éloignées de la surface (environ 40-50 nm si l'on prend en compte la séquence hybridée et la taille attendue pour une capside de CCMV). Or, comme l'intensité de l'onde évanescente décroît exponentiellement, cette dernière est donc plus faible qu'à la surface, ce qui va probablement ralentir l'apparition du photoblanchiment, et donc la décroissance de  $\langle n \rangle$ . Enfin, en orange je montre que l'allure de l'évolution de  $\langle n \rangle$  est mieux caractérisée par un ajustement en double exponentielle. Il semble néanmoins difficile de tirer une tendance des deux constantes de temps obtenues pour chaque concentration.



**FIGURE 5.13 – Dynamique à l'équilibre : influence du photoblanchiment sur l'estimation de  $\langle n \rangle$  au cours d'une expérience.** Évolution du nombre moyen de sous-unités marquées  $\langle n \rangle$  par ARN pour une concentration totale en protéines de A) 10, B) 50, C) 200 et D) 400 nM. Le paramètre  $\langle n \rangle_0$  indique la valeur initiale du nombre moyen de sous-unités marquées. Les données expérimentales sont en bleu, les ajustements exponentiels caractéristiques du photoblanchiment en rouge et ceux en double exponentielle en orange.

Ces observations ne remettent pas en cause l'équilibre thermodynamique mis en lu-

mière à la section précédente puisque le photoblanchiment a été réduit le plus possible en prenant une image toutes les deux minutes. De plus l'expression de  $p(\Delta t_{on})$  est toujours valable puisqu'elle traduit la distribution attendue de l'écart entre deux instants d'arrivée consécutifs, ce que le photoblanchiment ne va pas affecter. En revanche pour  $p(n)$ , la situation n'est pas aussi simple. En effet, même si l'équilibre thermodynamique, nécessaire pour obtenir la loi de Poisson décrite à la section 4.2.1, est avéré, l'estimation de  $n$  à chaque instant est faussée à cause du photoblanchiment. **Ce mécanisme fait que d'un point de vue photophysique, nous ne sommes pas à l'équilibre, ce qui rend caduque l'expression de  $p(n)$ .** Nous ne pouvons pas utiliser cette relation afin de traduire l'allure de la distribution de  $n$  à chaque instant. **En revanche, il nous est possible de l'appliquer en tout début de chaque acquisition, là où l'équilibre thermodynamique est établi et le photoblanchiment négligeable.** Ainsi, l'estimation de la taille moyenne des particules observées peut se faire en calculant la valeur de  $n$  à la première image *i.e.* à partir de la valeur d'intensité du premier plateau repéré par Ruptures.

Pour conclure, je rappelle les deux modèles utilisés afin de caractériser les distributions de  $\Delta t_{on}$  et  $n$  où l'expression de  $k_u$  est corrigée :

$$p(\Delta t_{on}) = k_{on} e^{-k_{on} \Delta t_{on}} \quad p(n) = \frac{\mu^n}{n!} e^{-\mu}$$

où  $k_{on}$  est la constante d'association et  $\mu = \langle n \rangle = \frac{k_{on}}{k_{off}}$ . Dans la définition de  $\mu$ , on retrouve la constante de temps des marches descendantes  $k_{off}$  qui s'exprime selon la relation suivante :  $k_{off} = k_u$  avec  $k_u = \frac{1}{\tau_u}$  où  $\tau_u$  est le temps de résidence moyen des protéines sur l'ARN. **L'expression de  $k_{off}$ , n'est valable qu'au début de chaque acquisition rouge, avant que le photoblanchiment ne vienne rompre l'équilibre concernant l'évolution de  $n$ .**

### 5.4.3 Effet de la concentration en protéines sur la dynamique d'assemblage

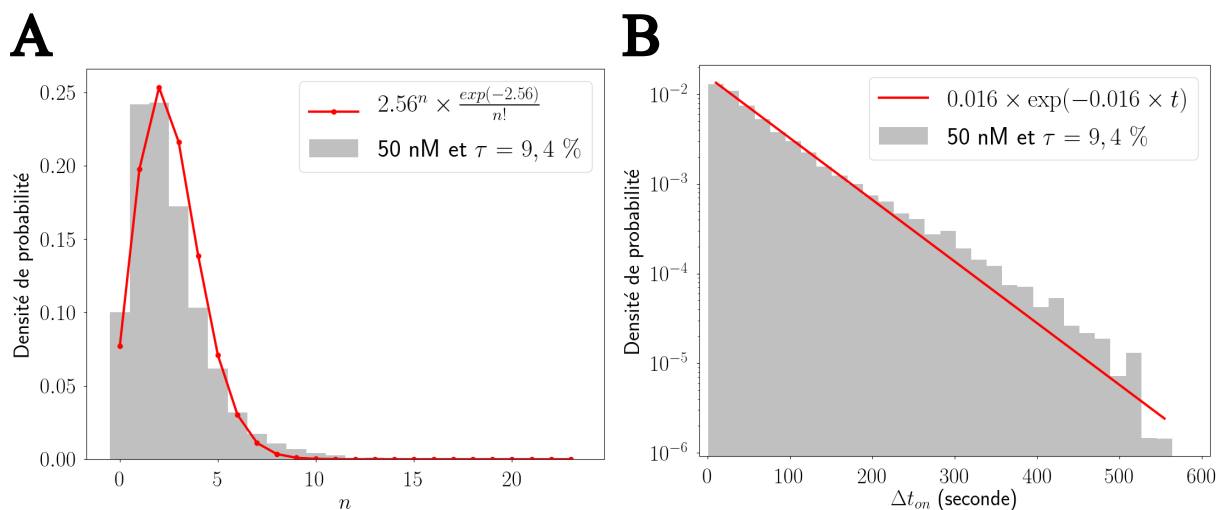
Les expériences de contrôle et corrections précisées dans les deux sections précédentes nous ont permis d'obtenir les résultats que je présente dans cette partie.

#### 5.4.3.1 Étude à concentration constante

Les expériences de dynamique d'assemblage ont été optimisées à partir d'une concentration totale en protéines de 50 nM et un taux de marquage d'environ 10 %. Je commence donc par présenter les résultats obtenus avec ces conditions expérimentales.

Les distributions respectives de  $n$  et  $\Delta t_{on}$  sont illustrées par les Figures 5.14.A et B. Ces deux paramètres d'assemblage ont été mesurés sur plusieurs cellules micro-fluidiques

et journées d'expériences, mais pour des protéines issues de la même purification. En ce qui concerne  $n$ , l'allure de la distribution obtenue expérimentalement correspond bien à la loi de Poisson tracée en rouge. En travaillant à l'échelle de la molécule unique, nous avons bien accès à la totalité des sites potentiels d'assemblage, ce qui nous permet d'affirmer que dans ces conditions expérimentales, les particules observées sont globalement assez éloignées d'un assemblage complet. En effet, pour une concentration en protéines de 50 nM, le nombre moyen de sous-unités marquées par ARN  $\langle n \rangle$  vaut environ 2,6. **Or, en connaissant le taux de marquage, qui est de 9,4 %, nous pouvons estimer le nombre moyen de sous-unités  $\langle n_{tot} \rangle$  par ARN à 27, soit près d'un tiers du nombre total de dimères présents au sein d'une capsid complète.** Même si les conditions d'assemblage sont différentes, ce résultat est cohérent avec ce qui a été observé en photométrie de masse à la Figure 5.4.D. Des capsides complètes sont difficiles à obtenir mais des particules ayant une masse proche de 1 - 1,2 MDa (25 à 30 dimères) peuvent être formées à partir de 50 nM en protéines.

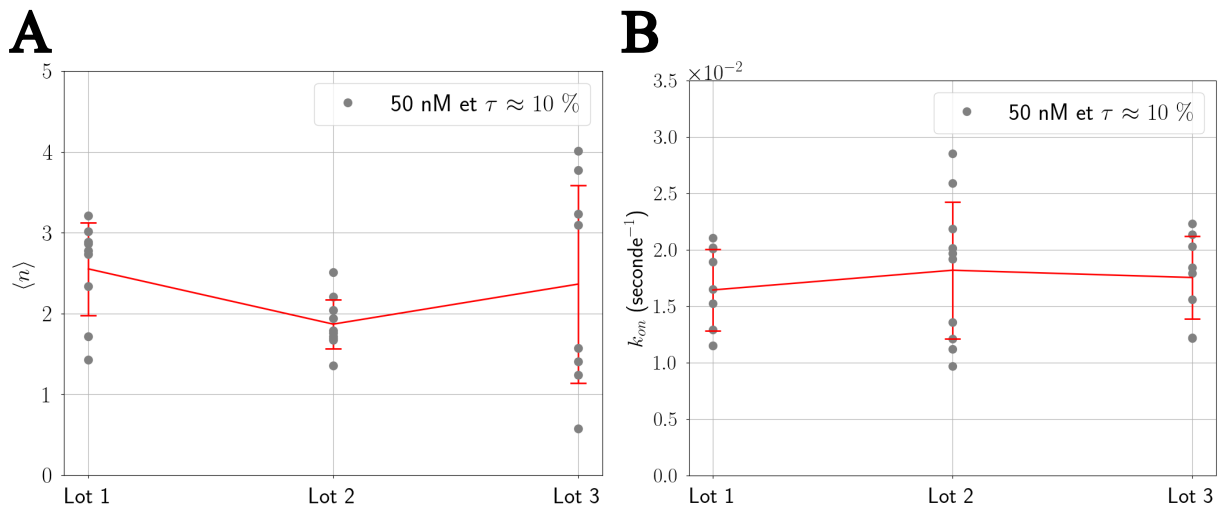


**FIGURE 5.14 – Dynamique à l'équilibre : distributions des paramètres d'assemblage mesurés expérimentalement pour 50 nM en protéines.** Histogrammes A) du nombre de sous-unités marquées  $n$  estimées en début d'acquisition et B) de la durée entre deux événements d'accroche consécutifs  $\Delta t_{on}$ . Les ajustements par une loi de Poisson et exponentielle sont tracés en rouge. Le taux de marquage est de 9,4 %.

Pour  $\Delta t_{on}$ , nous voyons sur la Figure 5.14.B que l'allure de la distribution est globalement très proche de la loi exponentielle attendue, malgré un léger décalage pour les durées plus longues. En s'appuyant sur l'ajustement exponentiel tracé en rouge, une constante d'association  $k_{on}$  de 0,016 s<sup>-1</sup> peut être estimée, ce qui correspond à une durée moyenne entre deux accroches en dimères marqués de près d'une minute. **De plus, en utilisant une nouvelle fois le taux de marquage, on obtient une constante d'association totale  $k_{on}^{tot}$  de 0,17 s<sup>-1</sup> soit une accroche en sous-unités toutes les six secondes.**

Ainsi, à partir de la relation  $\langle n \rangle = \frac{k_{on}}{k_u}$  et des résultats expérimentaux obtenus pour

$\langle n \rangle$  et  $k_{on}$ , on obtient  $k_u = 5,95 \times 10^{-3} \text{ s}^{-1}$ . **Cela correspond à un temps de résidence moyen des sous-unités de CCMV sur l'ARN de près de 168 secondes pour une concentration totale en protéines de 50 nM.** Ce résultat est valable pour le premier lot de protéines avec lequel j'ai travaillé sur ces expériences de dynamique d'assemblage. La Figure 5.15 nous montre que les mesures sont assez reproductibles d'un lot de protéines à un autre. En effet, à partir du nombre moyen de sous-unités marquées  $\langle n \rangle$  qui est représenté sur la Figure 5.15.A pour chacun des trois lots de protéines utilisés, on trouve un nombre moyen de sous-unités  $\langle n_{tot} \rangle$  par ARN de 27 - 19 et 24. De la même façon, nous pouvons voir que la constante d'association  $k_{on}$  est similaire d'un lot à un autre avec une valeur comprise entre 0,016 et 0,018  $\text{s}^{-1}$ . **En combinant toutes ces informations, on trouve respectivement un temps de résidence moyen des sous-unités par ARN pour 50 nM en protéines de près de 168 - 120 et 128 secondes respectivement pour les lots 1, 2 et 3.** Globalement les résultats sont donc cohérents d'un lot à un autre, malgré une certaine dispersion qu'il est possible d'observer, notamment dans la mesure de  $\langle n \rangle$  pour le lot 3. Ces différences d'une cellule à une autre peuvent s'expliquer, entre autres, par une chimie de surface moins efficace, favorisant plus ou moins l'apparition d'accroches non spécifiques ou une moins bonne estimation des paramètres photophysiques lors de l'expérience de contrôle.

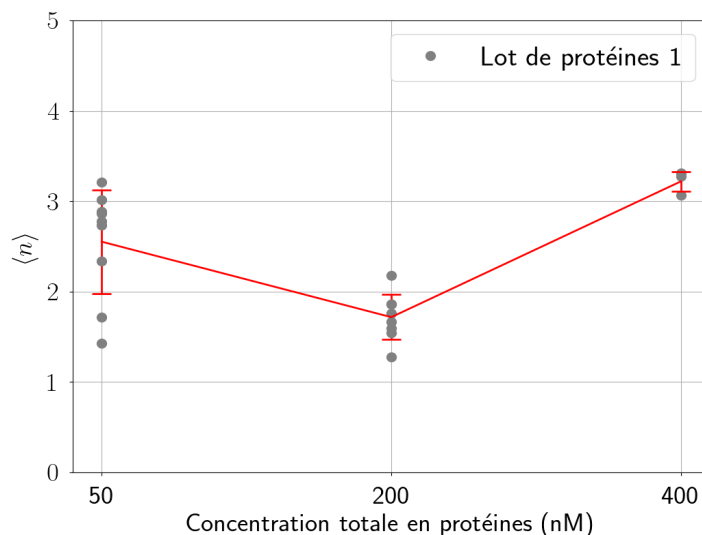


**FIGURE 5.15 – Dynamique à l'équilibre : caractérisation de la reproductibilité des mesures pour 50 nM en protéines.** Influence du lot de protéines sur l'estimation expérimentale de A)  $\langle n \rangle$  et B)  $k_{on}$ . Chaque disque gris représente le résultat d'une cellule micro-fluidique. Les traits verticaux rouges représentent l'écart-type autour des valeurs moyennes de  $\langle n \rangle$  et  $k_{on}$  calculées à chaque cellule micro-fluidique.

### 5.4.3.2 Influence de la concentration totale en protéines

L'étude de la dynamique d'assemblage a également été réalisée pour d'autres concentrations en protéines virales. Sur la Figure 5.16, l'évolution du nombre moyen de sous-unités marquées par ARN est donnée en fonction de la concentration totale  $c$  en protéines.

Les valeurs prises par  $\langle n \rangle$  sont d'environ 2,6 - 1,8 et 3,2 sous-unités marquées pour respectivement 50, 200 et 400 nM en protéines. En prenant en compte le taux de marquage qui est de 9,4 % pour la concentration la plus faible et de 2,5 % pour les deux autres, on retrouve un nombre total de dimères par ARN qui vaut en moyenne 27 - 71 et 129. **Le premier enseignement est que la taille moyenne des objets observés augmente bien avec la concentration en protéines.** Cette tendance a également été observée en photométrie de masse comme nous avons pu le voir à la section 5.1.2.

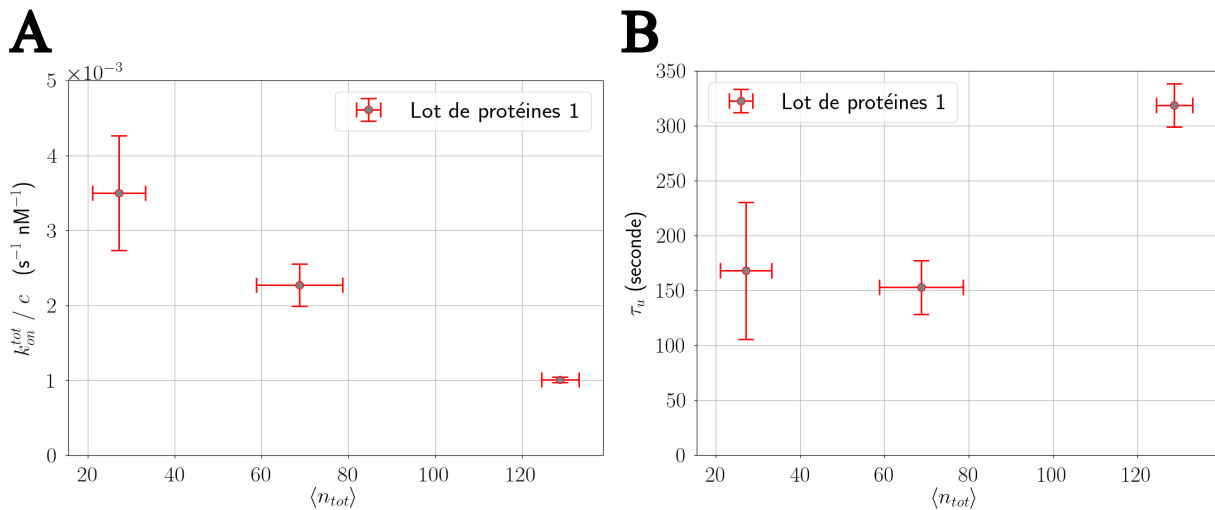


**FIGURE 5.16 – Dynamique à l'équilibre : influence de la concentration totale en protéines sur  $\langle n \rangle$ .** Chaque disque gris représente le résultat d'une cellule micro-fluidique. Les traits verticaux rouges représentent l'écart-type autour des valeurs moyennes de  $\langle n \rangle$  calculées à chaque cellule micro-fluidique. Ces résultats sont donnés pour un seul lot de protéines. Pour 50 nM le taux de marquage est de 9,4 % alors que pour 200 et 400 nM, il est de 2,5 %.

Cependant, même si la taille moyenne des capsides évolue dans la bonne direction, on constate une importante surestimation pour 400 nM. En effet, près de 130 dimères sont comptabilisés en moyenne par ARN alors qu'une capside complète n'en possède que 90. Il est difficile d'associer cette différence uniquement à ce que Garmann et al. (2014) ont observé pour des rapports en masse trop élevés et un assemblage réalisé en une seule étape. Nous avons effectivement observé des objets bien plus gros qu'une capside entièrement assemblée en photométrie de masse mais en présence de dimères marqués, cette sous-population ne représente pas la majorité des particules formées. Ici, en moyenne les capsides sont bien plus grosses qu'attendues et cette surestimation doit provenir en grande partie du taux de marquage qui a dû être réduit à 2,5 % pour éviter une présence trop importante des accroches non spécifiques. Le fait de multiplier la valeur de  $\langle n \rangle$  par 40 favorise la propagation d'erreurs dans l'estimation finale de  $\langle n_{tot} \rangle$ . **Cette surestimation doit également être élargie à une concentration de 200 nM en protéines mais aussi, dans une moindre mesure, à 50 nM.** Néanmoins, nos expériences de dyna-

mique à l'équilibre nous permettent d'observer une tendance globale de l'évolution de  $\langle n_{tot} \rangle$  dans la direction attendue. De plus, nous avons également accès à des informations sur la dynamique temporelle des échanges en sous-unités.

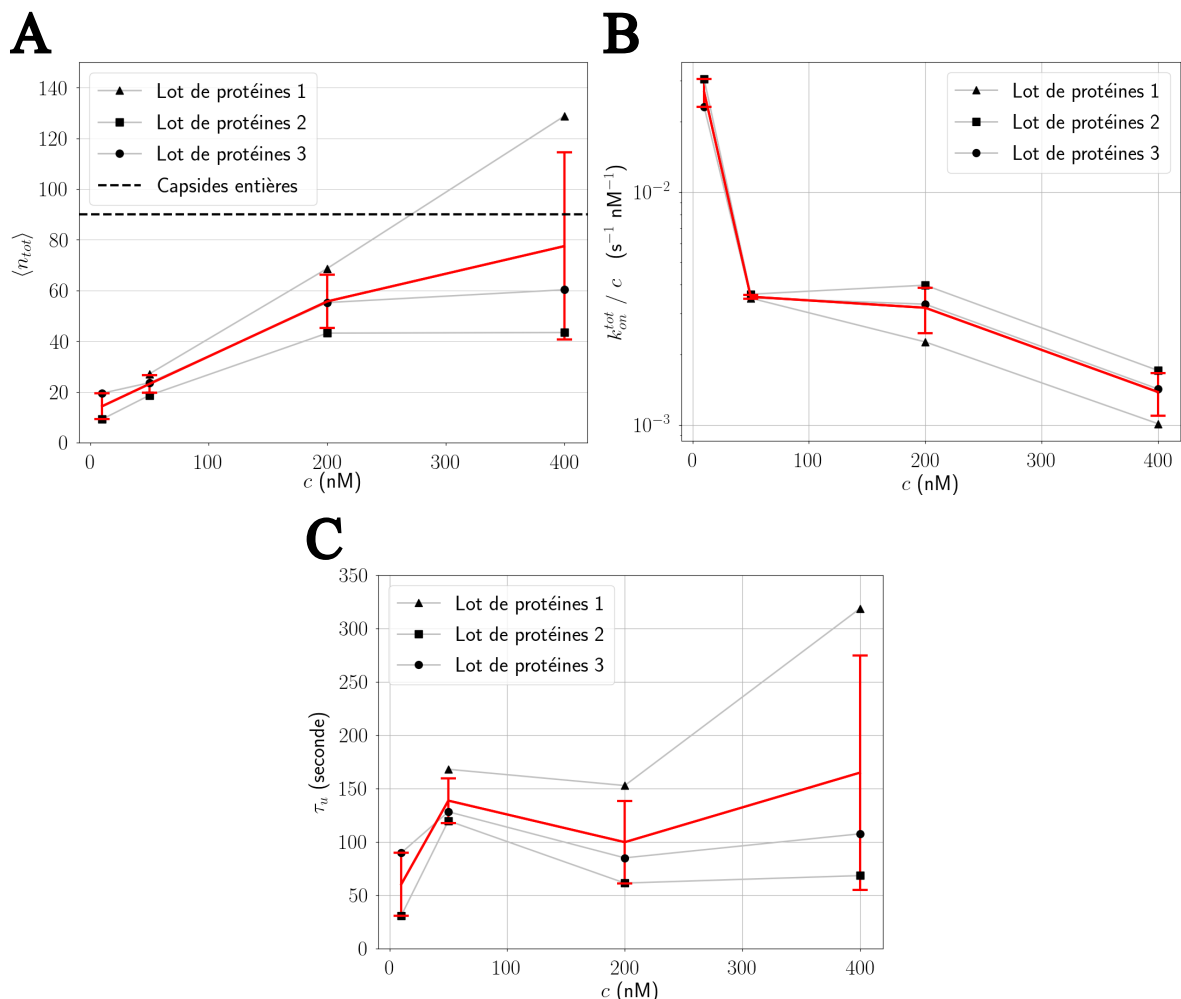
En effet, sur la Figure 5.17.A, nous représentons notamment l'évolution de la constante d'association  $k_{on}^{tot}$  en fonction de la taille moyenne  $\langle n_{tot} \rangle$  des particules observées. Ce paramètre caractéristique de la dynamique d'assemblage des capsides virales traduit la durée moyenne entre deux événements d'accroche sur un ARN. Il peut également être assimilé comme un taux de collision entre les sous-unités et le matériel génétique. En d'autres termes, cette constante d'association dépend en partie de deux paramètres importants : la concentration en protéines en solution mais également la taille des objets étudiés. Nous pouvons donc nous attendre à ce que d'un côté, plus le milieu est encombré en dimères de capside, plus le taux de collision augmente, mais de l'autre, que plus les particules se rapprochent de capsides complètes, moins les sous-unités ont tendance à interagir avec l'ARN. **Ici, nous voyons qu'en normalisant la constante d'association par la concentration  $c$  en protéines, le taux d'accroches diminue quand  $\langle n_{tot} \rangle$  augmente. Cela signifie que l'arrivée de sous-unités sur un ARN est moins fréquente quand les particules se rapprochent de capsides entièrement assemblées.**



**FIGURE 5.17 – Dynamique à l'équilibre : caractérisation de la dynamique des sous-unités de CCMV en fonction de la taille des objets observés.** Évolutions A) du rapport  $k_{on}^{tot}/c$  et B) du temps de résidence  $\tau_u$  des sous-unités en fonction de la taille  $\langle n_{tot} \rangle$  des particules assemblées. Le paramètre  $c$  indique la concentration totale en protéines. Les traits verticaux rouges représentent l'écart-type autour des valeurs moyennes de  $k_{on}^{tot}/c$  et  $\tau_u$  calculées à chaque cellule micro-fluidique. De la même façon, les traits horizontaux rouges représentent l'écart-type de  $\langle n_{tot} \rangle$ .

En combinant les résultats expérimentaux obtenus pour  $\langle n \rangle$  et  $k_{on}$ , nous pouvons calculer le temps de résidence moyen des sous-unités sur l'ARN. Ainsi, nous obtenons environ 168 - 153 et 319 secondes pour des concentrations respectives en protéines de 50 - 200 et 400 nM (Figure 5.17.B). **Si l'on compare directement les valeurs de  $\tau_u$**

pour 50 et 400 nM, le temps de résidence augmente avec la taille moyenne des particules étudiées. Ce comportement était attendu puisque plus l'assemblage tend à être complet, plus il est favorable pour les sous-unités de rester sur l'ARN. En revanche, le temps de résidence moyen à 200 nM en protéines est légèrement inférieur à celui calculé à 50 nM. Cette observation peut avoir deux explications possibles. En premier lieu, le calcul de  $n$  en début d'acquisition peut être légèrement faussé par une mauvaise estimation de la hauteur moyenne d'une marche lors du contrôle de photophysique. De plus, nous avons vu à la section 4.3.2 que la présence de clignotements entraîne une surestimation de  $k_{on}$ . L'association de ces biais peut donc très bien fausser le calcul final de  $\tau_u$ . Dans les deux cas, il est difficile de préciser si le temps de résidence moyen est sous-estimé pour 50 nM ou surestimé pour 200 nM en protéines.



**FIGURE 5.18** – Dynamique à l'équilibre : influence du lot de protéines étudié sur l'estimation des paramètres d'assemblage. Évolutions A) de  $\langle n_{tot} \rangle$ , B) du rapport  $k_{on}^{tot}/c$  et C) du temps de résidence  $\tau_u$  des sous-unités en fonction de la concentration totale en protéines. Les triangles, carrés et disques représentent les valeurs moyennes de chaque paramètre calculé pour les lots de protéines 1, 2 et 3 respectivement. Les moyennes calculées sur les trois lots sont tracées en rouge. Pour 10 et 50 nM, le taux de marquage est environ de 10 %, alors que pour 200 et 400 nM, il est d'environ 2,5 %.



Ces différents résultats ont tous été obtenus pour un même lot de protéines et à l'annexe H, je montre les distributions de  $n$  et  $\Delta t_{on}$  à 200 et 400 nM en protéines (Figure H.1). Or, à la section précédente, nous avons vu que les estimations des paramètres d'assemblage sont cohérentes d'une purification à une autre pour une concentration de 50 nM. Sur la Figure 5.18, je montre l'évolution de  $\langle n_{tot} \rangle$ ,  $\frac{k_{on}^{tot}}{c}$  et  $\tau_u$  en fonction de la concentration en protéines, et ce pour les trois lots utilisés.

Globalement, il est difficile de dire que nous observons la même reproductibilité pour les trois autres concentrations étudiées. Il est vrai que les valeurs obtenues notamment pour  $\langle n_{tot} \rangle$  et  $\tau_u$  sont assez éloignées entre les trois lots de protéines. Néanmoins, il est important de souligner que les mêmes tendances peuvent être observées. **En effet, nous retrouvons bien une augmentation globale de la taille des particules observées et une diminution de la constante d'association normalisée par  $c$  lorsque la concentration en protéines augmente (Figure 5.18.A-B).** En ce qui concerne le temps de résidence  $\tau_u$ , les mêmes difficultés sont observées pour constater une augmentation lorsque l'on passe de 50 à 200 nM en protéines (Figure 5.18.C). Ces différents résultats nous montrent bien que l'expérience de dynamique d'assemblage à l'équilibre que nous avons mise en place est robuste, malgré le grand nombre de paramètres à maîtriser. À l'annexe H, le Tableau H.1 donne les valeurs moyennes et les écarts-types de chaque paramètre d'assemblage pour les trois lots de protéines.

#### 5.4.4 Conclusion

Pour résumer, les expériences de dynamique d'assemblage ont d'abord nécessité plusieurs étapes de vérification et d'optimisation. En effet, il a fallu nous assurer que le système étudié à partir de quinze minutes après injection des dimères traduit bien un équilibre thermodynamique. De plus, ce n'est qu'après avoir obtenu nos premiers résultats que nous avons compris que le modèle présenté à la section 4.2.1 pour l'estimation de  $\langle n \rangle$  n'est valable que dans les premiers instants du film de dimères, avant que le photoblanchiment ne vienne rompre l'équilibre.

**En ce qui concerne les résultats expérimentaux, les expériences décrites précédemment nous ont permis de mettre en valeur l'avantage de travailler à l'échelle de la molécule unique : accéder à des distributions et donc observer, dans le détail, le comportement d'une population entière.** Ces histogrammes nous ont donné la possibilité de calculer des valeurs moyennes et d'en déduire des évolutions qui suivent le comportement attendu pour certains paramètres d'assemblage. En effet, le nombre moyen de sous-unités  $\langle n_{tot} \rangle$  par ARN augmente avec la concentration quand la constante d'association, normalisée par  $c$ , diminue. En d'autres termes, plus les particules observées sont proches de capsides complètes, plus le taux d'accroches en dimères sur le matériel génétique est faible puisque ce dernier est rendu moins accessible. Par ailleurs, la

formation d'objets plus proches de capsides complètes pour un rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$  croissant est cohérente avec les expériences menées en photométrie de masse. Néanmoins, nous avons pu noter une certaine limite dans la façon dont nous estimons  $\langle n_{tot} \rangle$  quand le taux de marquage est réduit à 2,5 %. En effet, à 400 nM en protéines pour le lot 1, un nombre moyen de près de 130 dimères a été estimé alors qu'une capside complète en possède au maximum 90. La plus grande source d'erreur ici se situe évidemment dans la multiplication par un facteur 40 de  $\langle n \rangle$ . C'est dans le but de réduire ce manque de précision que nous sommes en train d'associer au TIRFM un montage iSCAT comme je l'explique plus tard au chapitre 6.

**Une meilleure estimation de  $\langle n_{tot} \rangle$ , pourrait nous permettre d'affiner les résultats concernant le temps de résidence  $\tau_u$ .** En effet, les évolutions observées pour chacun des lots de protéines ne suivent pas totalement le comportement attendu à savoir une augmentation de  $\tau_u$  avec la taille des particules formées. Comme je l'ai dit dans la section précédente, le fait de ne pas observer cette tendance expérimentalement peut provenir à la fois d'un biais dans la mesure de  $\langle n \rangle$  et de  $k_{on}$ . Pour la constante d'association, nous avons vu que son évolution est cohérente avec la concentration en protéines mais également avec la taille des objets étudiés. Néanmoins, à la section 4.3.2, je montre que l'ajout de clignotements à des signaux simulés entraîne une surestimation de la constante d'association de 20 % avec Ruptures par rapport à ce que la *Ground-truth* renvoie pour ces mêmes données sans clignotements. Par exemple, lorsque je mesure une constante d'association de  $0,017 \text{ s}^{-1}$  pour 50 nM en protéines, ces valeurs sont en partie faussées par les clignotements. Ainsi, en appliquant une correction de 20 %, on se situe plutôt à  $k_{on} = 0,014 \text{ s}^{-1}$ . Cela correspond donc à une durée moyenne entre deux événements d'accroche en dimères marqués d'environ 71 secondes. Cependant, je rappelle que cette surestimation de 20 % n'est valable que pour la condition de clignotements suivante : 20 % des fluorophores qui clignotent, une durée moyenne de passage en OFF de 25 secondes et un nombre moyen de passage en OFF de 2,2. Il serait intéressant d'utiliser le contrôle de photophysique afin d'estimer les paramètres clignotements pour chaque expérience afin de connaître l'influence sur  $k_{on}$  grâce aux simulations.

**En somme, ces ajustements dans le calcul de  $\langle n_{tot} \rangle$  et  $k_{on}$  devraient nous permettre d'avoir une meilleure estimation du temps de résidence moyen. Cela viendrait affiner les tendances observées pour le nombre moyen de sous-unités par ARN ainsi que la constante d'association lorsque la concentration en protéines augmente. La reproductibilité des résultats obtenus à 50 nM pourrait également être étendue aux autres concentrations.**

## 5.5 Étude de la cinétique d'assemblage hors équilibre

Pour terminer ce chapitre, je me penche sur les expériences menées afin de suivre la cinétique des premières étapes d'assemblage des sous-unités de CCMV. Comme les conditions expérimentales changent grandement par rapport à celles de la section précédente, je commence par décrire le protocole utilisé ainsi que l'analyse mise en place. Dans un second temps, je présente les principaux résultats obtenus dans ces conditions, notamment en ce qui concerne l'évolution de l'intensité de fluorescence en fonction de la concentration en sel du tampon.

### 5.5.1 Description de l'expérience

Avant de regarder la dynamique d'assemblage à l'équilibre, les premières expériences effectuées en microscopie de fluorescence se sont déroulées hors équilibre. Cela signifie qu'au lieu de commencer l'acquisition dans la voie de détection rouge, quinze minutes après l'injection des dimères, celle-ci débute en même temps que l'arrivée des sous-unités au contact des ARN. Voici le protocole expérimental utilisé pour ces expériences :

**1** - Injection de l'ARN puis incubation de 2 minutes avant rinçage dans le tampon B (10 mM NaCl, pH = 6,1), B' (100 mM NaCl, pH = 6,1) ou B'' (200 mM NaCl, pH = 6,1).

**2** - Acquisition du film vert.

**3** - Injection de 65 nM en protéines avec un taux de marquage de 8 %, et première acquisition d'un film rouge : 1000 images, 100 ms de temps d'exposition et 1500 ms de temps de cycle. Le tampon utilisé pour diluer l'ARN est conservé.

**4** - Ensuite, le photoblanchiment est forcé en éclairant en continu pendant 3 minutes en augmentant la puissance nominale du laser à 150 mW, au lieu de 60 mW habituellement.

**5** - Nouvelle injection de protéines comme à l'étape **3** et seconde acquisition dans la voie rouge : 400 images, 100 ms de temps d'exposition et 1500 ms de temps de cycle.

Une différence notable par rapport aux expériences faites à l'équilibre est qu'un seul film est enregistré dans la voie de détection verte. Cela signifie que nous ne sommes pas sûrs de travailler avec des sites d'ARN présents tout au long de l'expérience. Néanmoins, afin de favoriser l'analyse sur des positions d'ARN avec lesquelles les sous-unités interagissent, nous nous sommes focalisés sur les sites ayant au moins une interaction spécifique parmi la première acquisition rouge. **Comme dans ces expériences Ruptures n'a pas été utilisé du fait d'une évolution trop rapide de l'intensité de fluorescence, une interaction spécifique est comptabilisée si pendant au moins 95 images consécutives, l'intensité bruitée émise par les dimères marqués est supérieure à la somme entre deux fois la hauteur moyenne d'une marche  $h$  et la valeur moyenne du fond  $\langle I_{fond} \rangle$ .** Ces deux paramètres photophysiques ont été obtenus grâce à une cellule de contrôle et j'explique à la section suivante pourquoi on considère ici au

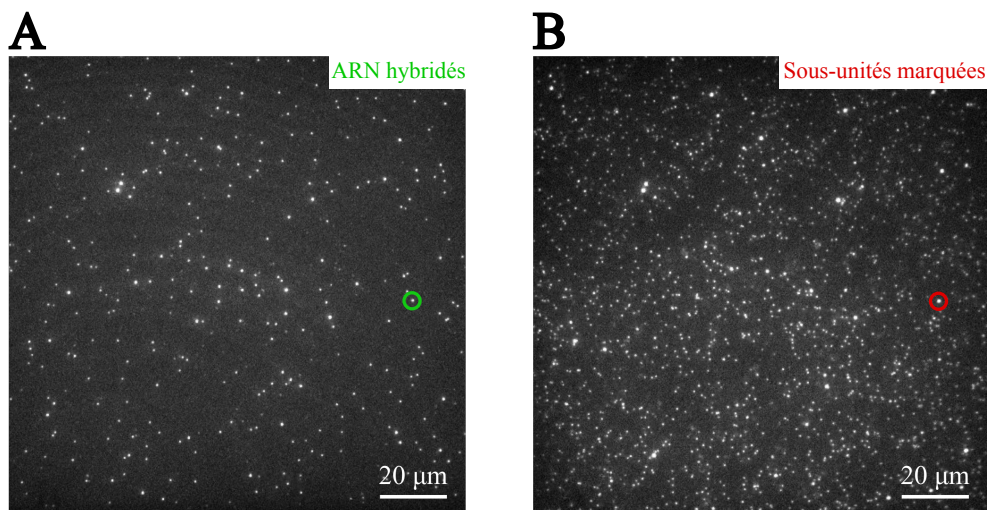
minimum deux hauteurs de marche et non une seule. Par ailleurs, puisque Ruptures n'a pas été appliqué sur les signaux expérimentaux, une étape de lissage des données a été rajoutée afin de limiter l'apparition de pic d'intensité pouvant affecter l'analyse mise en place à la section 5.5.2.2.

Enfin, mis à part ces changements, les mêmes algorithmes de détection de particules, de réduction de bruit et de correction d'une potentielle dérive mécanique ont été utilisés. Les différentes analyses décrites ci-dessous ont été faites à partir d'un seul lot de protéines.

## 5.5.2 Observation de la cinétique des premières étapes d'assemblage

### 5.5.2.1 Évolution typique des intensités de fluorescence

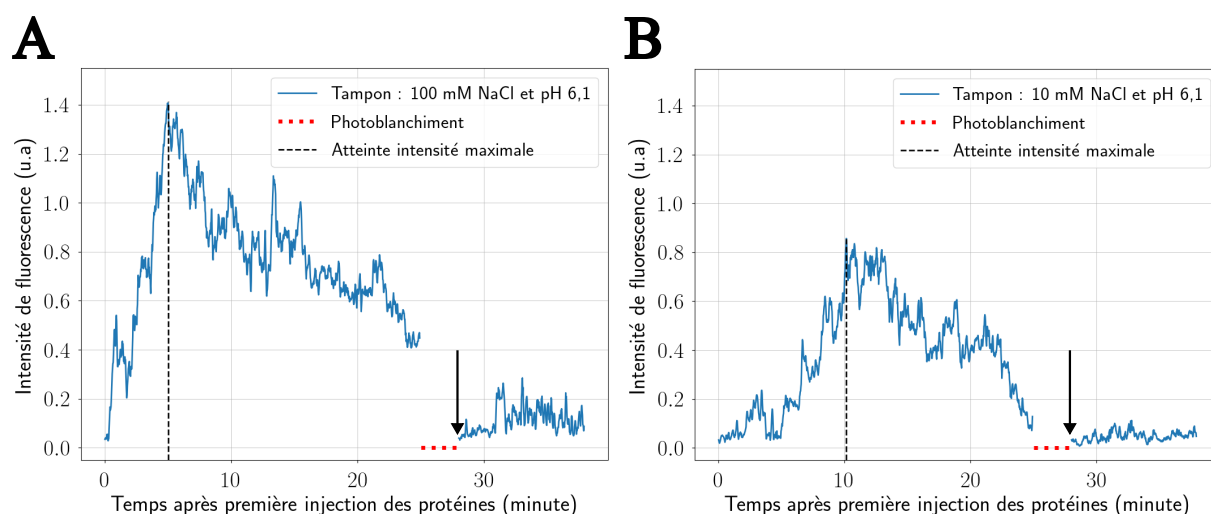
Une expérience habituelle de cinétique d'assemblage commence donc par une acquisition dans la voie de détection verte pour localiser les ARN (Figure 5.19.A) avant d'être suivie par un film de 25 minutes dans la voie rouge dès l'injection de dimères. La Figure 5.19.B illustre les accroches en sous-unités marquées après 7,5 minutes d'acquisition. Ces données ont été obtenues pour un tampon d'assemblage B' (100 mM NaCl) et nous voyons bien qu'il est possible de distinguer des accroches uniques dans chacune des voies de détection. Les deux cercles présents sur cette figure indiquent les mêmes coordonnées (X,Y) sur le champ de vue et mettent en évidence un site potentiel d'assemblage que nous regardons plus attentivement à la Figure 5.20.A.



**FIGURE 5.19 – Cinétique hors équilibre : exemples de couvertures en ARN et dimères marqués.** A) Accroches en ARN simple brin à la première image du film vert. B) Accroches en sous-unités marquées après 7,5 minutes. Les cercles vert et rouge sont positionnés aux mêmes coordonnées (X,Y) dans le champ de vue.

Sur ce site d'ARN, nous observons que l'intensité de fluorescence augmente assez rapidement, atteint son maximum  $I_{max}$  au bout de 5 minutes, puis diminue lentement jusqu'à la fin des 25 minutes d'acquisition. Au moment de réaliser ces expériences, nous

n'avions pas encore étudié la stabilité du système par des observations à l'équilibre. Nous ne savions donc pas si la diminution progressive de l'intensité était due à un départ progressif des sous-unités des capsides partiellement formées ou si le photoblanchiment en était la cause principale. Pour nous aider à comprendre, après 25 minutes d'acquisition, nous avons volontairement laissé le laser allumé pendant 3 minutes afin de photoblanchir les derniers dimères visibles sur le champ de vue. Ensuite, de nouvelles protéines ont été injectées 28 minutes après le début de l'expérience (flèches noires sur la Figure 5.20) et une seconde acquisition rouge a pu démarrer. Ainsi, sur de nombreux signaux comme ceux illustrés par la Figure 5.20, l'intensité de fluorescence n'augmentait que très peu, ce qui nous a convaincu que les sous-unités étaient toujours présentes sur l'ARN et que le photoblanchiment était la principale cause de la diminution observée. Finalement, les résultats illustrés précédemment sur la Figure 5.11 nous ont confortés au sujet de la stabilité des particules partiellement formées.



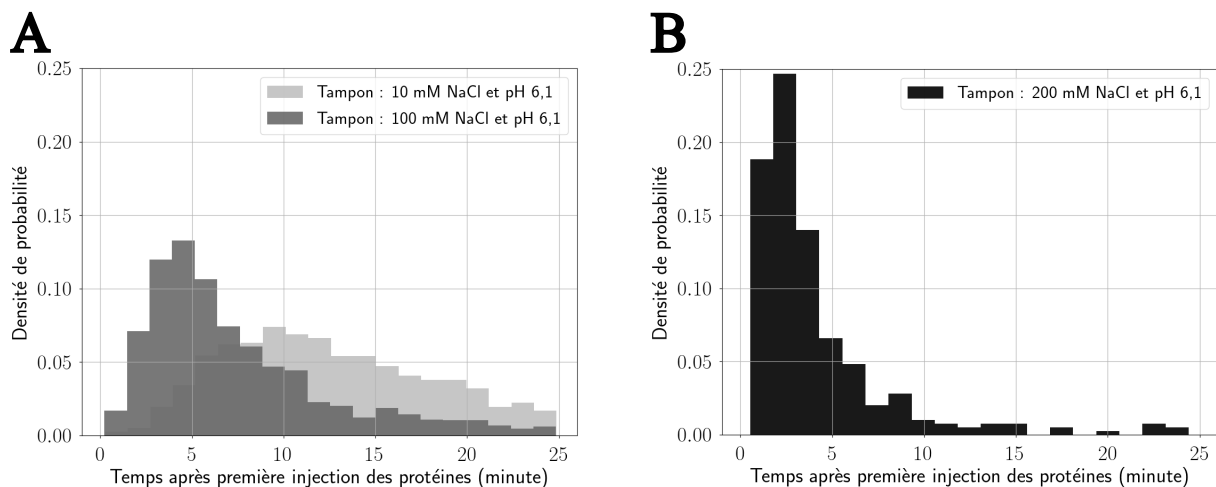
**FIGURE 5.20 – Cinétique hors équilibre : évolutions typiques de l'intensité de fluorescence émise par des sous-unités marquées sur un site d'ARN.** Suivi de l'intensité émise par des dimères marqués pour une concentration en NaCl dans la solution tampon de A) 10 mM et B) 100 mM. Les données expérimentales sont tracées en bleu, les pointillés rouges indiquent la période temporelle où le photoblanchiment a été forcé et les tirets noirs montrent l'instant où l'intensité de fluorescence est maximale. La flèche noire correspond au moment où la deuxième injection en protéines a lieu.

Ces expériences hors équilibre ont été réalisées pour différentes concentrations en NaCl comme évoqué à la section 5.5.1 : 10 - 100 et 200 mM. Sur la Figure 5.20.B, je montre l'évolution de l'intensité de fluorescence émise par des sous-unités marquées sur un site d'ARN pour une concentration en sel de 10 mM. **Sur ce signal, nous voyons que l'intensité maximale est moins élevée mais surtout qu'elle est atteinte plus tard qu'avec 100 mM.** L'observation d'une valeur de  $I_{max}$  plus faible est sans doute directement liée à une cinétique plus lente des premiers instants d'accroche. En effet, si les sous-unités arrivent plus lentement sur l'ARN, le photoblanchiment des premières

arrivées sera visible plus tôt. **Afin, de mettre en avant cette différence de cinétique, nous avons décidé de collecter l'instant  $t$  où  $I_{max}$  est atteinte pour tous les signaux d'intérêt.** Ces derniers devant caractériser un début d'assemblage assez net, nous avons conservé uniquement les données avec au moins deux accroches en dimères, *i.e.* où l'intensité est supérieure à  $2h + \langle I_{fond} \rangle$ . De plus pour nous éviter de collecter des signaux avec un pic d'intensité très élevé mais furtif, caractéristique d'un agrégat, cette valeur d'intensité seuil doit être conservée pendant au moins 95 images (143 secondes). Il s'agit de la durée moyenne avant photoblanchiment estimée pour ces expériences. Les résultats de l'analyse cinétique sont détaillés à la section suivante.

### 5.5.2.2 Influence de la salinité sur la cinétique d'assemblage

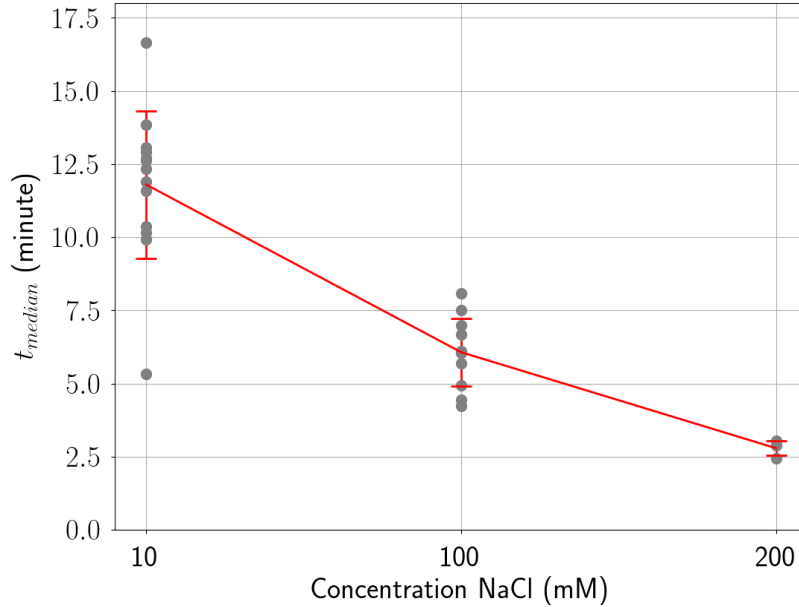
La Figure 5.21.A montre les deux histogrammes issus du protocole d'analyse cinétique détaillé précédemment. À partir de ces deux distributions, **il ressort clairement que les premières étapes de l'assemblage de capsides de CCMV sont plus rapides pour une concentration en sel de 100 mM.** En effet, en calculant la médiane  $t_{median}$  des instants temporels où  $I_{max}$  est atteinte, on trouve environ 6,1 et 11,8 minutes pour des concentrations respectives en NaCl de 100 et 10 mM.



**FIGURE 5.21 – Cinétique hors équilibre : influence du sel sur la distribution des instants où  $I_{max}$  est atteinte.** Les distributions sont tracées pour une concentration en NaCl de A) 10 mM en gris clair, 100 mM en gris foncé et B) 200 mM en noir. Le taux de marquage pour ces trois conditions est de 8 %.

Cette accélération de la cinétique des premiers moments de l'assemblage peut surprendre si l'on se rappelle que l'une des actions du sel consiste à réduire les interactions entre dimères de protéines et ARN. En écrantant le matériel génétique, le sel peut ralentir la formation de capsides voire totalement l'inhiber selon la concentration utilisée. Néanmoins, en passant de 10 à 100 mM en NaCl, la concentration en sel devient similaire à celle caractéristique des conditions physiologiques (100 à 150 mM). **Une première**

hypothèse est donc de considérer que le tampon B' constitue un environnement plus favorable pour l'assemblage de capsides de CCMV. Je précise que pour ces deux histogrammes, les instants temporels où  $I_{max}$  est atteinte ont été collectés sur plusieurs cellules micro-fluidiques et journées d'expérience.



**FIGURE 5.22** – Cinétique hors équilibre : évolution de  $t_{median}$  en fonction de la concentration en sel. Chaque disque gris représente le résultat d'une cellule micro-fluidique. Les traits verticaux rouges représentent l'écart-type autour des valeurs de  $t_{median}$  calculées à chaque cellule micro-fluidique. Pour les trois conditions, le taux de marquage est de 8 %.

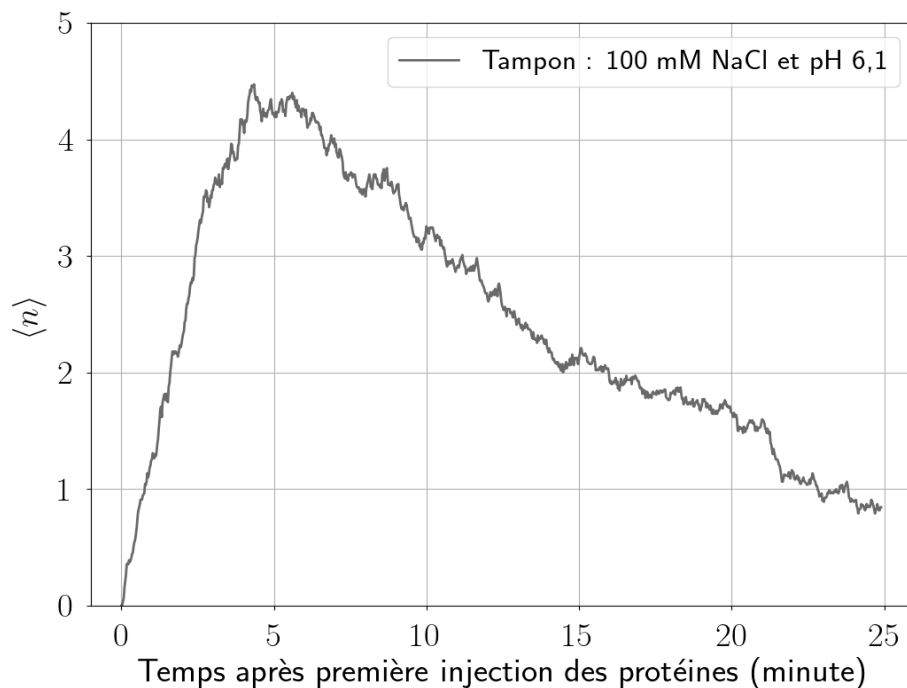
Cependant, ce qui est plus surprenant est de voir que cette accélération est toujours visible lorsque l'on passe de 100 à 200 mM en NaCl (Figure 5.21.B). **Pour cette nouvelle condition expérimentale, le calcul du temps médian  $t_{median}$  nous donne environ 2,8 minutes. L'intensité maximale a donc été atteinte deux fois plus rapidement avec le tampon B'' par rapport à B'.** Il semblerait donc qu'un autre mécanisme facilite l'arrivée des premiers événements d'accroche quand la concentration en sel augmente. Je rappelle que Garmann et al. (2022) ont montré à partir d'une étude sur le BMV, qu'en augmentant les interactions dimères - ARN, *i.e.* en réduisant la concentration en sel, la durée de nucléation est plus courte mais que la croissance des capsides n'est pas altérée. **Au contraire, de notre côté, il semblerait qu'en augmentant la concentration en sel, les premières accroches en dimères marqués sont facilitées et s'enchaînent plus rapidement.** Certes, les protéines de BMV et CCMV partagent près de 70 % de similitudes entre leurs séquences d'acides aminés, mais la répartition des charges n'est pas identique entre les deux virus. En effet, Duran-Meza et al. (2021), ont observé que le point isoélectrique du BMV est plus proche de la neutralité que celui du CCMV. En d'autres termes, à pH légèrement acide comme celui auquel nous travaillons,

le corps de la protéine de BMV peut être considéré comme globalement neutre alors que le CCMV sera plutôt négatif. Cet écart de charge pourrait alors jouer un rôle dans la différence de cinétique d'assemblage que nous avons observée avec le CCMV. Par ailleurs, comme nous l'avons vu en photométrie de masse à la section 5.1.2, il n'est pas exclu non plus que le marquage puisse légèrement influencer le mécanisme d'auto-assemblage.

Pour finir, la Figure 5.22 illustre la reproductibilité des expériences menées hors équilibre. En effet, pour les trois concentrations en sel nous voyons que les valeurs de  $t_{median}$  calculées à partir de chaque cellule utilisée pour l'analyse (1 disque gris = 1 cellule) sont relativement peu dispersées. On trouve un écart-type de 2,5 - 1,2 et 0,2 minutes pour respectivement 10 - 100 et 200 mM en NaCl. Je précise que les trois points mesurés pour le tampon B' ont été obtenus sur la même journée d'expérience.

### 5.5.2.3 Estimation du nombre de sous-unités marquées par ARN

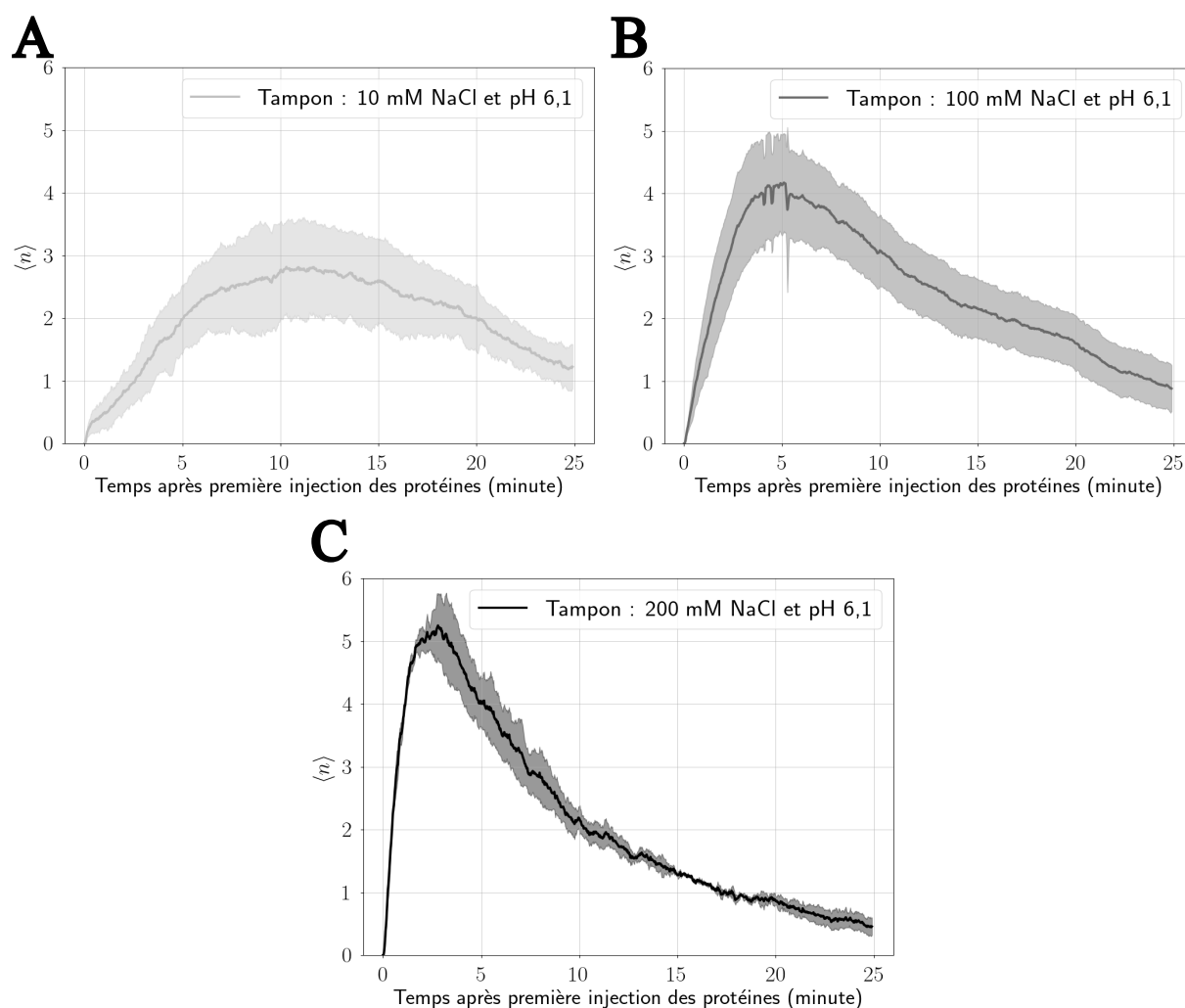
Une autre façon de regarder l'influence de la concentration en sel sur la cinétique des premiers moments de l'assemblage est de calculer le nombre moyen de sous-unités marquées  $\langle n \rangle$  par ARN. Cette estimation expérimentale suit le même protocole que celui présenté pour la dynamique à l'équilibre présenté à la section 5.4.1.2. Le seul changement est que j'utilise l'intensité de fluorescence bruitée, puisque Ruptures n'a pas été appliqué ici. Sur la Figure 5.23, je montre un exemple d'évolution de  $\langle n \rangle$  pour une cellule micro-fluidique étudiée dans le tampon B' (100 mM NaCl).



**FIGURE 5.23** – Cinétique hors équilibre : évolution de  $\langle n \rangle$  pour 100 mM en NaCl. Cette courbe a été tracée en moyennant les intensités de fluorescence à chaque image pour une seule cellule micro-fluidique. Le taux de marquage est de 8 %.



À la différence de ce que j'ai pu montrer sur la Figure 5.13, où le nombre moyen de sous-unités ne fait que diminuer dès le début de l'acquisition, trois phases peuvent être mises en évidence ici. Pendant environ 4 minutes,  $\langle n \rangle$  augmente assez rapidement, puis une certaine stabilité est observée durant près de 2 minutes avant qu'une décroissance progressive ait lieu jusqu'à la fin des 25 minutes. Il faut donc comprendre que pour cette cellule, le photoblanchiment met environ 6 minutes avant de prendre le dessus par rapport à l'arrivée de sous-unités marquées sur les différents sites d'ARN. Cette durée est donc bien en accord avec la valeur de 6,1 minutes calculée pour  $t_{median}$  à la section précédente avec 100 mM en sel.



**FIGURE 5.24 – Cinétique hors équilibre : influence de la concentration en sel sur  $\langle n \rangle$ .** Suivi de l'évolution de  $\langle n \rangle$  après la première injection de protéines pour une concentration en sel de A) 10 mM, B) 100 mM et C) 200 mM. Les trois courbes en traits pleins ont été moyennées sur l'ensemble des cellules analysées pour chaque condition expérimentale. Les zones plus claires correspondent aux écarts-types qui ont été calculés à partir des valeurs de  $\langle n \rangle$  obtenues pour chaque cellule micro-fluidique. Pour les trois conditions, le taux de marquage est de 8 %.

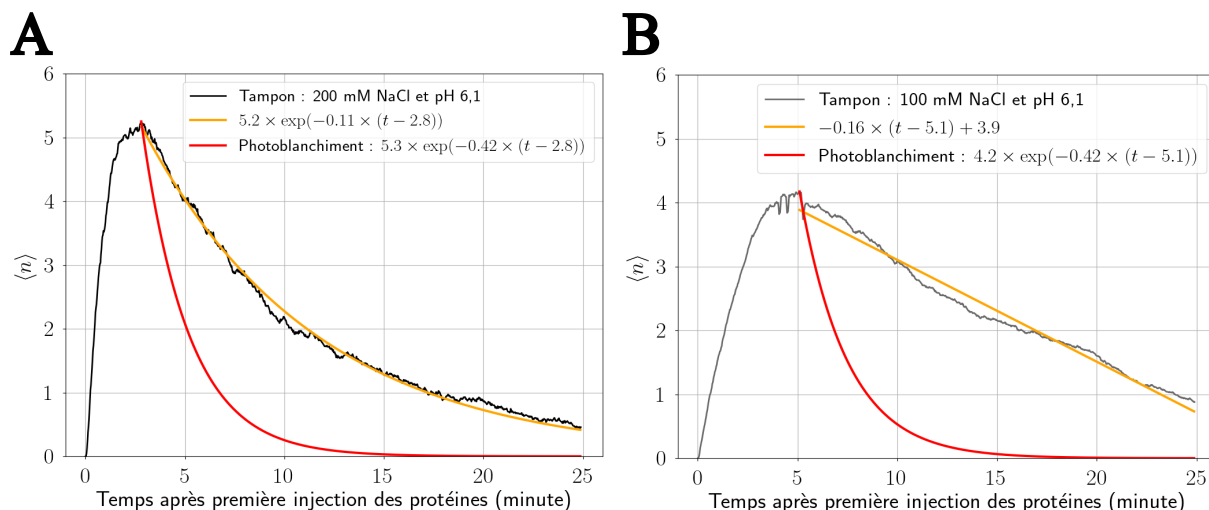
Du fait du photoblanchiment, nous ne pouvons pas mettre en évidence l'atteinte de

l'état d'équilibre thermodynamique comme à la section 5.4.2.1. **Toutefois, au bout de 4 minutes d'acquisition la valeur maximale de  $\langle n \rangle$  est d'environ 4,2 sous-unités marquées, ce qui est plus que tous les résultats décrits pour les expériences de dynamique à l'équilibre, au bout de 15 minutes d'incubation et sans photoblanchiment.** La concentration totale en protéines est certes, légèrement supérieure ici (65 nM) que pour les résultats illustrés en Figure 5.15.A (50 nM), mais cela ne suffit pas à expliquer cet écart. Une différence qu'il faut préciser est que les deux séries d'expériences menées à l'équilibre et hors équilibre ont été réalisées à un an d'intervalle, avec des protocoles de purification qui ont légèrement évolué et des lots de feuilles infectées qui ont changé. Il sera intéressant de refaire ces tests de cinétique avec un nouveau lot de protéines, mais à l'heure actuelle nous pouvons dire que cet écart dans la valeur de  $\langle n \rangle$  fait partie de la variabilité mise en évidence pour les expériences réalisées à l'équilibre.

Pour continuer sur l'influence de la concentration en sel, ces courbes d'évolution de  $\langle n \rangle$  ont été tracées pour chaque cellule micro-fluidique analysée puis moyennées pour une même condition expérimentale. Le résultat final est illustré par la Figure 5.24. Sur chacune des courbes obtenues, nous retrouvons les trois phases mentionnées plus haut : croissance - stabilité - décroissance. Dans un premier temps, nous avons vu que le temps médian  $t_{median}$  pour atteindre  $I_{max}$  diminue quand la concentration en sel augmente. Le photoblanchiment est donc visible après l'arrivée de plus de sous-unités marquées pour le tampon B". **C'est principalement pour cette raison que nous trouvons une valeur maximale de  $\langle n \rangle$  de plus en plus grande : 2,8 - 4,2 - 5,3 sous-unités marquées respectivement pour 10 - 100 et 200 mM en NaCl.**

En ce qui concerne la phase de stabilité, nous voyons qu'elle semble de plus en plus longue quand les interactions ARN - dimères augmentent, ce qui retarde par conséquent le début de la décroissance. En d'autres termes, le photoblanchiment semble plus rapidement prendre le dessus sur l'accroche de nouvelles sous-unités marquées quand la concentration en sel est importante. Cette différence de comportement entre les deux premières phases de l'évolution de  $\langle n \rangle$  peut paraître contradictoire. D'un côté, une plus forte concentration en sel semble faciliter l'arrivée de dimères marqués mais de l'autre, cela fige plus rapidement l'état d'assemblage des capsides. **Une hypothèse ici est qu'à 200 mM en NaCl, un état d'équilibre global est atteint plus tôt et pour une courte durée, avant que le photoblanchiment ne prenne définitivement le dessus. À l'inverse, pour des interactions plus fortes entre dimères de protéines et ARN, la mise en place de cet équilibre global prend plus de temps.** Cette interprétation étant grandement influencée par l'apparition du photoblanchiment, il est difficile de généraliser ce comportement à tous les dimères de capside et d'associer le maximum de chacune des courbes de la Figure 5.24 à l'atteinte de l'équilibre thermodynamique étudié à la section 5.4. Par ailleurs, comme l'étude sur le BMV faite par Garmann et al. (2022) ne montre

pas la même évolution de la cinétique d'assemblage en fonction de la salinité du milieu, des expériences supplémentaires nécessitent d'être réalisées pour mieux comprendre les résultats que nous avons obtenus.



**FIGURE 5.25 – Cinétique hors équilibre : évaluation de l'influence du photoblanchiment sur la décroissance de  $\langle n \rangle$ .** Comparaison entre l'évolution caractéristique du photoblanchiment (en rouge) et celle de  $\langle n \rangle$  pour A) 200 mM en NaCl (courbe noire) et B) 100 mM en NaCl (courbe grise). En orange, je trace respectivement un ajustement exponentiel et linéaire de la décroissance de  $\langle n \rangle$ . Le taux de marquage est de 8 % dans les deux conditions.

Pour terminer, sur la Figure 5.25.A, nous voyons que même si la phase de stabilité est obtenue plus rapidement avec le tampon B'', la décroissance de  $\langle n \rangle$  est plus lente que celle du photoblanchiment. En effet, les deux ajustements exponentiels tracés en rouge et en orange nous montrent que la constante de temps du mécanisme photophysique est environ 4 fois grande que celle associée au nombre moyen de dimères marqués. Comme je l'ai expliqué à la section 5.4.2.2, une partie de cet écart peut s'expliquer par une mauvaise estimation de la constante de temps du photoblanchiment en conditions d'assemblage. Néanmoins, les phases de décroissance observées sur les Figures 5.24.A-B semblent encore plus lentes, avec notamment un comportement plutôt linéaire dans le cas du tampon B' (Figure 5.25.B). De ce fait, cette différence de décroissance pour les trois solutions tampon peut mettre en évidence une arrivée de nouvelles sous-unités marquées sur le matériel génétique qui est simplifiée lorsque les interactions dimères - ARN sont plus importantes (*i.e.* pour des concentrations en sel plus faibles). Une façon d'être plus quantitatif sur ce point serait de faire de nouvelles expériences dans ces conditions mais en continuant les acquisitions plus longtemps, afin d'observer la totalité de la décroissance dans chaque cas.

### 5.5.3 Conclusion

Au cours de cette cinquième partie, je me suis penché sur le comportement des sous-unités marquées hors équilibre. Les expériences que j'ai décrites nous ont donné des informations temporelles sur l'assemblage global des dimères de capsid autour de leur ARN en fonction de la concentration en sel. En effet, **nous avons pu voir qu'en passant de 10, à 100 puis 200 mM en NaCl, la cinétique des premières accroches en sous-unités marquées est plus rapide, tout comme l'obtention d'un équilibre global entre accroches en dimères marqués et photoblanchiment.** Le fait que ces interprétations soient valables quand on compare les tampons B' et B'' marque une opposition notamment avec les conclusions établies par Garmann et al. (2022) sur le comportement du BMV. Les différences de charge entre le CCMV et le BMV nous laissent penser que les deux sous-unités de capsid peuvent interagir différemment avec leur matériel génétique respectif pour une même force ionique. Par ailleurs, comme nous avons déjà pu le remarquer en photométrie de masse (5.1.2), il est possible que le marquage influence légèrement le mécanisme d'auto-assemblage du CCMV.

Dans le but de mieux comprendre les résultats obtenus, il serait intéressant de réaliser des expériences complémentaires. **En effet, les tests détaillés à 200 mM en sel montrent des accroches très rapides en sous-unités, mais rien ne nous dit que ces événements ne se sont pas produits de manière totalement désordonnée, menant à des particules partiellement assemblées et piégées cinétiquement.** Des expériences de microscopie électronique peuvent être envisagées afin de voir l'état des particules obtenues quelques minutes après le début de l'assemblage. Toutefois, ces tests impliqueraient de devoir augmenter les concentrations en protéines et matériel génétique, ce qui pourrait modifier les comportements observés en microscopie TIRF. Une autre possibilité serait de refaire des expériences de photométrie de masse, pour une concentration fixe en protéines pour différents tampons d'assemblage. De cette façon, nous pourrions comparer les masses des particules formées pour des conditions expérimentales plus représentatives de celles utilisées pour nos mesures.

Par ailleurs, je rappelle que toutes ces expériences ont été réalisées pour un seul lot de protéines. **Nous avons pu voir que le nombre moyen de sous-unités calculé à l'équilibre pour 50 nM en protéines est plus faible que celui estimé hors équilibre à partir de 65 nM en protéines, alors que le tampon B' a été utilisé dans les deux cas.** L'écart est difficile à réellement quantifier à cause de l'effet du photoblanchiment sur les expériences de cinétique mais il est d'au moins une sous-unité en moyenne. Une partie de cette différence peut être associée à une concentration légèrement supérieure, mais elle n'explique pas tout. Il est donc important de refaire ces expériences avec d'autres lots de protéines purifiées pour avoir de nouvelles estimations de  $\langle n \rangle$  et du temps médian  $t_{median}$  en fonction de la concentration en sel. Ces dernières peuvent être

élargies à d'autres valeurs pour voir si l'assemblage finit par être ralenti au bout d'un certain temps.

Enfin, comme je l'ai précisé à la section 5.5.1, Ruptures n'a pas pu être utilisé du fait de données trop riches en événements d'arrivée en dimères marqués. Dans le but d'obtenir des informations temporelles à l'échelle de l'accroche unique, nous pouvons envisager d'augmenter la fréquence d'acquisition pour chaque concentration en sel. **Cela nous permettrait de mieux mettre en évidence les arrivées consécutives, surtout dans le cas où elles se produisent très rapidement avec le tampon B'', et ainsi faciliter l'application de Ruptures.** En revanche, cette modification du protocole se ferait au détriment d'un photoblanchiment encore plus présent.

## 5.6 Conclusion du chapitre

Ce chapitre m'a permis de mettre en avant l'avantage des expériences réalisées à l'échelle de la molécule unique. En effet, que ce soit la photométrie de masse ou la microscopie TIRF, nous pouvons caractériser l'hétérogénéité d'un paramètre, ou d'un comportement, là où les mesures d'ensemble ne renvoient que des moyennes. Il est vrai qu'en fin de compte, des tendances sont extrapolées à partir de valeurs moyennées sur plusieurs expériences, mais nous pouvons toujours revenir travailler à l'échelle du signal unique et observer la diversité du mécanisme d'auto-assemblage.

En travaillant en microscopie de fluorescence, nous avons pu obtenir des informations temporelles sur la dynamique d'assemblage à l'équilibre des sous-unités de CCMV. Ces expériences nous ont fourni des premières estimations sur le temps de résidence moyen des dimères mais également sur leur constante d'association. En normalisant cette dernière par la concentration en protéines, il ressort que plus la taille des objets formés est importante, plus les accroches se produisent rarement. À l'heure actuelle, il nous faut encore travailler sur l'aspect quantitatif de nos expériences, notamment dans l'estimation du nombre moyen de sous-unités  $\langle n_{tot} \rangle$  par ARN, mais les premiers résultats sont prometteurs. Une des solutions envisagées est détaillée au chapitre 6, où j'explique que nous sommes en train de coupler le TIRF avec un montage iSCAT.

Enfin, certaines interprétations données dans ce chapitre ont également rappelé la question habituelle lorsque l'on travaille en microscopie de fluorescence : est-ce que le marquage influence les résultats obtenus? À défaut de montrer que des capsides complètes ne peuvent pas être obtenues en présence de sous-unités marquées, les tests de photométrie de masse ont mis en avant un comportement différent de ce qui est obtenu quand le marquage est absent. Pour les expériences hors équilibre, les différences de cinétique d'assemblage du CCMV et du BMV en fonction de la salinité du milieu peuvent potentiellement provenir d'une différence au niveau de la charge portée par chaque virus. Toutefois, la présence d'un Atto647N sur un dimère de CCMV peut modifier la charge

globale de ce dernier et ainsi influencer son interaction avec les autres sous-unités ou le matériel génétique. Ce questionnement doit être gardé en tête pour les prochaines expériences, mais il sera important de réaliser des tests complémentaires, pour apporter de nouveaux éléments de réponse à ce sujet.



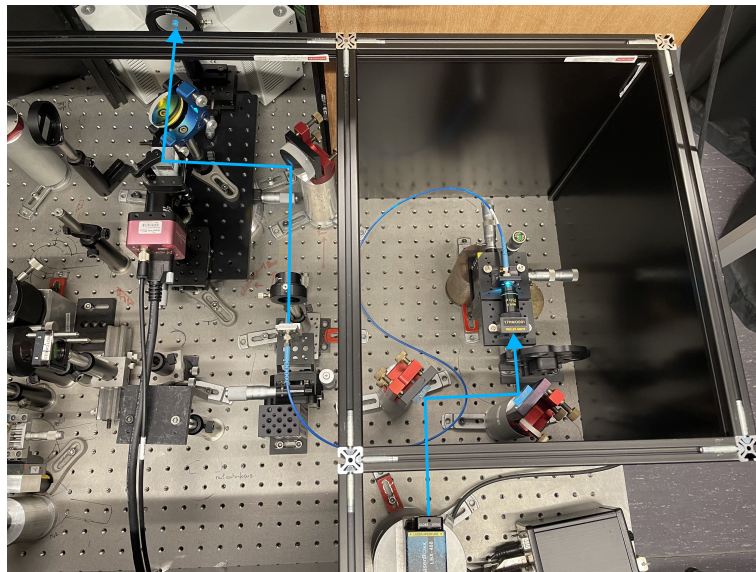
## Chapitre - 6

---

### Vers l'implémentation d'un montage iSCAT

---

Ce dernier chapitre est consacré à la microscopie iSCAT. Pour commencer, je reviens sur certains aspects importants de cette technique optique en m'appuyant sur quelques travaux réalisés dans le domaine. Dans un second temps, je présente le montage expérimental que nous avons décidé de coupler avec le microscope TIRF (Figure 6.1). Je montre également les premières images que nous avons pu obtenir.



**FIGURE 6.1** – Illustration du montage expérimental iSCAT couplé au TIRFM. Les flèches bleues indiquent le trajet du laser jusqu'à l'entrée du microscope. Elles sont reliées par une fibre monomode.



## 6.1 La microscopie à diffusion interférométrique

La photométrie de masse présentée à la section 5.1, m'a permis d'illustrer une des applications possibles en microscopie à diffusion interférométrique (*Interferometric Scattering* ou iSCAT). Nous avons vu que cette technique se base sur une information d'origine lumineuse pour quantifier la masse de nanoparticules inconnues. Afin de mieux comprendre les principes sous-jacents, je propose une brève revue bibliographique concernant la microscopie iSCAT. J'insiste notamment sur certains aspects théoriques importants avant de décrire quelques-uns des montages expérimentaux rencontrés dans la littérature.

### 6.1.1 Les principes théoriques

À la fin du chapitre 2, j'évoque la microscopie iSCAT en tant que méthode expérimentale permettant de quantifier, à l'échelle de la molécule unique, l'assemblage de particules virales. Cette technique optique se base sur les interférences créées entre un faisceau de référence, dont le champ électrique est noté  $E_r$ , et le signal diffusé par la particule d'intérêt, noté  $E_s$ . Sans tenir compte des conditions d'illumination utilisées, il est possible d'écrire l'intensité captée par le détecteur,  $I_{det}$ , selon la relation suivante :

$$I_{det} \propto |E_r + E_s|^2 = E_i^2 \times [r^2 + s^2 + 2rs \cos \Phi] \quad (6.1)$$

où  $E_i$  est le champ électrique incident,  $r$  l'amplitude relative associée au signal de référence considéré,  $s$  celle du signal diffusé et  $\Phi$  la différence de phase entre  $E_r$  et  $E_s$ . À partir de cette équation, il est possible de distinguer les deux principales configurations retrouvées en microscopie par diffusion : celle entièrement basée sur le signal issu de la particule étudiée et celle qui s'intéresse aux interférences (Ortega Arroyo et al. (2016)). Dans le premier cas, le montage généralement utilisé concerne la microscopie en champ sombre dont un exemple est illustré en Figure 6.2.A. Ici, l'illumination en réflexion totale interne permet de suffisamment séparer le champ incident du champ diffusé. Dans ce cas de figure, le champ de référence correspond à celui associé au fond et l'intérêt de cette méthode est de réduire le plus possible sa contribution pour le rendre négligeable. Ainsi, le terme en  $|s^2|$  est prédominant par rapport aux deux autres et l'équation 6.1 devient :

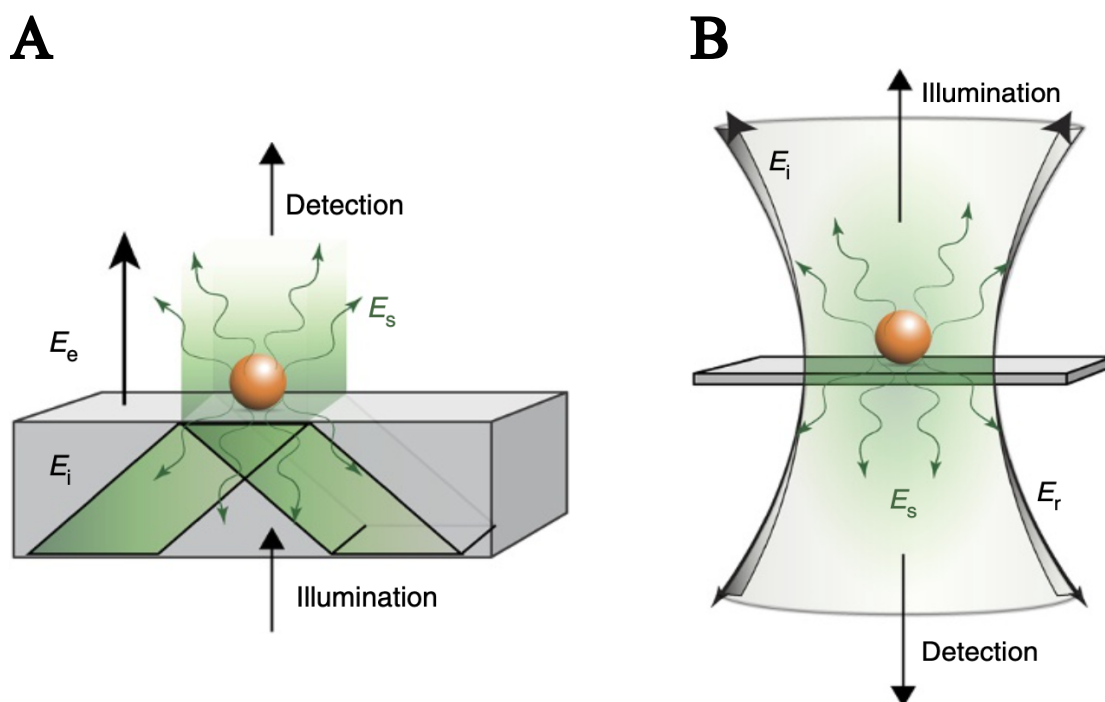
$$I_{df} \propto E_i^2 \times s^2 \quad (6.2)$$

où  $I_{df}$  indique l'intensité captée au niveau du détecteur en configuration de microscopie en champ sombre. Par ailleurs, il est possible de mettre en évidence les paramètres dont va dépendre l'intensité collectée en sortie du montage. En effet, l'amplitude du signal diffusé peut s'écrire selon la relation  $s^2 = \frac{\sigma_{scat}}{A}$  où  $\sigma_{scat}$  est la section efficace de diffusion d'une particule et  $A$  l'aire de cette dernière en considérant le rayon de sa tache limitée par la diffraction. L'expression de la section efficace de diffusion est donnée ici en travaillant

dans le régime de Rayleigh :

$$\sigma_{scat} = \frac{128\pi^5 R^6}{3} \times \left(\frac{n_m}{\lambda}\right)^4 \times \left(\frac{n_p^2 - n_m^2}{n_p^2 + 2n_m^2}\right)^2 \quad (6.3)$$

où  $n_m$  et  $n_p$  sont les indices de réfraction respectifs du milieu et de la particule,  $R$  le rayon de cette dernière et  $\lambda$  la longueur d'onde de la source lumineuse utilisée. **Le signal capté par le détecteur étant proportionnel à  $R^6$ , cela signifie que l'intensité collectée est donc très dépendante de la taille l'échantillon.** Par exemple, en passant d'une particule avec un diamètre de 100 à 10 nm, l'amplitude du signal est divisée par un facteur  $10^6$ .



**FIGURE 6.2 – Deux exemples de configurations de microscopie par diffusion.** A) Microscopie en champ sombre à partir d'une illumination en réflexion totale interne. B) Microscopie iSCAT. On note :  $E_i$  le champ incident,  $E_s$  le champ diffusé,  $E_r$  le champ de référence (réfléchi ici) et  $E_e$  le champ évanescent. Les figures sont adaptées de Ortega Arroyo et al. (2016).

Dans le but de limiter cette dépendance et de pouvoir détecter plus facilement des objets de l'ordre de la dizaine de nanomètres, la microscopie par diffusion interférométrique (iSCAT) est préférée (Figure 6.2.B). Dans cette configuration, nous regardons les interférences entre le champ diffusé et celui réfléchi par une interface entre deux milieux d'indices différents. L'amplitude relative du champ de référence est environ égal à 0,067 pour une réflexion à une interface entre du verre ( $n_v = 1,52$ ) et de l'eau ( $n_e = 1,33$ ). En comparaison, si l'on considère un virion natif de CCMV, d'indice de réfraction moyen pour une protéine  $n_p = 1,45$  (Faez et al. (2015)) et de rayon  $R = 15$  nm, éclairé par un laser de longueur

d'onde  $\lambda = 488$  nm, on trouve une section efficace de diffusion  $\sigma_{scat} = 2,9 \times 10^{-8} \mu\text{m}^2$ . De plus, en considérant une aire associée à la limite de diffraction  $A = \pi(0,2 \mu\text{m})^2$ , on trouve une amplitude relative pour le signal diffusé  $s = 4,8 \times 10^{-4}$ . Ainsi, en microscopie iSCAT, le terme en  $s^2$  est négligeable devant les deux autres et l'équation 6.1 devient :

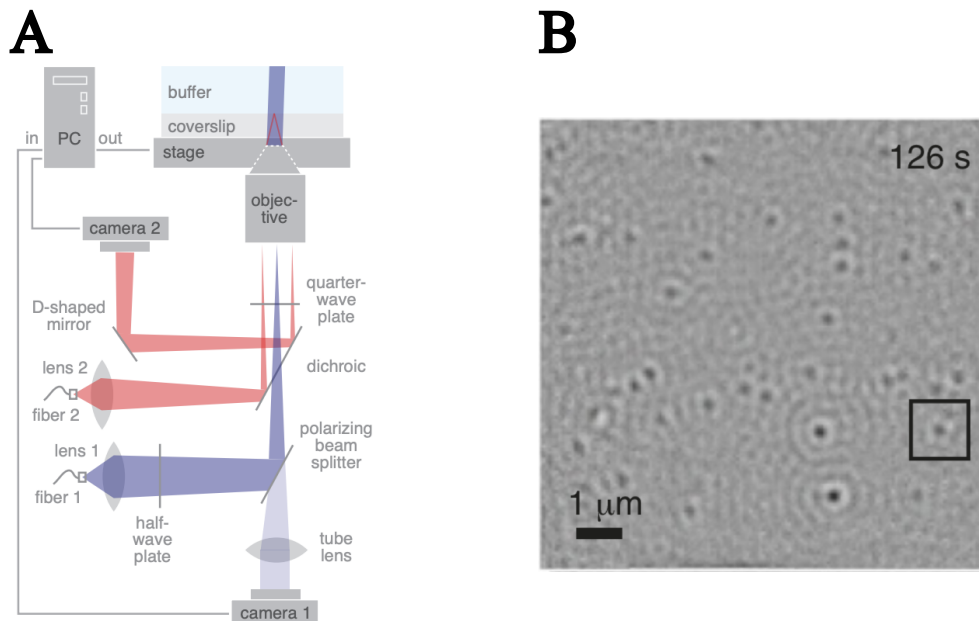
$$I_{iSCAT} \propto E_i^2 \times [r^2 + 2rs \cos \Phi] \quad (6.4)$$

où  $I_{iSCAT}$  est l'intensité collectée en microscopie par diffusion interférométrique. Cette expression nous permet d'estimer le contraste  $C$  de la particule observée à partir de la relation :  $C = \frac{I_{iSCAT}}{I_r} = 1 + \frac{2s \cos \Phi}{r}$  où  $I_r$  est l'intensité du champ réfléchi. **Avec cette expression il apparaît clairement que le signal d'intérêt est désormais proportionnel à  $s$  et donc seulement à  $R^3$ , ce qui atténue la dépendance vis-à-vis de la taille de la particule.** En appliquant cette formule à la situation précédente, on trouve un contraste théorique d'environ 1,4 % pour un virion natif de CCMV. Ce résultat se base sur la formule générale de la section efficace de diffusion (6.3) sans tenir compte des propriétés de détection propres à chaque montage. Certaines d'entre elles sont mises en évidence dans la section suivante.

### 6.1.2 Exemples de configurations expérimentales

Les configurations expérimentales rendant possible la microscopie à diffusion interférométrique sont nombreuses. En effet, l'objectif étant de faire interférer deux sources lumineuses, la façon dont l'échantillon est illuminé peut varier. Néanmoins, en parcourant la littérature, il est possible de retrouver principalement deux montages, qui nécessitent tous les deux de travailler en réflexion. **Le premier consiste à éclairer l'échantillon en champ large comme l'illustre la Figure 6.3.A.**

Sur ce montage, nous voyons que le laser bleu est focalisé dans le plan focal arrière de l'objectif dans le but d'avoir un faisceau collimaté en sortie. La réflexion de ce même laser sur l'interface verre - eau est récupérée sur la caméra 1, tout comme le signal diffusé (non représenté sur la Figure 6.3.A) par chaque particule. Dans leur étude, Garmann et al. (2019) ont suivi l'assemblage de capsides de MS2 en temps réel comme le montre la Figure 6.3.B. Chaque tache noire visible sur le champ de vue correspond, *a priori*, à une capsidie partiellement ou totalement formée. Par ailleurs, l'étudiant ayant participé à ces travaux précise dans sa thèse que le contraste mesuré pour un virion natif de MS2 est proche de 0,8 % (Goldfain (2018)). En théorie, à partir de leurs conditions expérimentales ( $R = 15$  nm,  $\lambda = 450$  nm), on trouve environ 1,8 %. Cette différence s'explique notamment par l'efficacité de détection qui dépend du montage optique utilisé. Elle correspond aux pertes lumineuses pouvant avoir lieu tout au long de la propagation du faisceau (collection par l'objectif, réflexions parasites) et entraîne une perte d'information. Malgré cela, la Figure 6.3.B nous montre bien que le suivi de l'assemblage de capsides est possible.

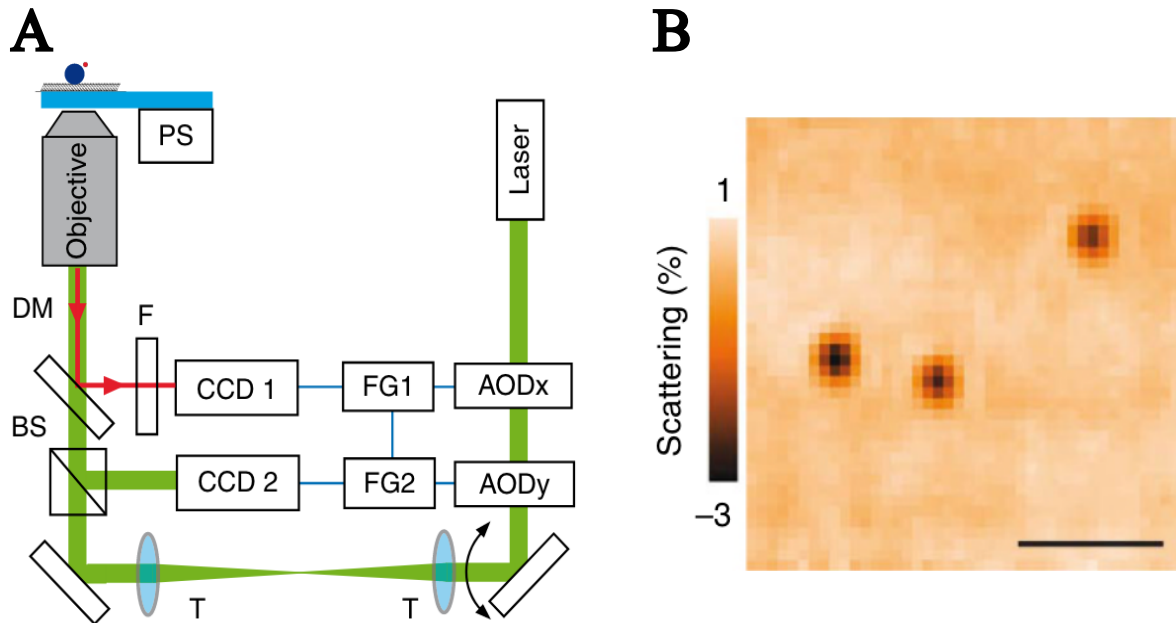


**FIGURE 6.3** – Observation de virus avec un montage iSCAT à partir d’une illumination en champ large. A) Schéma d’un montage iSCAT en champ large : le laser bleu participe à la microscopie interférométrique, le rouge est utilisé pour la stabilisation axiale du montage. B) Observation de l’auto-assemblage de capsides de MS2 en temps réel. Les figures sont adaptées de Garmann et al. (2019).

Le montage précédent est assez simple à réaliser puisque, sans tenir compte de la caméra, il n’utilise pas d’éléments optiques particuliers au niveau de leur prise en main. Par ailleurs, la réflexion au niveau de l’interface verre - eau permet d’atténuer le faisceau de référence et donc d’augmenter le contraste théorique. Cependant, comme l’ont observé Jacobsen et al. (2006), une illumination en champ large a tendance à détériorer la qualité de l’image. **En effet, des interférences parasites provenant de réflexions secondaires peuvent être observées sur le champ de vue. De plus, il est possible d’avoir une certaine inhomogénéité dans l’illumination. Pour améliorer ces deux aspects, certains montages utilisent un balayage ultra-rapide du faisceau laser (Figure 6.4.A).**

Dans cette configuration, Kukura et al. (2009) utilisent deux défecteurs acousto-optiques (*Acousto-optic deflector* ou AOD). L’onde acoustique générée au sein de ce composant permet de dévier le faisceau laser dans une direction. En utilisant deux AOD, il est possible d’agir sur la position du laser selon les axes X et Y et donc de réaliser le motif d’illumination souhaité au niveau de l’échantillon. Les auteurs ne fournissent pas le modèle des AOD mais Ortega Arroyo et al. (2016) mentionnent une période de balayage de  $12,5 \mu\text{s}$ . En comparaison, le temps d’exposition de la caméra utilisée étant supérieure à  $500 \mu\text{s}$ , il est possible de réaliser plusieurs fois le même motif pendant une image afin de réduire les fluctuations d’intensité. Sur la Figure 6.4.B, nous pouvons voir des virions natifs de SV40 observés dans cette configuration. Le contraste expérimental mesuré vaut

environ 3 % contre 3,5 % en théorie ( $R = 22,5$  nm et  $\lambda = 532$  nm). Le signal mesuré est plus élevé que pour un virion de MS2 du fait d'un diamètre plus grand. De plus, les pertes semblent légèrement moins importantes ici.



**FIGURE 6.4 – Observation de virus avec un montage iSCAT à partir d'une illumination à balayage.** A) Schéma d'un montage iSCAT à partir d'un balayage du laser : le faisceau vert participe à la microscopie interférométrique, la flèche rouge correspond à un signal de fluorescence. On note : AOD<sub>X</sub> et AOD<sub>Y</sub> les défecteurs acousto-optiques dans les directions X et Y, FG1 et FG2 les générateurs associés, BS la lame séparatrice, DM le miroir dichroïque, PS l'actuateur piézoélectrique, F le filtre passe-haut et CCD 1 et 2 les caméras *Charge-Coupled Device*. B) Observation de virus de SV40, barre d'échelle de 1  $\mu$ m. Les figures sont adaptées de Kukura et al. (2009).

En utilisant cette deuxième configuration, la contribution des réflexions parasites est atténuée, facilitant la détection de particules virales. **Le principal inconvénient concerne l'implémentation d'un tel montage. En effet, ce dernier nécessite de synchroniser la caméra, les AOD mais également le laser si l'iSCAT est couplé avec une autre technique optique.**

Les deux exemples illustrés dans cette section sont utilisés en réflexion. Il est également possible de travailler en transmission pour ne pas être gêné par les réflexions secondaires créées, entre autres, par l'objectif. Cependant, dans cette configuration le faisceau de référence est moins atténué que celui utilisé en réflexion, ce qui diminue le contraste théorique. De plus, le microscope TIRF que nous utilisons nécessite de coupler l'iSCAT en réflexion.

## 6.2 Stratégies de couplage avec la microscopie TIRF

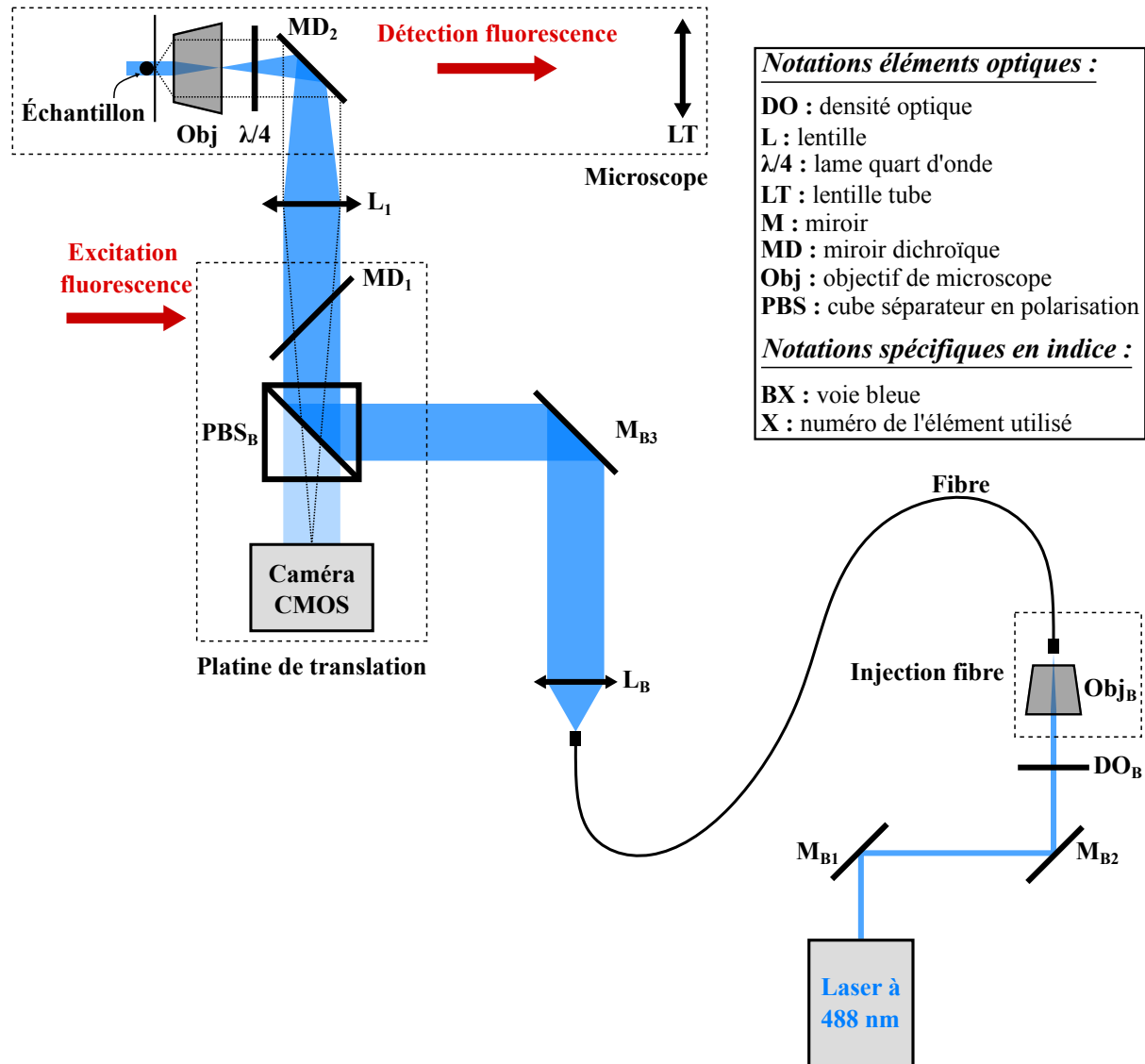
Cette partie a pour objectif de présenter le montage iSCAT que nous avons mis en place au laboratoire. Je montre également les premières images qui ont été obtenues dans cette configuration et les améliorations envisagées.

### 6.2.1 Choix d'une illumination en champ large

L'objectif premier était d'utiliser le protocole proposé par Ortega Arroyo et al. (2016) afin de coupler microscopies TIRF et iSCAT sur le même montage. Comme je l'ai dit à la section précédente, les auteurs utilisent deux AOD afin de balayer rapidement le laser au niveau de l'échantillon. En raison de contraintes temporelles, nous les avons remplacés par une matrice de micro-miroirs (*Digital Micromirror Device* ou DMD) qui nous a été fournie par une autre équipe du LuMIn. Ce système électromécanique est composé de centaines de milliers de miroirs micrométriques qui peuvent tous s'orienter selon deux positions afin de projeter un laser dans une direction souhaitée. En contrôlant individuellement chaque miroir, il est donc possible de déplacer le faisceau dans un plan donné et donc de former un motif d'illumination comme avec les AOD. Cependant, le fait que le DMD se comporte comme un réseau de diffraction a entraîné une modification du profil spatial du laser. En effet, dans le plan de l'échantillon, le profil du laser est devenu linéaire, au lieu de gaussien, avec une taille variable en fonction de la puissance nominale choisie. Ce problème étant apparu pour plusieurs lasers, nous avons donc mis de côté la configuration par balayage pour nous orienter sur une illumination en champ en large. Le montage actuel est schématisé par la Figure 6.5.

Ici, nous utilisons un laser de longueur d'onde 488 nm ayant une puissance réglable pouvant aller jusqu'à 100 mW (*Oxxius*, LBX 488). Le faisceau est dirigé jusqu'à une roue à densités DO dont l'objectif est d'imposer un second contrôle sur la puissance incidente. Le laser est ensuite injecté dans une fibre monomode (*Thorlabs*, P1-488PM-FC-1) afin d'améliorer le profil du faisceau. En sortie de fibre, le laser est collimaté à l'aide de la lentille  $L_B$  et le diamètre obtenu mesure environ 8 mm. De plus, en sachant que la lumière émergente du laser est polarisée  $s$ , le cube PBS permet de réfléchir le faisceau vers le miroir dichroïque noté  $MD_1$  (*Thorlabs*, DMSP550L). Ce dernier est très important dans le montage puisque c'est lui qui permet de coupler microscopies TIRF et iSCAT. En effet, il s'agit d'un filtre passe-bas qui transmet toutes les longueurs d'onde inférieures à 550 nm et réfléchit les autres. Il remplace donc le miroir noté  $M_2$  dans le montage du TIRF décrit à la Figure 3.4. Ensuite, le faisceau d'excitation est focalisé dans le plan focal arrière de l'objectif avec la lentille  $L_1$  (focale de 300 mm). Auparavant, le laser est réfléchi par le miroir dichroïque noté  $MD_2$  (*Chroma*, ZT405/488/561/640rpcv2-UF1) puis la polarisation est modifiée avec la lame quart d'onde. Les axes de cette dernière sont à

45° de la polarisation incidente  $s$ , ce qui nous donne une polarisation circulaire en sortie. Enfin, le faisceau est de nouveau collimaté grâce à l'objectif Olympus dont l'ouverture numérique est de 1,5. Le système utilisé possède un grandissement  $\times 100$  (pour une lentille tube de 180 mm, le grandissement de l'objectif est  $\times 60$ ), ce qui donne un diamètre du laser au niveau de l'échantillon d'environ 80  $\mu\text{m}$ .



**FIGURE 6.5** – Schéma du montage iSCAT en champ large couplé au TIRFM. Les pointillés noirs qui partent de l'échantillon et vont vers la caméra correspondent à la lumière diffusée par la particule observée. Le faisceau collimaté représenté en bleu clair et qui arrive sur la caméra indique la lumière réfléchie par l'interface verre - eau.

Sur le trajet retour, le faisceau réfléchi par l'interface suit le même chemin jusqu'au cube PBS. Toutefois, la réflexion ayant entraîné un changement dans la polarisation (circulaire gauche à droite), le passage par la lame quart-d'onde conduit à une polarisation linéaire  $p$  en amont du cube, qui est donc transmise jusqu'à la caméra CMOS (*Complementary Metal Oxide Semiconductor*). Cette dernière (*Photon Focus*, MV1-D1024E-

160-CL-12) est caractérisée par un capteur de taille 1024x1024 pixels de 10,6  $\mu\text{m}$  de côté chacun. En considérant le grandissement du système optique, cela correspond à un champ de vue total, au niveau de l'échantillon, d'environ 110x110  $\mu\text{m}^2$ . Pour finir, en parallèle du signal réfléchi, celui diffusé par les particules est focalisé au niveau de la caméra CMOS.

Remarques :

- En premier lieu, la lame quart d'onde devait être positionnée à un endroit où le faisceau est collimaté, afin de se placer dans des conditions de fonctionnement optimales. Cependant, nous nous sommes rendus compte que les miroirs dichroïques MD<sub>1</sub> et MD<sub>2</sub> transforment la polarisation circulaire en linéaire! Nous avons donc dû la placer après MD<sub>2</sub>, pour que la polarisation circulaire soit conservée jusqu'à l'échantillon.
- Pour limiter les perturbations de la ventilation, le montage est protégé par une boîte.

### 6.2.2 Les premières images

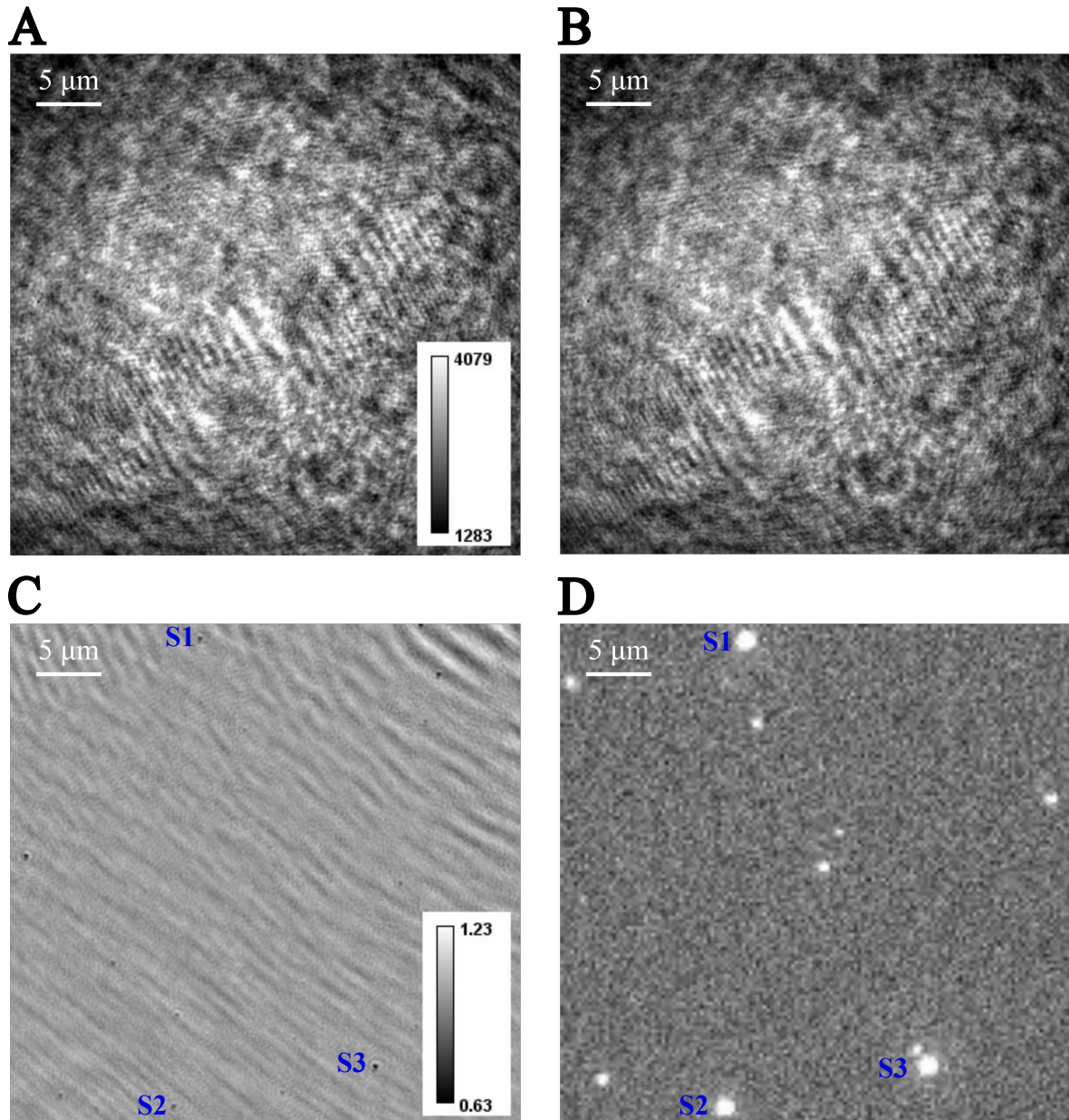
Désormais, je souhaite montrer les premières images obtenues en microscopie iSCAT. Pour ces tests, j'ai utilisé des sphères fluorescentes de 40 nm de diamètre et dont l'indice de réfraction vaut 1,52. L'avantage de ces billes est qu'elles peuvent également être observées en microscopie de fluorescence grâce au laser d'excitation à 638 nm. L'objectif est donc de voir si les mêmes particules sont visibles avec les deux techniques optiques.

En iSCAT, nous avons enregistré un film de 400 images avec une fréquence d'acquisition de 100 Hz et l'une des images obtenues est illustrée en Figure 6.6.A. Il apparaît assez nettement que l'image brute n'est pas exploitable en l'état. En effet, l'illumination n'est pas uniforme sur le champ de vue étudié, des motifs assez particuliers sont visibles tout comme des franges d'interférence parasites. Cependant, l'avantage est que ces différents éléments sont invariants dans le temps et qu'ils proviennent du microscope en lui-même et non de l'échantillon. Ortega Arroyo et al. (2016) proposent une étape de traitement qu'ils nomment *flat-fielding* afin d'améliorer la qualité de l'image. Le but est d'acquérir un film pendant que l'échantillon est translaté dans une direction puis de calculer la médiane temporelle en chaque pixel. Dans notre cas, une rampe de tension est envoyée à la platine piézoélectrique sur la voie Y afin de parcourir un aller-retour d'environ 40  $\mu\text{m}$  en une seconde. La Figure 6.6.B montre le résultat de la médiane temporelle réalisée sur 100 images acquises à 100 Hz, sur le même champ de vue que précédemment.

Enfin, chaque image brute est divisée par la médiane temporelle calculée à l'aide du protocole de *flat-fielding*. Le résultat de cette normalisation peut être observé sur la Figure 6.6.C. Nous voyons que le fond est bien plus uniforme que sur l'image brute, malgré la persistance de certaines franges. De plus, des taches circulaires noires sont visibles en différents endroits du champ de vue. Pour savoir s'il s'agit bien de sphères fluorescentes, il faut utiliser l'image acquise en TIRFM qui est représentée en Figure 6.6.D. Sur ce champ de vue, trois objets ont attiré notre attention et sont indiqués par S1 - S2 et S3. En effet,



sur les deux figures d'intérêt, nous observons le même motif triangulaire que forment les trois objets. Pour vérifier qu'il s'agit bien des mêmes particules, il faut utiliser le facteur de grandissement de 2,53 qui permet de passer d'une distance mesurée sur la caméra de fluorescence à la même longueur évaluée en iSCAT. Par exemple, entre les objets S1 et S2, la distance est de 136,4 pixels en TIRF et 340,1 pixels en iSCAT, ce qui donne un rapport de 2,49. De plus, pour S1-S3 et S2-S3, on trouve respectivement 2,50 et 2,51. **Il est donc clair que les trois particules S1, S2 et S3 sont bien des sphères fluorescentes.**



**FIGURE 6.6 – Observation expérimentale de sphères fluorescentes en iSCAT et TIRFM.** A) Image brute acquise en iSCAT. B) Médiane temporelle réalisée à partir d'un film iSCAT de 100 images pendant lequel l'échantillon est translaté. C) Image brute iSCAT normalisée par la médiane temporelle. D) Image brute acquise en microscopie TIRF. Les notations S1, S2 et S3 montrent des sphères fluorescentes.

Toutefois, nous pouvons remarquer qu'il y a des points vus en TIRF qui ne semblent pas être visibles en iSCAT et inversement. Pour le premier cas, il est intéressant d'évaluer le contraste mesuré en microscopie interférométrique. En traçant une coupe d'intensité sur les trois sphères, on trouve des valeurs de 21,9 - 16,7 - 33,9 % respectivement pour S1, S2 et S3. Or, en théorie, le contraste d'une sphère de 40 nm de diamètre et d'indice 1,52 pour une observation avec un laser de 488 nm est d'environ 5,3 %. **Cela signifie que, pour le moment, nous observons des agrégats de sphères, et dans ce cas-ci, au minimum 3 objets pour S2, ce qui correspond environ à une centaine de nanomètres.** Il est donc cohérent que les particules uniques qui émettent moins d'intensité sur la Figure 6.6.D ne sont pas visibles en iSCAT. Celles-ci doivent être cachées dans les fluctuations d'intensité notamment causées par les franges toujours présentes sur le champ de vue.

Pour le second cas de figure, les objets visibles en iSCAT mais pas en TIRF mettent en évidence **la principale limitation de la microscopie interférométrique : son manque de spécificité.** En effet, comme le précise Ortega Arroyo et al. (2016), sans contrôle supplémentaire, il n'est pas possible de différencier des objets nanométriques de taille similaire. La fluorescence est donc importante afin de différencier une particule d'intérêt d'une impureté.

### 6.2.3 Les améliorations envisagées

Pour conclure, je souhaite présenter certaines des améliorations auxquelles nous avons pensé. Ces dernières concernent principalement deux aspects : l'augmentation du contraste et l'amélioration de la qualité des images.

Dans le premier cas, puisque la taille du CCMV, son indice de réfraction et celui du tampon d'assemblage sont fixes, il ne reste qu'un seul paramètre qui peut être ajusté dans l'équation 6.3 : la longueur d'onde. En effet, la section efficace de diffusion varie en  $\frac{1}{\lambda^4}$ , ce qui peut avoir une influence non négligeable sur le contraste final. **Il peut donc être envisagé de passer d'un laser de longueur d'onde 488 nm à 405 nm, ce qui conduirait à un contraste théorique de 2,5 %, soit une augmentation de près d'un facteur 2 par rapport à la valeur précédente (1,4 %).** Néanmoins, l'utilisation d'une source lumineuse très proche des ultraviolets peut participer à la dégradation des échantillons biologiques. Il ne faudrait donc pas utiliser des temps d'exposition trop longs pour ne pas favoriser ce mécanisme.

Une seconde amélioration correspond à l'utilisation d'un masque spatial afin d'améliorer le contraste mesuré. En effet, Cole et al. (2017) montrent que le fait de placer cet élément absorbant dans le plan focal arrière de l'objectif, *i.e.* là où le faisceau incident est focalisé, permet d'atténuer la voie d'illumination et donc le signal de référence. Par

ailleurs, sachant que ce masque a un diamètre légèrement supérieur à celui du faisceau lorsqu'il est focalisé, l'élément absorbant n'altère que très peu le signal diffusé par les particules. **Dans leur étude, les auteurs montrent que cette configuration permet d'améliorer le contraste d'un facteur 10.** Cependant, en même temps qu'elle rend possible l'observation de particules d'environ 400 kDa, elle augmente l'intensité provenant du fond et notamment des aspérités de la surface de verre. Il est donc important de pouvoir efficacement atténuer cette contribution dans le but de repérer les particules d'intérêt.

En ce qui concerne l'amélioration de la qualité de l'image, nous pouvons nous pencher sur le traitement des franges toujours visibles après l'étape de *flat-fielding*. Ces dernières ne peuvent pas être corrigées simplement en appliquant un filtre gaussien ou équivalent puisqu'elles bougent dans le temps. Néanmoins, nous avons récemment remarqué que les fluctuations de ces franges sont causées par la géométrie des cellules micro-fluidiques. En effet, en déposant un échantillon sur une lamelle seule, elles étaient toujours visibles mais ne bougeaient plus. Cela confirme ce que précise Goldfain (2018) dans sa thèse : pour éviter d'avoir des interférences parasites, il faut que l'écart entre lame et lamelle soit de plusieurs centaines de micromètres. Or, il est inférieur à 200  $\mu\text{m}$  pour une cellule micro-fluidique complète. Il est donc important de réaliser de nouveaux tests avec une lamelle seule pour voir si les franges restantes peuvent être retirées numériquement par un filtre gaussien, tout en conservant le signal issu des particules d'intérêt. Si c'est le cas, nous réfléchirons à une façon d'augmenter l'épaisseur des cellules micro-fluidiques. **Par ailleurs, Dulin et al. (2014) ont montré qu'en modulant le courant d'une diode laser, la cohérence temporelle de cette dernière est réduite. Cela permet ainsi d'éliminer des interférences parasites tout en conservant le signal des objets étudiés.** Ces deux améliorations ont l'avantage de pouvoir être testées assez facilement sans avoir besoin de modifier en profondeur le montage.

Une dernière option consiste à changer totalement de configuration expérimentale en remplaçant l'illumination en champ large par celle réalisée à partir d'un balayage du faisceau laser. Il serait alors nécessaire de prendre en main le pilotage des déflecteurs acousto-optiques et la synchronisation de ces derniers avec la caméra et le laser.

# Chapitre - 7

---

## Conclusions et perspectives

---

Ce manuscrit présente les travaux que j'ai menés durant trois ans au sein du LPS et du LuMIn, dans le but de suivre l'assemblage de capsides du CCMV en microscopie de fluorescence par réflexion totale interne (TIRFM) à l'échelle de la molécule unique. L'utilisation de cette méthode se justifie par le fait que la grande majorité des techniques employées pour caractériser le mécanisme d'auto-assemblage de virus donne des résultats moyennés sur toute une population. Un des objectifs était donc de pouvoir observer l'hétérogénéité de ce processus en visualisant des sites d'assemblage uniques. De plus, comme nous l'avons vu à plusieurs reprises au sein de ce manuscrit, le fait de marquer les dimères de protéines par un fluorophore nous a permis de travailler à l'échelle de l'accroche unique. Ces événements nous ont donné accès à des informations temporelles sur la dynamique d'assemblage des sous-unités de capside du CCMV.

Avant de pouvoir mettre en place les expériences correspondantes, une phase d'optimisation du protocole de fonctionnalisation des lamelles de verre a été nécessaire comme le montre la troisième partie du manuscrit. De nombreuses tentatives se sont succédé afin de réduire le plus possible les accroches non spécifiques en dimères fluorescents tout en permettant l'adsorption spécifique de l'ARN viral marqué en certains endroits de la surface de verre. L'architecture ainsi obtenue est assez complexe et peut expliquer pourquoi de temps en temps, l'un des deux objectifs a été moins bien atteint. Néanmoins, ce protocole, ainsi que les différentes expériences de contrôle, nous ont assuré que la grande majorité des événements d'accroche en sous-unités fluorescentes sur les ARN simples brins traduit bien des interactions spécifiques.

Suite à cette première phase, des expériences de cinétique d'assemblage ont pu être réalisées. En travaillant hors équilibre, il est apparu que les premières accroches en dimères marqués sur leur matériel génétique sont facilitées et se succèdent de plus en plus en

---

rapidement quand la concentration en sel de la solution tampon augmente. Par ailleurs, l'évolution temporelle du nombre moyen de sous-unités fluorescentes par site d'assemblage nous indique qu'un équilibre global est atteint plus tôt pour des interactions plus faibles entre ARN et dimères. Il est néanmoins difficile d'évoquer un équilibre thermodynamique à proprement parler puisque le photoblanchiment altère grandement nos mesures ici. Ces résultats, obtenus notamment en nous éloignant des concentrations physiologiques en sel, doivent désormais être accompagnés par de nouvelles expériences afin de mieux comprendre l'évolution des interactions électrostatiques en fonction des conditions utilisées. Cela nous permettrait de mieux appréhender les différences observées avec le BMV par exemple.

Afin d'avoir accès plus facilement à la dynamique temporelle des accroches en dimères, le protocole d'acquisition a été orienté vers une étude à l'équilibre. Dans ces conditions, nous avons pu appliquer l'algorithme Ruptures sur les signaux de fluorescence pour obtenir un ajustement en marches d'escalier des échanges entre sous-unités marquées au niveau d'un ARN. En amont de son application expérimentale, les performances de Ruptures, évaluées sur des signaux simulés selon l'équilibre thermodynamique attendu en expérience, nous ont donné des indications sur le niveau de confiance que nous pouvions avoir vis-à-vis des prédictions de l'algorithme. Ce dernier tend à manquer plus d'événements qu'il n'en repère de faux, ce qui réduit le nombre d'échanges pris en compte, mais semble mieux retranscrire les événements simulés à l'équilibre que le second algorithme testé, quickPBSA. De plus, nous avons mesuré l'influence des clignotements sur le calcul de paramètres d'assemblage tels que la constante d'association. Sans surprise, l'ajout artificiel d'événements montants a entraîné une sous-estimation par Ruptures du taux d'accroches en dimères marqués, et ce d'environ 20 % en considérant les conditions de simulations les plus proches de l'expérience.

En appliquant Ruptures sur des données expérimentales, nous avons obtenu des premières estimations en termes de dynamique temporelle d'arrivée des dimères sur leur ARN viral. En effet, pour une concentration totale en protéines de 50 nM, nous avons mesuré une accroche en sous-unité toutes les six secondes et un temps de résidence moyen supérieur à deux minutes. Dans ces conditions, nous avons estimé la présence d'un peu plus de vingt dimères par ARN. De plus, lorsque la concentration totale en protéines croît, la taille moyenne des particules observées augmente ce qui se traduit également par une diminution du taux d'accroches en dimères. Cette évolution montre bien que le matériel génétique est rendu moins accessible quand les particules se rapprochent de nucléocapsides. Dans le but d'être plus quantitatives, les expériences de dynamique d'assemblage à l'équilibre nécessitent encore quelques optimisations, notamment lorsque le taux de marquage des protéines diminue.

En ce qui concerne les prochaines améliorations envisagées, la priorité concerne la réduction des clignotements. En effet, ce mécanisme photophysique, de fréquence très variable, altère directement la mesure de la constante d'association et par extension, l'estimation du temps de résidence moyen. Une solution envisagée est de modifier le tampon d'assemblage en utilisant un système oxydo-réducteur désoxygénant. Les premiers tests réalisés ont bien montré une limitation du photoblanchiment comme ce qui était attendu, mais aucune amélioration n'a été constatée vis-à-vis des clignotements. D'autres protocoles vont être appliqués prochainement pour éliminer totalement les passages en OFF de l'Atto647N. En empêchant l'apparition des clignotements, cela nous permettrait sans doute d'être plus quantitatifs dans l'estimation de la constante d'association des dimères de protéines. Par ailleurs, une amélioration des performances de Ruptures pourrait également nous aider. Les pistes discutées avec Charles Truong consistent à rajouter la hauteur moyenne d'une marche dans l'ajustement réalisé par l'algorithme pour que ce dernier ait une estimation du nombre de sous-unités marquées en temps réel pour chaque signal. Ruptures pourrait utiliser cette information pour corriger ses prédictions s'il s'avère que le nombre de dimères marqués devient négatif au cours de l'ajustement.

De plus, nous avons remarqué que les valeurs du nombre moyen de sous-unités par ARN pouvaient être dispersées, surtout en utilisant un taux de marquage de 2,5 %. En effet, dans le but de limiter les accroches non spécifiques en dimères marqués, nous avons fixé la concentration maximale de ces derniers à 10 nM, ce qui nous oblige à réduire le taux de marquage dès que la concentration totale en protéines est supérieure à 100 nM. Une solution envisagée consiste à travailler avec deux lots de dimères marqués avec deux fluorophores différents. Cependant, le seul essai réalisé à partir d'un autre marqueur que l'Atto647N s'est avéré être un échec puisque un faible taux de marquage a été obtenu en fin de purification, probablement à cause d'un problème d'interactions électrostatiques. Ainsi, pour améliorer notre précision dans l'estimation du nombre de dimères par site d'assemblage, nous avons commencé à mettre au point un montage de microscopie à diffusion interférométrique (iSCAT). Cette technique, qui se base notamment sur la lumière diffusée par l'ensemble des sous-unités, doit être couplée à notre microscope TIRF avec pour objectif de colocaliser les sites d'assemblage observés par les deux méthodes, et donc d'avoir deux estimations de la taille des particules étudiées. Pour le moment, les premières images que nous avons obtenues nous indiquent que nous arrivons à observer des agrégats de sphères fluorescentes d'une centaine de nanomètres. Une meilleure stabilité du montage ainsi qu'une optimisation du traitement d'images devraient nous permettre d'observer des virions uniques, puis des capsides partiellement assemblées. Ainsi, les améliorations conjointes dans le calcul de la constante d'association et du nombre moyen de dimères par ARN nous aideraient à affiner les estimations du temps de résidence moyen des sous-unités.

---

Enfin, il est important de se pencher également sur l'influence du marquage. En effet, même si la formation de nucléocapsides à partir de dimères marqués n'a pas été remise en cause par la photométrie de masse, les particules étudiées ont une masse plus faible que dans des conditions sans marquage. Réaliser de nouvelles expériences de photométrie de masse, avec d'autres conditions expérimentales, pourrait nous permettre d'avoir une idée globale sur l'influence du marquage, mais nous n'aurons pas la confirmation absolue que des dimères marqués sont présents, ou non, sur les objets observés. À l'inverse, le montage iSCAT que nous mettons en place, semble être la meilleure solution pour pleinement évaluer l'influence du marquage sur le mécanisme d'auto-assemblage. Cela nous permettrait également d'affirmer que l'Atto647N ne joue pas un rôle dans la cinétique d'assemblage étudiée hors équilibre pour le CCMV.

Les différentes optimisations proposées ici devraient nous permettre d'être plus quantitatifs dans nos estimations du nombre moyen de sous-unités par site d'assemblage, de leur constante d'association ou temps de résidence sur l'ARN simple brin. À plus long terme, il serait intéressant d'appliquer les expériences de microscopie de fluorescence, combinées à l'iSCAT, sur le suivi du mécanisme d'auto-assemblage en milieu encombré. L'objectif consisterait à se rapprocher petit à petit des conditions rencontrées au sein d'une cellule eucaryote afin de comprendre comment l'encombrement affecte la diffusion des sous-unités et donc la formation de capsides. Cette volonté pourrait également être l'occasion de réaliser des expériences de compétition entre différents ARN (viraux ou non) dans le but de comparer les cinétiques et dynamiques d'assemblage. Enfin, nous envisageons d'appliquer le protocole expérimental mis en place en microscopie de fluorescence à d'autres virus étudiés dans la littérature tels que le BMV ou HBV.

## Annexe A

# Des virions natifs de CCMV aux sous-unités de capsid marquées

C'est au sein de l'équipe SOBIO (Structure et Dynamique d'Objets Biologiques Auto-assemblés) du LPS que j'ai pu me familiariser avec la préparation du CCMV. Durant ces trois années, j'ai participé aux étapes suivantes : plantation des graines, infection des plantes, purification des virions et enfin purification des protéines. Dans cette annexe, je ne vais décrire que les deux derniers protocoles de purification puisque les étapes de plantation et infection peuvent être retrouvées dans la thèse de Maelenn Chevreuil (Chevreuil (2020)). La Figure A.1 montre le changement d'aspect des feuilles une fois l'infection réalisée : les taches jaunes sont le signe d'une infection réussie.



**FIGURE A.1** – Feuilles de cornille. Avant (A) et après infection (B).

Dans le Tableau A.1, je donne la composition des solutions tampons utilisées au cours des protocoles de purification des virions et des protéines. Par ailleurs, à plusieurs reprises je vais parler de mesures d'absorbance. Ces étapes sont réalisées avec un spectrophotomètre (*Nanodrop, ThermoFisher*) qui dispose de plusieurs modes dont notamment : "Nucleic acid", "Protein  $A_{280}$ " et "UV-Visible". La première option est utilisée pour calculer les concentrations en virions et en ARN, alors que la seconde est préférée lorsque l'on souhaite estimer la concentration de protéines non marquées. Le mode "UV-Visible"



remplace les précédents dès que les solutions de virions ou de protéines sont marquées afin de couvrir le spectre d'absorption du marqueur fluorescent.

NOM	COMPOSITION	pH
Tampon A	150 mM d'acétate de sodium	4,8
Tampon Vir	50 mM d'acétate de sodium	4,8
Tampon M	100 mM HEPES, 5 mM MgCl <sub>2</sub>	8,2
Tampon D	50 mM HEPES, 500 mM CaCl <sub>2</sub> , 1 mM EDTA (pH = 8), 1 mM DTT et 0,5 mM PMSF	7,5
Tampon Dim	50 mM HEPES, 500 mM NaCl, 1 mM EDTA (pH = 8), 1 mM DTT et 0,5 mM PMSF	7,5

**TABLEAU A.1 – Composition des tampons de purification.**

## Purification des virions :

Les étapes décrites ci-dessous ont été adaptées du protocole utilisé par Ali and Roosinck (2007) et sont valides pour une masse initiale de feuilles infectées de 140 g :

**1** - Broyer les feuilles infectées à l'aide d'un mixeur tout en incorporant au fur et à mesure 300 mL de tampon A.

**2** - Filtrer le mélange tout en ajoutant 300 mL de chloroforme glacé avant de laisser agiter pendant une heure en chambre froide.

**3** - Verser la solution dans des bouteilles de centrifugation (volume 250 mL) puis lancer le cycle suivant : 10 minutes, 4°C, 7800 rotations par minute (rpm) dans un rotor GSA.

**4** - Une fois la centrifugation terminée, on distingue trois phases : un culot solide, une première phase liquide verte au fond de la bouteille et une seconde phase liquide jaunâtre. Récupérer la phase liquide jaunâtre puis y incorporer, sous agitation, 8 g de PEG 8000 Da et 115 mg de NaCl pour 100 mL. Laisser agiter pendant une heure en chambre froide.

**5** - Verser de nouveau la solution dans des bouteilles de centrifugation puis lancer le même cycle qu'à l'étape **3**.

**6** - À la fin de la centrifugation, on distingue deux phases : un culot solide blanc et un liquide jaunâtre. Éliminer le liquide jaunâtre, re-suspendre les culots des différentes bouteilles de centrifugation dans 21 mL de tampon Vir au total puis laisser agiter en chambre froide pendant une heure.

**7** - Verser la solution précédente dans des petites bouteilles de centrifugation (volume 30 mL) puis lancer le cycle suivant : 10 minutes, 4°C, 8200 rpm dans un rotor SS-34.

**8** - Après la centrifugation, on distingue de nouveau deux phases : un culot sombre et une phase liquide assez claire. Récupérer uniquement la phase liquide et la répartir dans des tubes d'ultra-centrifugation (volume 13 mL). Ajouter 2 mL d'une solution de

sucrose au fond de chaque tube, équilibrer ces derniers au milligramme près en ajoutant du tampon Vir si nécessaire puis lancer le cycle suivant : 2 heures, 4°C, 35 000 rpm dans un rotor SW41 Ti.

**9** - Dès que l'ultra-centrifugation est terminée, repérer le culot brun au fond du tube et éliminer rapidement la phase liquide. Resolubiliser chaque culot dans 1 mL de tampon Vir.

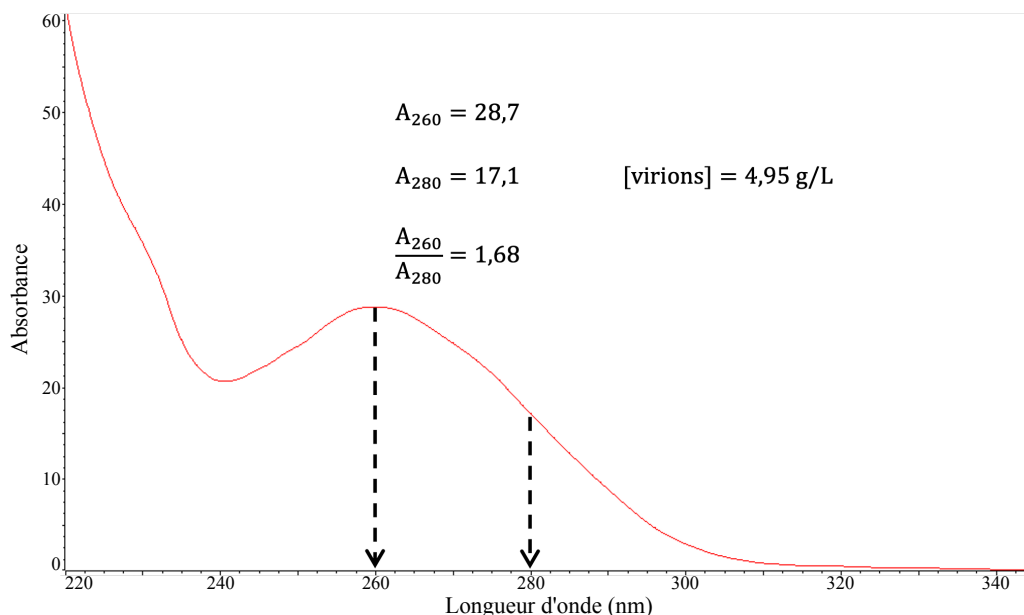
**10** - Pour attester de la pureté de la purification, mesurer l'absorbance à 260 nm (signal caractéristique du matériel génétique) et à 280 nm (signal typique pour les protéines). Le rapport  $\frac{A_{260}}{A_{280}}$  est calculé et doit être compris entre 1,5 et 1,7 (Michel et al. (2004)). Ensuite, faire une estimation de la concentration en virions en divisant l'absorbance à 260 nm par 5,8. Si la concentration en virions est supérieure à 5 g/L, diluer la solution avec du tampon Vir pour éviter les risques d'agrégation. La Figure A.2 montre une courbe expérimentale d'absorbance suite à une purification de virions de CCMV.

**11** - Une fois la pureté et la concentration contrôlées, faire des aliquots de 500  $\mu$ L ou 1 mL et les stocker à -80°C.

Remarques :

- Pour 140 g de feuilles, on estime que la purification présente un bon rendement si la masse de virions est supérieure à 10 mg.

- Les volumes des deux premières étapes peuvent être ajustés en fonction de la masse de feuilles utilisée mais ceux donnés à partir de l'étape **6** seront conservés quelles que soient les conditions initiales.



**FIGURE A.2** – Absorbances d'une solution de virions. Les valeurs d'absorbance sont données pour une longueur de cuve de 1 cm et les flèches indiquent les absorbances à 260 et 280 nm.

---

## Purification des protéines :

Le deuxième protocole de purification a pour but de marquer les protéines de CCMV et de désassembler les virions afin de ne collecter que les protéines. La stratégie de marquage utilisée est décrite plus en détails dans la section 3.3.4. Par ailleurs, ce protocole est effectué le plus souvent à partir d'une solution de virions de 1 mL concentrée à 4,9 g/L :

**1** - Décongeler le ou les aliquot(s) de virions puis verser la solution dans deux à trois cellules de dialyse pouvant contenir un volume maximal de 500  $\mu$ L (*Thermo Scientific*, Slide-A-Lyzer MINI Dialysis Units, 3 500 MWCO (*Molecular Weight Cutoff*)). Laisser dialyser sur la nuit (au moins 12 heures) dans 500 mL de tampon M en chambre froide.

**2** - Mesurer l'absorbance de la solution à 260 nm et calculer la concentration en virions afin d'estimer les éventuelles pertes.

**3** - Marquer les virions : on souhaite marquer 25 % des protéines de CCMV avec un Atto647N-NHS-Ester (*ATTO-TEC*) de concentration initiale notée  $C_{Atto}$  (2,37 mM habituellement). Pour ce faire, il faut d'abord calculer la concentration massique en protéines en multipliant celle en virions par 0,75 puisque nous avons vu à la section 2.2.1 que les protéines correspondent à 75 % de la masse molaire d'un virus complet. Le passage à une concentration molaire en protéines, notée  $C_{prot}$ , s'effectue en utilisant la masse molaire d'un monomère qui est estimée à 20 300 Da. On peut donc calculer la concentration molaire en protéines que l'on souhaite marquer en multipliant  $C_{prot}$  par 0,25. Le but est désormais de déterminer le volume  $V_{Atto}$  à injecter dans une solution de volume  $V_{vir}$ . Comme cette étape correspond à une dilution du marqueur fluorescent et que la quantité de matière se conserve au cours de ce processus, on utilise la relation suivante :

$$V_{Atto} = \frac{V_{vir} \times C_{prot} \times 0,25}{C_{Atto}}$$

Cette équation est écrite généralement sous la forme  $C_i V_i = C_f V_f$  avec i l'état initial et f l'état final. Ici, le volume final est approximé au volume de la solution de virions vu que ce dernier est de l'ordre du millilitre contre un volume à injecter en marqueur d'une dizaine de microlitres. Après avoir effectué ce calcul, injecter le volume de marqueur correspondant, homogénéiser la solution et laisser incuber à température ambiante pendant 2 heures en entourant le tube d'aluminium. En effet, la solution marquée est protégée de la lumière afin de ne pas endommager le fluorophore.

**4** - Répartir la solution marquée dans deux à trois cellules de dialyse de 500  $\mu$ L et laisser dialyser dans 500 mL de tampon M pendant 24 heures en chambre froide. Cette étape a pour but d'éliminer les marqueurs libres en solution.

**5** - Après la dialyse, mesurer l'absorbance à 260, 280 et 647 nm afin d'estimer la concentration molaire en protéines  $C_p$  et le taux de marquage  $\tau$  en utilisant les formules suivantes :

$$C_p = \frac{\varepsilon_{ARN}(280) \times [A_{260} - A_{647} \times CF_{260}] - \varepsilon_{ARN}(260) \times [A_{280} - A_{647} \times CF_{280}]}{\varepsilon_{ARN}(280) \times \varepsilon_p(260) - \varepsilon_{ARN}(260) \times \varepsilon_p(280)} \quad (\text{A.1})$$

$$\tau = \frac{A_{647}}{\varepsilon_{f,max} \times C_p} \quad (\text{A.2})$$

où :

-  $CF_{260/280} = \frac{\varepsilon_{260/280}}{\varepsilon_{f,max}}$  donne le coefficient de correction propre au marqueur fluorescent à 260 et 280 nm avec  $CF_{260} = 0,04$  et  $CF_{280} = 0,03$  pour l'Atto647N.

-  $\varepsilon_{f,max}$  correspond au coefficient d'extinction molaire du fluorophore à sa longueur d'onde maximale d'absorption, à savoir 646 nm ici.

-  $\varepsilon_{ARN/p}(260/280)$  décrit le coefficient d'extinction molaire de l'ARN (respectivement des protéines) à 260 et 280 nm. Le tableau A.2 donne les valeurs numériques des différents coefficients d'extinction molaire utilisés dans notre cas :

Notation	Valeur ( $\text{L.mol}^{-1}.\text{cm}^{-1}$ )
$\varepsilon_{f,max}$	$1,5 \times 10^5$
$\varepsilon_{ARN}(260)$	$2,5 \times 10^7$
$\varepsilon_{ARN}(280)$	$1,25 \times 10^7$
$\varepsilon_p(260)$	$1,68 \times 10^4$
$\varepsilon_p(280)$	$2,4 \times 10^4$

**TABLEAU A.2** – Valeurs des coefficients d'extinction molaire utilisés en purification.

L'équation A.1 est tirée des travaux de Porterfield and Zlotnick (2010). Dans leur article, ils précisent que les virus diffusent la lumière incidente du spectrophotomètre, ce qui entraîne une surestimation des absorbances mesurées. Comme une mesure précise de la concentration en protéines n'est pas nécessaire ici, je n'utilise pas la correction qu'ils proposent. En revanche, on corrige les absorbances à 260 et 280 nm en retirant la contribution du marqueur fluorescent dans cette plage de longueur d'onde (termes  $A_{647} \times CF_{260/280}$ ). À ce stade, on mesure généralement un taux de marquage d'environ 20 %.

**6** - Une fois les calculs effectués, verser la solution de virions marqués dans une cellule de dialyse pouvant contenir un volume maximal de 5 mL (*Spectra/Por*®<sup>®</sup>, Float-A-Lyzer G2 Dialysis Device, 3500-5000 MWCO) et compléter avec du tampon D jusqu'à 1 cm du bord supérieur. Placer cette cellule de dialyse dans une éprouvette contenant 500 mL de tampon D. Juste avant de débiter la dialyse, ajouter 77 mg de DTT (masse molaire de 154,25 Da) et 2,5 mL de PMSF à 100 mM fraîchement préparé dans une solution d'éthanol pur. Cette étape ne peut pas être faite lors de la préparation de la solution tampon puisque le PMSF et le DTT perdent rapidement leur propriété en solution aqueuse. Cette dialyse dite de dissociation doit durer 24 heures en chambre froide et permet de préparer le désassemblage : l'EDTA capture les ions divalents qui forment des ponts entre les protéines

---

les empêchant de se dissocier et le DTT empêche la création de ponts disulfure pouvant stabiliser la capsid.

**7** - À la fin de la dialyse, verser la solution marquée (environ 5 mL) dans un tube d'ultra-centrifugation pouvant contenir un volume maximal de 13,2 mL (*Beckman Coulter*®), Ultra-Clear™, Centrifuge Tubes 14 x 89 mm) et compléter avec du tampon D jusqu'à 1 cm du bord supérieur. Équilibrer le tube avec un contrepois au milligramme près puis lancer le cycle suivant : 18 heures, 4°C, 35 000 rpm dans un rotor SW41 Ti.

**8** - Dès la fin de l'ultra-centrifugation, collecter des fractions de 1 mL de la solution marquée et les répartir dans des tubes différents. La collecte des fractions s'effectue dans la glace. À ce stade, on récupère généralement 11 fractions de 1 mL et un culot bleu est présent au fond du tube d'ultra-centrifugation. Ce culot doit contenir la grande majorité des ARN puisque le CaCl<sub>2</sub> du tampon de dissociation a pour but de précipiter le matériel génétique du CCMV. Par ailleurs la couleur bleue indique également la présence d'Atto647N dans le culot. Il est possible qu'à cause du marquage, certaines capsides ont eu du mal à se désassembler ce qui favorise la précipitation des marqueurs associés.

**9** - Pour chaque fraction, mesurer le rapport  $\frac{A_{260}}{A_{280}}$  afin de vérifier la pureté de la solution. Pour des protéines non marquées, les fractions qui présentent un rapport inférieur à 0,65 peuvent être conservées alors que les autres doivent être jetées afin de diminuer le risque de contamination de la solution par des ARN natifs. En revanche, comme je l'ai mentionné à l'étape **5**, le marqueur fluorescent contribue à l'absorbance de la solution à 260 et 280 nm et nous empêche d'utiliser ce critère de sélection. Je me suis donc basé sur les précédentes purifications réalisées avec des protéines non marquées au cours desquelles les deux dernières fractions étaient éliminées (les plus proches du culot). Pour plus de sûreté, je décide d'éliminer les trois dernières fractions.

**10** - Rassembler les fractions conservées (environ 8 mL) dans un Falcon de dialyse pouvant contenir au un volume maximal de 12 mL (*Merck*, Amicon®Ultra-15, Centrifugal Filters, 10 000 NMWL (*Nominal Molecular Weight Limit*)) et rajouter du tampon Dim jusqu'à 1 cm du bord. Lancer un premier cycle de centrifugation de 15 minutes à 4°C et 4000 g. À la fin de cette étape, vider la partie inférieure du Falcon et compléter la solution de protéines marquées avec du tampon Dim jusqu'à 1 cm du bord. Refaire un second cycle de centrifugation, vider la partie inférieure du Falcon et compléter cette fois-ci avec 2 mL de tampon Dim. Lancer un dernier cycle de centrifugation. Cette étape permet de concentrer la solution en protéines en réduisant le volume à environ 1 mL.

**11** - Mesurer la concentration en protéines avec l'équation suivante :

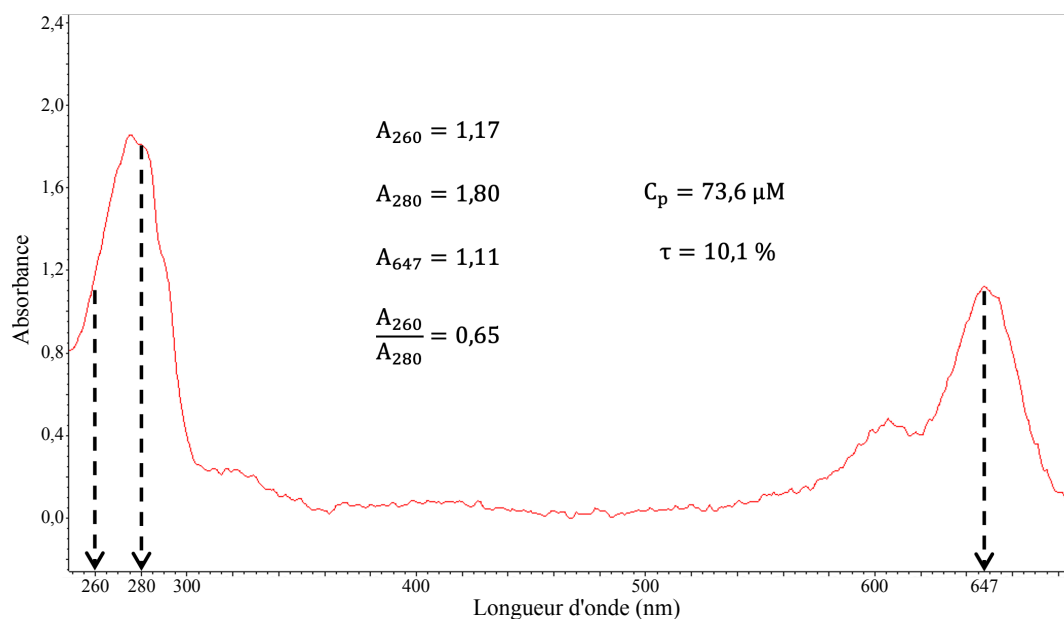
$$C_p = \frac{A_{280} - A_{647} \times CF_{280}}{\varepsilon_p(280)} \quad (\text{A.3})$$

Sachant que l'ARN a été éliminé, toutes les contributions liées à ce dernier ont été négligées. L'équation A.3 est donc une simplification de l'équation A.1. Le taux de marquage

peut être évalué à l'aide de l'équation A.2.

**12** - Répartir la solution de protéines dans deux à trois cellules de dialyse de 500  $\mu$ L et laisser dialyser dans 500 mL de tampon Dim sur la nuit (au moins 12 heures). Cette dernière étape de dialyse a pour but de remplacer les ions divalents par du sel monovalent.

**13** - Une fois la dialyse terminée, mesurer la concentration molaire en protéines et le taux de marquage en utilisant de nouveau les équations A.3 et A.2 puis stocker l'échantillon entouré d'aluminium au réfrigérateur. La Figure A.3 montre une courbe expérimentale d'absorbance suite à une purification de protéines de CCMV pour lesquelles 10 % sont marquées.



**FIGURE A.3** – Absorbances d'une solution de protéines marquées avec un Atto647N. Les absorbances sont données pour une longueur de cuve de 1 cm. Les flèches indiquent les absorbances à 260 - 280 et 647 nm.

*Remarques :*

- Pour 4,9 mg de virions, j'obtiens généralement entre 1,2 et 1,6 mg de protéines. Par ailleurs, le rapport  $\frac{A_{260}}{A_{280}}$  de la solution finale est souvent inférieur à 0,7.

- Pour un taux de marquage initial de 25 %, les protéines purifiées sont finalement marquées à hauteur de 10 à 15 %. Avant mon arrivée en thèse, des tentatives pour marquer 100 % des protéines ont été effectuées, ce qui a conduit à une impossibilité de désassembler les virions marqués. Ainsi, pour ne pas trop perturber le comportement des virions / protéines, nous n'avons pas souhaité viser un taux de marquage initial supérieur à 25 %.

- Les protéines sont stockées au plus 3 mois au réfrigérateur.

- Durant les étapes de dialyse, les béchers et les éprouvettes sont entourés d'aluminium une fois que la solution est marquée pour limiter le photoblanchiment.

- Le pH du tampon M est volontairement choisi supérieur à 7,5 puisque le fournisseur

---

du marqueur fluorescent, ATTO-TEC GmbH, conseille de travailler à un pH compris entre 8 et 9 pour un marquage efficace.

- La purification de protéines marquées se déroule généralement en parallèle de la purification d'un lot de protéines non marquées. Ce second lot est utilisé pour ajuster le taux de marquage choisi durant les expériences d'assemblage. La deuxième purification démarre à l'étape **6** avec la dialyse de dissociation et suit le même protocole jusqu'à la fin. Les bains de dialyse sont différents pour chaque lot de protéines et le calcul final de la concentration en protéines non marquées est effectué avec l'équation A.3 en posant  $A_{647} = 0$ .

## Annexe B

# Synthèse de l'ARN C2 du CCMV

Les expériences détaillées au Chapitre 5 se basent toutes sur un assemblage des protéines de capsid autour de l'ARN C2 du CCMV. La séquence composée de 2774 nucléotides (*référence* : NC – 00351) est présente dans le plasmide ppC2 (8 ng/ $\mu$ L) qui nous a été fourni par William Gelbart puis transcrit à l'aide d'un kit commercial *MEGAscript*<sup>®</sup> *Kit*. Il est important de s'assurer que le plasmide utilisé possède le promoteur, *i.e.* la séquence de nucléotides correspondant à un site d'accroche, de l'ARN polymérase fourni par le kit. Parmi les ARN polymérases proposées par le kit, nous travaillons généralement avec la T7. Le protocole complet est décomposé selon les trois parties suivantes :

### I - Linéarisation du plasmide :

Il est conseillé de linéariser le plasmide dans le but d'avoir une plus grande homogénéité dans les ARN transcrits.

**1** - Utiliser l'enzyme de restriction XbaI (20 U/ $\mu$ L, référence : R0145S) qui va permettre de cliver le plasmide en aval de la séquence à transcrire. Réaliser la digestion de 100  $\mu$ g de plasmide dans un volume total de 200  $\mu$ L. Vérifier la digestion sur un gel d'agarose 1 %.

**2** - Pour purifier l'ADN, faire une extraction phénol/chloroforme.

**3** - Mélanger la solution et la mettre à -80°C pendant 15 minutes afin de faire précipiter l'ADN.

**4** - Culotter l'ADN en centrifugeant, éliminer le surnageant, centrifuger de nouveau quelques secondes avant de retirer délicatement la phase liquide restante.

**5** - Re-suspendre le culot dans un tampon TE (10 mM Tris-HCl à pH 7,5 et 1 mM EDTA) pour avoir une concentration de 1  $\mu$ g/ $\mu$ L.

### II - Transcription :

**1** - Décongeler les différents échantillons : le mélange d'enzymes est stocké dans du glycérol à -20°C, il n'est donc pas gelé et peut être stocké dans la glace en attendant.



---

Agiter les quatre solutions de ribonucléotides (ATP, GTP, CTP, UTP) et le tampon de réaction 10X jusqu'à décongélation. Garder les échantillons de ribonucléotides dans la glace mais laisser le tampon de réaction à température ambiante.

**2** - Mélanger les échantillons à température ambiante : ajouter dans un tube vide 2  $\mu\text{L}$  de chaque ribonucléotide puis incorporer le tout dans le tube contenant 0,5 à 1  $\mu\text{g}$  de plasmide linéaire. Ajouter 2  $\mu\text{L}$  du mélange d'enzymes puis 6  $\mu\text{L}$  d'eau certifiée sans nucléases et enfin 2  $\mu\text{L}$  du tampon de réaction.

**3** - Agiter doucement la solution puis centrifuger brièvement afin d'avoir la totalité du volume au fond du tube.

**4** - Incuber à 37°C pendant quatre heures.

**5** - Afin d'enlever les traces résiduelles d'ADN, ajouter 1  $\mu\text{L}$  de TURBO DNase puis bien mélanger. Incuber de nouveau à 37°C pendant 15 minutes.

### III - Purification de l'ARN et contrôle :

**1** - Ajouter 115  $\mu\text{L}$  d'eau certifiée sans nucléases puis 15  $\mu\text{L}$  d'une solution fournie par le kit : 5 M d'acétate d'ammonium et 100 mM d'EDTA.

**2** - Extraire avec un volume égal de phénol acide/chloroforme saturé en eau, vortexer une minute et centrifuger à 12000 g pendant deux minutes. Récupérer la phase aqueuse et la transférer dans un nouveau tube, ajouter un volume égal de chloroforme, vortexer une minute et centrifuger une nouvelle fois à 12000 g pendant deux minutes. Récupérer la phase aqueuse et la transférer dans un nouveau tube.

**3** - Précipiter l'ARN en ajoutant 1 volume d'isopropanol puis bien mélanger.

**4** - Stocker la solution pendant 15 minutes à -80°C puis centrifuger à 4°C pendant 15 minutes pour culotter l'ARN.

**5** - Enlever délicatement la phase liquide et re-suspendre le culot dans de l'eau certifiée sans nucléases.

**6** - Afin de connaître la concentration en ARN  $C_{ARN}$  de l'échantillon, il est conseillé de diluer ce dernier d'un facteur  $F$  afin de ne pas saturer le spectrophotomètre et obtenir une valeur précise de l'absorbance à 260 nm. La valeur de la concentration initiale d'ARN en  $\mu\text{g}/\text{mL}$  s'obtient avec la relation suivante :

$$C_{ARN} = 40 \times A_{260} \times F$$

où on considère que pour un ARN simple brin, une unité d'absorbance à 260 nm équivaut à 40  $\mu\text{g}/\text{mL}$ .

**7** - Stocker l'échantillon à -80°C.

Remarques :

- Le séquençage du plasmide ppC2 montre une homologie de 100 % avec l'ARN C2 du CCMV.

- L'étape **II.2** est détaillée pour un volume de réaction total de 20  $\mu$ L. Il est possible de modifier les quantités de chaque réactif selon la quantité d'ARN souhaité. Pour 10 réactions, on obtient au plus 1 mg d'ARN.

- L'étape **II.5** est optionnelle puisque la concentration résiduelle en ADN est négligeable comparée à celle de l'ARN. Pour nos expériences, cette étape est réalisée puisque l'on souhaite un ARN pur.

- L'étape **III.5** peut directement avoir lieu dans un tampon utilisé en expérience.

- Les ARN étant assez fragiles, il est conseillé de faire plusieurs aliquots aux concentrations requises afin de ne pas décongeler la solution stock plusieurs fois.



## Annexe C

# Protocole de chimie de surface

Le protocole détaillé ci-dessous fait suite aux explications données dans la section 3.2.1. Il décrit la dernière version utilisée à ce jour afin de réaliser correctement une chimie de surface à partir du couple PLL-g-PEG et PLL-biotine :

**1** - Passer les lamelles de verre au bain à ultrasons dans des solutions d'acétone pure (15 minutes) et d'eau ultra-pure (deux fois 15 minutes) au plus tard la veille d'une expérience.

**2** - Le matin d'une journée d'expérience, après avoir collé les scotchs sur les lames de microscope, passer les lamelles de verre au *plasma cleaner* à oxygène pendant 10 minutes.

**3** - Pendant ces 10 minutes, préparer une solution de chimie de surface dans du tampon H2 (10 mM HEPES, 150 mM NaCl à pH = 7,4) : 1,5 g/L en PLL-g-PEG et  $10^{-4}$  g/L en PLL-biotine.

**4** - Une fois le cycle du *plasma cleaner* terminé, coller rapidement une lamelle de microscope sur chaque cellule micro-fluidique, connecter une seringue Hamilton et injecter entre 35 et 50  $\mu$ L de chimie de surface. Le but est de ne pas trop attendre entre la fin du *plasma cleaner* et l'injection de la chimie de surface pour profiter du mieux possible de l'activation de la surface de verre.

**5** - Laisser incuber pendant environ 30 minutes puis rincer avec 50  $\mu$ L de tampon H2.

**6** - Pour que la cellule soit prête à être utilisée, injecter 50  $\mu$ L de NeutrAvidine préparée dans le tampon d'assemblage qui va servir pour l'expérience du jour. La concentration en NeutrAvidine est adaptée en fonction de l'expérience réalisée mais globalement j'ai travaillé avec trois valeurs bien distinctes : 400 nM, 80 nM et 20 nM. Le but était de trouver la bonne concentration afin de ne pas avoir trop d'ARN fixés à la surface. Pour les cellules qui serviront plus tard dans la journée, mettre 50  $\mu$ L de tampon H2 dans chaque réservoir sans injecter, cela permet de ne pas assécher le canal en cas de fuite.

**7** - Laisser incuber la NeutrAvidine pendant 10 minutes puis rincer avec 50  $\mu$ L de tampon d'assemblage.

À partir de là, le protocole dépend de l'expérience choisie. Plus de détails sont donnés

---

à la section 3.3.2 pour les contrôles de photophysique et aux sections 5.4 et 5.5 pour les expériences d'assemblage.

Remarques :

- Les aliquots de PLL-g-PEG et PLL-biotine, concentrés à 50 g/L, sont stockés dans le tampon H2 au -80°C pendant au plus un an.
- La chimie de surface est préparée le matin même de chaque expérience et n'est pas conservée au-delà.

Les Figures C.1 et C.2 résument les tests d'optimisation de la chimie de surface de type PLL-g-PEG et PLL-biotine qui sont décrits à la section 3.2.1. Dans ces deux tableaux, les notations sont les suivantes :

- Dans les colonnes "Chimie de surface" ou "Passivation supplémentaire", si une durée est notée entre parenthèse sans indication particulière, elle correspond à la durée d'incubation de l'espèce chimique ajoutée.
- La colonne "Passivation supplémentaire" fait surtout référence à une injection en plusieurs étapes ou non en protéines marquées. Elle peut également indiquer l'ajout de BSA ou lysozyme.
- Les concentrations notées dans les colonnes "Non spécifique" et "Spécifique" indiquent la dernière concentration ajoutée en protéines marquées, Atto565-biotinylés ou ARN.
- Ici, la PLL-g-PEG est notée PLL et la PLL-biotine, PLLbiot.
- "Tp" signifie tampon : H1 ou H2 pour la chimie de surface mais également B' pour les protéines.
- CP\* ou protéines\* font référence aux protéines marquées alors que CP aux protéines sans Atto647N.

	Chimie de surface	Tp	Neutra	Passivation supplémentaire	Non spécifique : [CP*]	Nb points	Spécifique : [Atto565]	Nb points	Figure
Passivation protéines virales	PLL 1g/L	H1	X	150 pM, puis 1,5 nM CP*	<b>15 nM</b>	2500	X		14A
	PLL 1g/L	H1	X	150 pM, puis 1,5 nM CP*	<b>15 nM</b>	2300	X		16A
	PLL 1g/L	H1	X	X	<b>15 nM</b>	tapis	X		16B
	PLL 1g/L	H1	X	12 nM CP	<b>15 nM</b>	1800	X		16C
	PLL 1g/L PLLbiot 1g/L	H1	40 nM	150 pM, puis 1,5 nM CP*	<b>15 nM</b>	3500	X		14C
	PLL 1g/L PLLbiot 1g/L	H1	400 nM	150 pM, puis 1,5 nM CP*	<b>15 nM</b>	2800 + agrégats	X		14B
Passivation Lysozyme	PLL 0,75g/L	H1	X	X	<b>15 nM</b>	4800	X		17A1
	PLL 0,75g/L	H1	X	lysozyme 15nM	<b>15 nM</b>	3500	X		17A2
	PLL 1g/L	H1	X	X	<b>15 nM</b>	3500	X		17B1
	PLL 1g/L	H1	X	lysozyme 56nM	<b>15 nM</b>	4400	X		17B2
Passivation BSA	PLL 1g/L PLLbiot 0,5g/L	H1	400 nM	BSA 10 nM (4h), puis 150 pM CP*	1,5 nM	tapis	X		18A1
	PLL 1g/L PLLbiot 0,5g/L	H1	400 nM	BSA 10 nM (6h), puis 150 pM CP*	1,5 nM	tapis	X		18A2
BSAbiot au lieu de PLLbiot	PLL 1g/L BSAbiot 10 nM (4h, 1h après PLL)	H1	400 nM	150 pM prot marquées	1,5 nM	1000	100 pM	1000	18B
	PLL 1g/L BSAbiot 10 nM (6h, 1h après PLL)	H1	400 nM	150 pM prot marquées	1,5 nM	800	100 pM	700	18C
	PLL 1g/L BSAbiot 100 nM (6h, 1h après PLL)	H1	400 nM	150 pM prot marquées	1,5 nM	286	<b>1 nM</b>	960	19A1/B1
	PLL 1g/L BSAbiot 100 nM (6h, 1h après PLL)	H1	400 nM	150 pM prot marquées	1,5 nM	142	<b>1 nM</b>	263	19A2/B2
	PLL 1g/L BSAbiot 100 nM (6h, 1h après PLL)	H1	400 nM	150 pM prot marquées	1,5 nM	35	<b>1 nM</b>	114	19A3/B3
	PLL 1g/L BSAbiot 100 nM (6h, 1h après PLL)	H1	400 nM	Injection de 5 nM CP* durant une expérience d'assemblage			1200	<b>ARN : 1 nM</b>	100

FIGURE C.1 – Résumé de l'optimisation de la chimie de surface PLL-g-PEG et PLL-biotine, partie 1.

Chimie de surface		Tp	Neutra	Passivation supplémentaire	Non spécifique : [CP*]	Nb points	Spécifique : [Atto565]	Nb points	Figure
Passivation Lysine	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L	H2	400 nM	X	5 nM	4000	X		20A1
	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L Lysine 0,1 g/L (30 min, 10 min après PLL)	H2	400 nM	X	5 nM	4200	X		20A2
	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L Lysine 1 g/L (30 min, 10 min après PLL)	H2	400 nM	X	5 nM	4300	X		20A3
	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L Lysine 10 g/L (30 min, 10 min après PLL)	H2	400 nM	X	5 nM	4000	X		20A4
Injection Lysine avec protéines*	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L	H2	400 nM	X	5 nM	4500	X		20B1
	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L	H2	400 nM	X	5 nM + Lysine 0,05g/L dans Tp B'	3500	X		20B2
	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L	H2	400 nM	X	5 nM + Lysine 0,5g/L dans Tp B'	3500	X		20B3
	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L	H2	400 nM	X	5 nM + Lysine 5g/L dans Tp B'	2300	X		20B4
Optimisation finale	PLL 1g/L PLLbiot 1g/L	H1	400 nM	150 pM prot marquées	1,5 nM	tapis	<b>1 nM</b>	500	21A1/B1
	PLL 1g/L PLLbiot 10 <sup>-2</sup> g/L	H1	400 nM	150 pM prot marquées	1,5 nM	1200	<b>1 nM</b>	tapis	21A2/B2
	PLL 1g/L PLLbiot 10 <sup>-5</sup> g/L	H1	400 nM	150 pM prot marquées	1,5 nM	500	<b>1 nM</b>	3000	21A3/B3
	PLL 1g/L PLLbiot 10 <sup>-5</sup> g/L	H1	400 nM	Injection de 5 nM CP* durant une expérience d'assemblage		700	<b>ARN : 1 nM</b>	100	22
	PLL 1g/L PLLbiot 10 <sup>-4</sup> g/L	H1	400 nM	Injection de 5 nM CP* durant une expérience d'assemblage		700	ARN : 100 pM	1400	23

FIGURE C.2 – Résumé de l'optimisation de la chimie de surface PLL-g-PEG et PLL-biotine, partie 2.

## Annexe D

# Double hybridation de l'ARN simple brin

Dans cette annexe, je présente le protocole d'hybridation de l'ARN simple brin avant de revenir sur les accroches non spécifiques en oligo-Atto565 sur la lamelle de verre.

### Protocole d'hybridation :

Le protocole détaillé ci-dessous fait suite aux explications données dans la section 3.2.3. Il décrit comment préparer un volume d'hybridation de 11  $\mu\text{L}$  contenant 1  $\mu\text{M}$  en ARN simple brin :

**1** - Sortir du congélateur à  $-20^{\circ}\text{C}$  les quatre solutions suivantes : un échantillon d'oligo-Atto565 (noté OA) à 10  $\mu\text{M}$ , un échantillon d'oligo-biotinylé (noté OB) à 10  $\mu\text{M}$ , du tampon d'hybridation (100 mM Tris-HCl, 2 M NaCl et 10 mM EDTA à  $\text{pH} = 7,5$ ) concentré dix fois plus que nécessaire et de l'eau ultra-pure certifiée sans ribonucléases.

**2** - Sortir du congélateur à  $-80^{\circ}\text{C}$  un aliquot d'ARN C2 contenant quelques microlitres.

**3** - Ajouter à l'aliquot d'ARN : 3,3  $\mu\text{L}$  d'OA, 1,1  $\mu\text{L}$  d'OB, 1,1  $\mu\text{L}$  de tampon d'hybridation et compléter avec le volume approprié d'eau ultra-pure pour atteindre 11  $\mu\text{L}$ . La solution contient 1  $\mu\text{M}$  d'ARN, 3  $\mu\text{M}$  d'OA, 1  $\mu\text{M}$  d'OB et du tampon d'hybridation dilué dix fois comme attendu.

**4** - Placer l'échantillon dans un bain à sec à  $68^{\circ}\text{C}$  pendant 5 minutes.

**5** - Placer l'échantillon dans un nouveau bain à sec mais à  $55^{\circ}\text{C}$ , toujours pendant 5 minutes.

**6** - Laisser reposer la solution d'hybridation pendant au moins 5 minutes à  $4^{\circ}\text{C}$  avant de commencer l'expérience.

À partir de là, l'ARN peut être dilué à la concentration souhaitée selon les besoins de l'expérience. La dilution a lieu directement dans le tampon d'assemblage du jour.



---

### Remarques :

- Les aliquots d'ARN sont obtenus après avoir dilué dans de l'eau ultra-pure certifiée sans ribonucléases la solution stock synthétisée selon le protocole décrit à l'annexe B. La concentration des aliquots dilués est mesurée au Nanodrop et varie généralement entre 3 et 4  $\mu\text{g}/\mu\text{L}$ . Ensuite, ils sont stockés au  $-80^\circ\text{C}$  directement dans le volume adapté pour obtenir 1  $\mu\text{M}$  dans 11  $\mu\text{L}$ .

- Au bout de 20 minutes à  $4^\circ$ , j'ajoute 0,5  $\mu\text{L}$  de SUPERase-In (*Thermo Fisher Scientific*, 20 U/ $\mu\text{L}$ ) qui est un inhibiteur de ribonucléases, dans la solution d'hybridation. De plus, j'attends généralement 30 minutes après le deuxième bain à sec avant de diluer l'ARN

- Pour les dilutions, je prélève 1  $\mu\text{L}$  du volume d'hybridation afin d'atteindre la concentration souhaitée. À la fin d'une cellule micro-fluidique, les dilutions sont jetées et je repars de la solution concentrée à 1  $\mu\text{M}$  en ARN pour la prochaine. La concentration nécessaire en ARN pour avoir une bonne couverture dépend énormément de la réussite de la chimie de surface et de la concentration en NeutrAvidine mais généralement, je peux monter à quelques centaines de picomolaires. Par ailleurs, il arrive souvent que la concentration en ARN à injecter augmente en fin d'expérience, ce qui pourrait s'expliquer par une hybridation qui se dégrade au fur et à mesure.

- Durant la préparation de l'hybridation, j'utilise un excès d'OA (3  $\mu\text{M}$ ) par rapport à l'ARN (1  $\mu\text{M}$ ). Le but ici est de faire en sorte que tous les ARN hybridés avec un oligo-biotinylé possède également un Oligo-Atto565. De cette manière, tous les ARN qui se sont adsorbés à la chimie de surface seront visibles en illumination TIRF.

- Que ce soit pour une simple ou double hybridation, les oligo-nucléotides sont des séquences d'ADN qui ont été synthétisées par l'entreprise *Eurogentec*.

## Retour sur les accroches non spécifiques des ARN hybridés

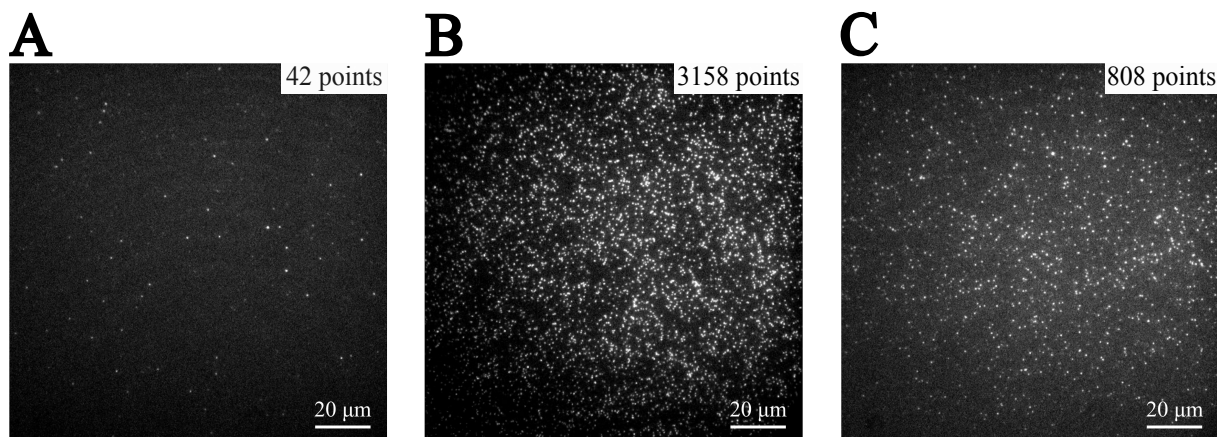
À la section 3.2.3, j'explique que la double hybridation de l'ARN nous assure que les particules détectées avec le laser d'excitation vert indiquent bien la présence du matériel génétique sur la lamelle. Cette affirmation n'est valable que si l'oligo-Atto565 utilisé ne se fixe pas non spécifiquement sur la surface. Au cours des différents tests réalisés sur ce point, j'ai pu généralement obtenir une absence presque totale d'adsorptions non spécifiques provenant des oligo-Atto565. Cependant, depuis les changements conjoints des lots de NeutrAvidine et de chimie de surface au début de l'année 2024, il m'est arrivé d'en voir apparaître plus souvent et de manière très variée en fin de thèse. En effet, sur deux journées d'expérience consécutives et sept cellules utilisées, les résultats sont les suivants :

- **Pour quatre cellules**, le nombre d'accroches est inférieur à 100 points, ce qui peut être assimilé à des lamelles sales.

- **Sur deux autres cellules**, on retrouve respectivement près de 250 et 450 points, ce qui commence à être assez important.

- **Enfin sur la dernière cellule**, près de 1500 oligo-Atto565 se sont fixés non spécifiquement à la surface.

Ces trois exemples ont tous été obtenus à partir d'une injection de 300 pM en oligo dans le tampon B' (50 mM MES, 100 mM NaCl à pH = 6,1) sur une lamelle passivée avec une chimie de surface de type PLL-g-PEG, PLL-biotine et NeutrAvidine. Dans la majorité des cas, les accroches en oligo-Atto565 sont suffisamment peu nombreuses pour ne pas être gênantes, mais il est également possible d'obtenir une couverture conséquente et surtout, semblable à celle provenant d'ARN hybridés. Pour essayer de limiter leur présence, trois autres cellules ont été étudiées dans les mêmes conditions expérimentales que les sept précédentes. Les oligos ont été incubés pendant deux minutes dans chaque cellule microfluidique avant d'être rincés avec 50  $\mu$ L de tampon B'. La Figure D.1 illustre les accroches observées dans ces trois cas.



**FIGURE D.1** – Interactions non spécifiques issues des oligo-Atto565. A-B-C) Images acquises, la même journée, sur trois cellules différentes après avoir injecté 300 pM en OA sur une chimie de surface de type PLL-g-PEG, PLL-biotine et NeutrAvidine.

Sans que les conditions expérimentales n'aient été modifiées d'une cellule à une autre, nous pouvons voir que la couverture en oligos varie énormément et qu'elle peut être beaucoup trop importante comme le montre la Figure D.1.B. Ce type de résultat n'est évidemment pas souhaité mais il semblerait donc bien qu'il soit possible. Pour les deux cas les plus problématiques, chaque cellule a été rincée une deuxième fois, mais avec 100  $\mu$ L en tampon B'. À la suite de ce rinçage, une nouvelle acquisition a été réalisée et ce processus a été répété une seconde fois. Les nombres d'accroches non spécifiques en oligo-Atto565 pour les cellules illustrées par les Figures D.1.B et C sont donnés dans le Tableau D.1. **Nous pouvons voir que les adsorptions non souhaitées ont été respectivement divisées par un facteur 4 et 5 après les deux rinçages à 100  $\mu$ L. Le nombre d'accroches final est toujours assez élevé, surtout pour la cellule de la Figure**

D.1.B, mais l'amélioration est significative.

Nombre d'oligo-Atto565			
Cellule	Post-injection + rinçage 50 $\mu$ L	Post-rinçage 1 à 100 $\mu$ L	Post-rinçage 2 à 100 $\mu$ L
1	42	/	/
2	3158	1239	739
3	808	226	153

Nombre d'ARN			
Cellule	Incubation 2 min + rinçage 50 $\mu$ L	t = 15 min + protéines 30 $\mu$ L	t = 25 min
1	431	283	289
2	2215	1137	897
3	780	466	434
4	2784	1719	1418
5	704	343	315
6	1133	553	468

**TABLEAU D.1** – Effet du rinçage sur les accroches en oligo-Atto565 et ARN hybridés. Les oligo-Atto565 et les ARN ont été incubés pendant 2 minutes avant d'être rincés avec 50  $\mu$ L de tampon B'. Pour les tests avec les OA, les trois cellules ont été réalisées le même jour et il n'y a aucune attente particulière entre les différents films ou rinçages. En ce qui concerne les ARN, les résultats sont tirés de deux expériences de dynamique d'assemblage effectuées dans les mêmes conditions. Les notations "t = 15 min" et "t = 25 min" signifient que les acquisitions 2 et 3 ont eu lieu respectivement 15 et 25 minutes après la première.

Cette solution pourrait donc être envisagée afin de limiter le plus possible les adsorptions non spécifiques en oligo-Atto565 dans le cas où ces dernières sont présentes. Néanmoins, l'ARN semble également sensible à une étape de rinçage supplémentaire comme le montre la seconde partie du Tableau D.1. Ici, je détaille l'évolution du nombre d'ARN hybridés présents sur les lamelles de six cellules utilisées en expérience de dynamique d'assemblage. Dans ce cas, les 100 pM en ARN ont été injectés dans le tampon B' puis localisés à trois reprises : à la suite de la combinaison incubation / rinçage puis 15 et 25 minutes après l'injection de protéines. Les deux derniers films sont enregistrés en amont et en aval de l'acquisition permettant d'observer les sous-unités marquées dans le cas d'une expérience de dynamique d'assemblage comme je l'explique à la section 5.4. **En regardant la statistique en ARN, nous pouvons remarquer que cette dernière a diminué de 40 à 50 % entre les deux premières acquisitions alors qu'elle est quasiment stable entre le deuxième et le troisième film.** Si le problème était lié à un défaut dans l'hybridation, le nombre d'ARN devrait continuer à décroître tout au long de la durée d'utilisation d'une cellule, ce qui n'est pas le cas ici. En revanche, la diminution du nombre de points détectés dans la voie verte intervient entre les deux premiers films d'ARN, soit après l'injection de 30  $\mu$ L en protéines. Il est donc possible qu'une partie

de ces pertes puisse être attribuée au fait d'ajouter les sous-unités au sein de la cellule micro-fluidique. Une autre hypothèse serait que les objets qui ne sont plus détectés à la deuxième acquisition correspondent aux oligos accrochés non spécifiquement à la surface. Malheureusement, dans ces conditions de marquage, nous ne pouvons pas le vérifier expérimentalement.

Ainsi, malgré les précautions visant à limiter le non spécifique dans la voie de détection verte, il semblerait que les oligo-Atto565 puissent se fixer à la surface. Par ailleurs, le côté aléatoire de ces accroches rend difficile la compréhension et l'étude plus approfondie de ce phénomène. Cependant, afin d'en apprendre davantage sur le potentiel biais introduit par l'adsorption non souhaitée des oligo-Atto565, je me suis intéressé à leurs interactions avec les sous-unités marquées. Pour ce faire, j'ai réutilisé l'oligo qui nous permet d'effectuer une hybridation simple de l'ARN. Ce dernier possède à la fois une biotine et un Atto565 : il peut donc se fixer spécifiquement à la surface et être aussi observé dans la voie de détection verte. La seule différence entre ces deux oligos concerne leur taille (53 nucléotides pour celui participant à la double hybridation contre 40 pour l'autre) mais nous faisons l'hypothèse que les interactions avec les dimères devraient être similaires. Voici le protocole expérimental utilisé :

- 1 - Injection des oligos puis incubation de 2 minutes avant rinçage.
- 2 - Acquisition du film vert.
- 3 - Injection de 50 nM en protéines avec un taux de marquage de 10 %.
- 4 - Incubation pendant 15 minutes puis second film vert.
- 5 - Acquisition rouge dans la foulée : 1600 images, 100 ms de temps d'exposition et 375 ms de temps de cycle.
- 6 - Troisième film vert.

En ce qui concerne l'analyse, les sites d'intérêt excités par le laser vert sont sélectionnés en colocalisant ceux présents juste avant et juste après l'acquisition dans la voie rouge. Dans un second temps, nous regardons si des sous-unités marquées viennent se fixer aux oligos. Les signaux collectés sont fournis à Ruptures et une interaction non spécifique est comptabilisée si à au moins une reprise, l'intensité du signal en marches d'escalier est supérieure à la somme entre la valeur moyenne du fond et l'intensité moyenne d'émission d'un fluorophore  $h$ . Cette valeur nous permet d'accéder au pourcentage de colocalisations en la divisant par le nombre total de sites détectés dans la voie verte.

Une fois l'analyse effectuée, il vient que sur les trois cellules utilisées, près de 1700 sites d'oligo ont été étudiés et environ 13 % d'entre eux présentent au moins une interaction avec une sous-unité marquée. Ce résultat est très proche de celui obtenu à la section 5.3.1 entre les Atto565-biotinylés et ces mêmes dimères marqués. Pour aller plus loin, j'ai compté le nombre d'images par oligo où les interactions avec des sous-unités se produisent. Au total, on en compte 8533 sur les  $2,6 \times 10^6$  que représentent les 1653 oligos étudiés (1600 images

---

par site considéré). Cela correspond donc à seulement 0,3 % du total d'images enregistrées. Pour avoir un point de comparaison, à la section 5.3.1, je montre que lorsque l'on injecte 50 nM en protéines, dont 10 % sont marquées, sur une lamelle avec des ARN, près de 32 % du nombre total d'images acquises montrent des interactions spécifiques entre sous-unités marquées et le matériel génétique (au total on compte  $1,3 \times 10^7$  images pour 8175 sites d'ARN avec 1600 images chacun).

**Pour conclure, même si des oligo-Atto565 sont fixés à la surface, l'analyse faite ici nous montre que les protéines ont peu tendance à s'y accrocher. Ainsi, au cours d'une expérience d'assemblage, le pourcentage d'échantillons marqués dans le vert avec au moins une accroche en sous-unité marquée sera simplement plus faible si les oligo-Atto565 sont présents en abondance sur la surface.**

## Annexe E

# Protocole d'analyse des données de photophysique

Le but de cette annexe est d'expliciter les différents seuils choisis afin d'analyser les expériences de photophysique.

L'analyse commence, comme toutes les expériences de fluorescence, par une étape de traitement d'images que j'ai présentée à la section 3.4. Généralement, la taille du carré permettant de faire le floutage de l'image brute mesure 13x13 pixels et le diamètre de la *Rolling Ball* est de 5 pixels. En ce qui concerne les paramètres de la fonction *trackpy.locate*, le diamètre  $d$  considéré vaut 3 ou 5 pixels et la valeur de *minmass* dépend des conditions d'illumination mais est généralement comprise entre 0,13 et 0,17 u.a. Les différents paramètres que je viens de donner sont également valables pour les expériences d'assemblage. Par ailleurs, à la suite de chaque traitement d'images, une zone de l'image ne contenant aucun marqueur fluorescent est isolée afin de mesurer la nouvelle valeur moyenne du fond et l'écart-type associée. Pour ce faire, la fonction *Plot Z-axis Profile* du logiciel *ImageJ* est utilisée. Elle renvoie la valeur moyenne du fond et l'écart-type de la zone isolée pour les  $N$  images contenues dans l'acquisition. Ce tableau de valeurs est récupéré afin de moyenniser cette grandeur sur les  $N$  images du film.

**Comme je l'ai expliqué à la section 3.3.2, le but des expériences de photophysique est de caractériser la photostabilité du fluorophore d'intérêt à savoir l'Atto647N. Il est donc important de travailler avec des marqueurs fluorescents et non des agrégats afin d'avoir les informations les plus précises possibles.** Pour ce faire, on travaille avec l'ajustement en marches d'escalier estimé par Ruptures et trois conditions doivent être respectées par les intensités initiales  $I_0$  de chaque fluorophore détecté :

- $I_0$  doit être supérieure à la somme entre la valeur moyenne du fond ( $I_{fond}$ ) et son écart-type ( $\sigma_{fond}$ ), ce qui correspond généralement à une intensité de 0,06 - 0,07 u.a.
- $I_0$  doit être inférieure à la somme entre la valeur moyenne de toutes les  $I_0$  détectées

---

sur le champ de vue et son écart-type.

- L'écart entre  $I_0$  et  $I_{max}$  (intensité maximale) de chaque signal doit être inférieur à  $\sigma_{fond}$ .

Tous les fluorophores respectant ces conditions sont considérés comme étant uniques. Ensuite, les seuils choisis dépendent du paramètre que l'on veut étudier : le photoblanchiment, le pourcentage de marqueurs immortels et les clignotements. L'analyse de ces trois paramètres utilise également l'ajustement en marches d'escalier estimé par Ruptures.

**Le photoblanchiment** : ici on souhaite collecter deux informations particulières : l'intensité émise par le fluorophore lors de sa désexcitation, à savoir la hauteur de la marche descendante, et l'instant auquel le photoblanchiment se produit. Sachant que nous travaillons avec des marqueurs uniques, nous nous attendons à repérer une seule marche descendante bien nette. Cependant, comme nous l'avons vu au chapitre 4, Ruptures peut ajouter quelques marches qui n'ont pas lieu d'être. De plus, nous ne sommes pas à l'abri que le filtre de sélection des marqueurs uniques ait laissé passer des fluorophores doubles. Ainsi trois nouvelles conditions doivent être respectées pour comptabiliser un signal dans la case "photoblanchiment" :

- Si Ruptures repère N marches, il faut que N-1 marches soient plus petites que  $\sigma_{fond}$ .
- En valeur absolue, la hauteur d'une seule marche descendante doit être supérieure à la somme entre  $I_{fond}$  et  $\sigma_{fond}$ .
- À la dernière image, l'intensité doit être inférieure à la somme entre  $I_{fond}$  et  $\sigma_{fond}$ .

**Fluorophores qui ne s'éteignent pas** : dans un cas idéal, le signal ne doit pas évoluer au cours du temps et Ruptures ne devrait donc repérer aucune marche. Comme ce n'est jamais vraiment le cas, trois conditions doivent être respectées pour considérer qu'un fluorophore est immortel :

- $I_{min}$  (intensité minimale) doit être supérieure à la somme entre  $I_{fond}$  et  $\sigma_{fond}$ .
- Toutes les marches potentiellement repérées par Ruptures doivent être inférieures à  $\sigma_{fond}$ .
- L'écart entre  $I_{min}$  et  $I_{max}$  doit être inférieur à la somme entre  $I_{fond}$  et  $\sigma_{fond}$ .

**Pour ces deux paramètres, Ruptures a été entraîné afin de repérer au plus une marche. Ainsi, une pénalité d'environ 25,1 a été choisie, ce qui favorise la détection d'un faible nombre de marches. Cela nous permet de travailler également avec les fluorophores qui présentent des clignotements courts. En effet, avec cette pénalité, généralement seule la dernière marche caractéristique du photoblanchiment sera repérée.**

**Les clignotements** : ce mécanisme photophysique est plus compliqué à analyser puisque, comme je l'ai présenté à la section 3.3.2.2, la durée des passages en OFF est très hétéro-

gène. Par ailleurs, j'ai déjà pu observer des signaux de fluorophore unique qui ne clignotaient pas totalement. En effet, pendant plusieurs images consécutives, l'intensité émise par le marqueur est réduite de moitié ou de deux tiers de son intensité totale alors que tous les autres passages en OFF et le photoblanchiment montrent une absence totale d'émission. Ce genre d'événements combinés à certains fluorophores doubles qui ont pu passer le premier filtre rendent l'analyse du clignotement assez complexe. **Afin de repérer au mieux les différents passages en OFF, un nouvel entraînement de Ruptures a été réalisé sur des données bien choisies et la pénalité d'analyse est d'environ 2,95.** Enfin, deux conditions doivent être également respectées :

- Il faut que la hauteur du clignotement soit supérieure à la somme entre  $I_{fond}$  et  $\sigma_{fond}$ .
- L'écart entre les valeurs moyennes des plateaux estimés par Ruptures et situés juste autour du passage en OFF doit être inférieur à  $\sigma_{fond}$ .





## Annexe F

# Évaluation des performances de Ruptures et quickPBSA

Dans cette annexe, je reviens sur quelques points concernant le chapitre 4 et la détection de marches sur des données obtenues en microscopie de fluorescence. Je décris notamment la fonction permettant d'ajouter des clignotements aux données simulées, l'influence de ces derniers sur la phase d'entraînement de Ruptures ainsi que le choix du seuil pour quickPBSA.

### Simulation : ajout des clignotements

Dans cette partie je ne rentre pas en détails dans la présentation de la fonction utilisée mais j'explique brièvement son principe.

L'objectif de cette fonction est de rajouter, sur certains fluorophores d'un même signal, des passages en OFF temporaires avant qu'ils ne quittent potentiellement le site d'ARN simulé. En amont de chaque signal,  $N$  entiers valant 0 ou 1 sont tirés au sort avec la fonction *random.choices* à partir d'une répartition des poids qui n'est pas uniforme. En effet, le 0 possède 20 % de chances d'être choisi contre 80 % pour le 1. **Cela signifie qu'en associant le 0 à un fluorophore présentant des clignotements, nous choisissons de générer des données pour lesquelles 20 % des fluorophores clignent.** Ici, l'entier  $N$  correspond au nombre de marqueurs fluorescents présents dans un même signal. Par ailleurs, pour qu'un clignotement ait bien lieu, il faut que la durée de vie du fluorophore étudié soit supérieure à un seuil choisi en amont. Ce dernier a été fixé à 8 images vis-à-vis de Ruptures puisqu'en l'état, l'algorithme ne peut détecter que des plateaux de 2 images minimum. Il faut donc qu'un passage en OFF dure au moins 2 images et que les plateaux situés en amont et en aval d'un clignotement soient eux aussi longs d'au moins 2 images, ce qui donne 6 images en tout. Une marge supplémentaire de 2 images a été ajoutée, ce qui fait bien 8 images au minimum. **Sachant que l'intervention de ce seuil a lieu après**

le tirage au sort des 0 et 1, ce ne sont pas 20 % des marqueurs qui clignent mais plutôt 19 %.

Une fois que les fluorophores pouvant clignoter ont été identifiés le nombre de passages en OFF  $n_{cl}$  ainsi que la durée  $t_{cl}$  de ces derniers sont tirés au sort selon deux lois exponentielles de constantes respectives 0,45 et 0,04. Le nombre moyen de passages en OFF vaut donc environ 2,2 et leur durée moyenne plutôt 25 images soit un peu plus de 9 secondes. Ces deux paramètres peuvent être tirés au sort plusieurs fois pour un même fluorophore puisque que le but est de respecter la contrainte mentionnée plus haut : avoir au moins deux images d'écart entre deux marches consécutives pour offrir la possibilité à Ruptures de repérer ces événements. Après avoir suivi ce protocole pour toute la série de données étudiée, l'intensité bruitée de chaque signal est corrigée en retirant l'intensité émise par chaque fluorophore qui doit clignoter sur une durée  $t_{cl}$ . Les positions des clignotements sont choisies presque aléatoirement, la seule contrainte étant encore une fois de laisser au moins deux images d'écart entre deux marches consécutives. **Ainsi, dans cette première version de simulation des clignotements, deux fluorophores ne peuvent pas clignoter en même temps, ce qui est tout à fait possible en réalité.** La prise en compte de ce mécanisme photophysique n'est donc pas encore parfaite puisque des améliorations peuvent être ajoutées. Cependant, comme je l'ai montré à la section 4.3.2, il est quand même possible de comprendre comment les clignotements affectent notre estimation des paramètres d'assemblage à partir de données expérimentales.

À la fin de l'ajout des clignotements, toutes les positions des marches d'une même série de données sont stockées dans un tableau afin d'avoir accès à la *Ground-truth* de chaque signal.

## Ruptures : influence des clignotements sur l'entraînement

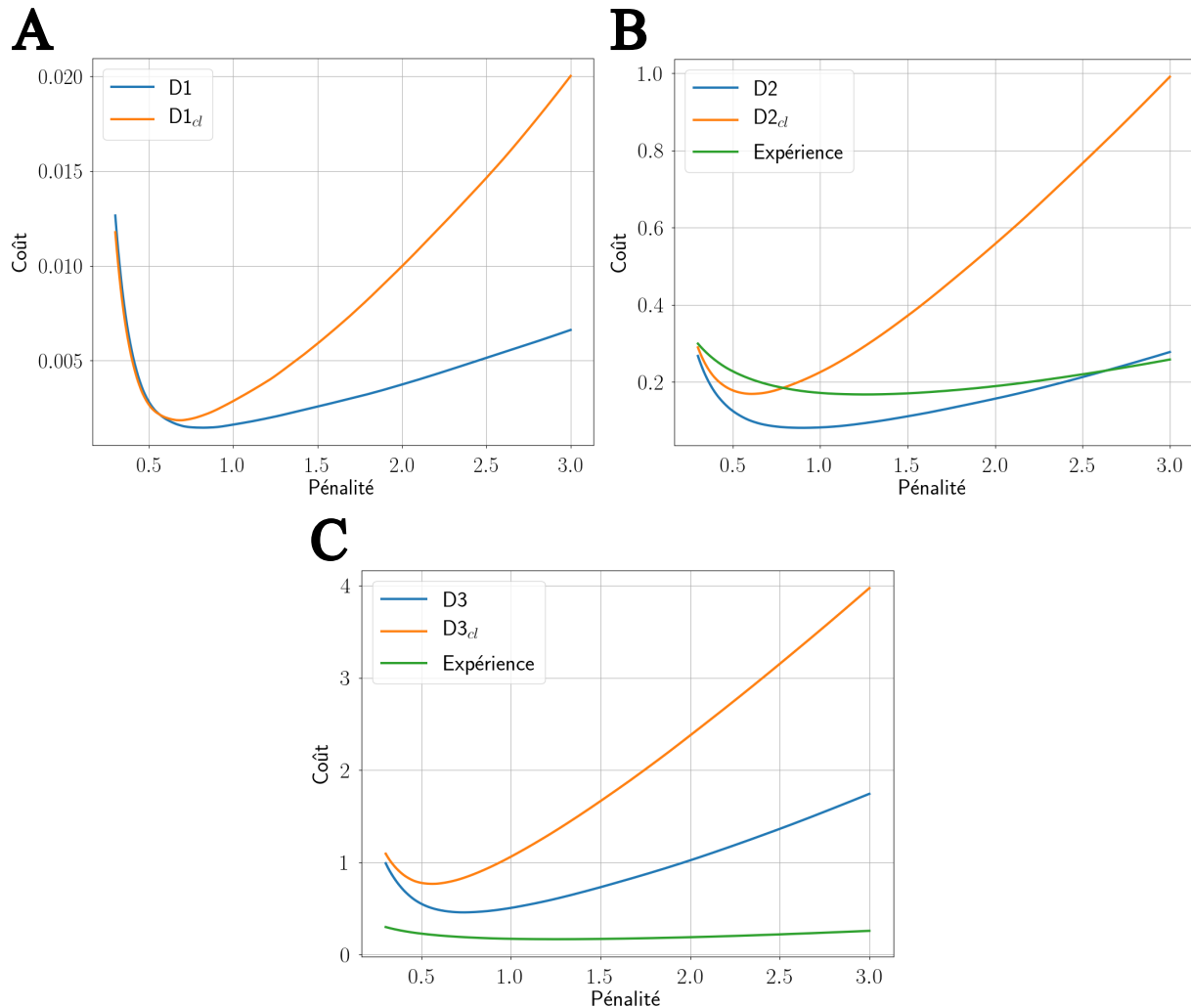
À la section 4.3.2, je précise que les pénalités ont évolué du fait que des clignotements ont été ajoutés aux signaux simulés. Dans le Tableau F.1, je regroupe les valeurs de pénalités pour les trois séries de données d'intérêt.

	D1	D2	D3
Pénalité Ruptures	0,819	0,895	0,735
Pénalité Ruptures <sub>cl</sub>	0,684	0,605	0,559
Pénalité $F_{score}$	1,22	1,28	1,70
Pénalité $F_{score,cl}$	0,930	1,06	1,25

**TABLEAU F.1 – Données d'entraînement : comparaison des pénalités avec et sans clignotements.** La "Pénalité Ruptures" est celle qui minimise le coût alors que la "Pénalité  $F_{score}$ " maximise le  $F_{score}$  pour chaque jeu de données. Les résultats en rouge correspondent aux données avec clignotements et sont indexés par "cl".

Comme attendu, pour chaque jeu de données d'entraînement, la pénalité donnée par

Ruptures et celle qui maximise le  $F_{score}$  ont diminué. Cette différence a pour but de faciliter la détection d'un nombre de marches plus grand. Enfin, à la Figure F.1, les fonctions de coût pour chaque série sont représentées avec et sans clignotements.



**FIGURE F.1 – Fonctions de coût pour des données d'entraînement avec et sans clignotements.** Séries A) D1, B) D2 et C) D3. Les fonctions de coût ont été tracées en bleu pour les données sans clignotements et en orange avec. En vert, il s'agit de la fonction de coût obtenue après avoir entraîné Ruptures sur 50 signaux choisis aléatoirement parmi les 819 collectés au cours d'une expérience d'assemblage réalisée à l'équilibre. Les paramètres d'acquisition sont les suivants : 1600 images, 100 ms de temps d'exposition, 375 ms de temps de cycle et puissance après le cube PBS de 3,7 mW.

De manière générale, nous pouvons remarquer qu'en présence de clignotements les coûts calculés pour chaque pénalité durant la phase d'entraînement sont plus élevés. Cela signifie qu'il est plus difficile de se rapprocher de la *Ground-truth*. Par ailleurs, même si ce n'est pas parfait, nous voyons bien sur la Figure F.1.B que la série D2 est à l'origine de la fonction de coût la plus proche de celle obtenue en expérience (pénalité selon Ruptures de 1,26).

---

## quickPBSA : détermination du seuil pour quickPBSA

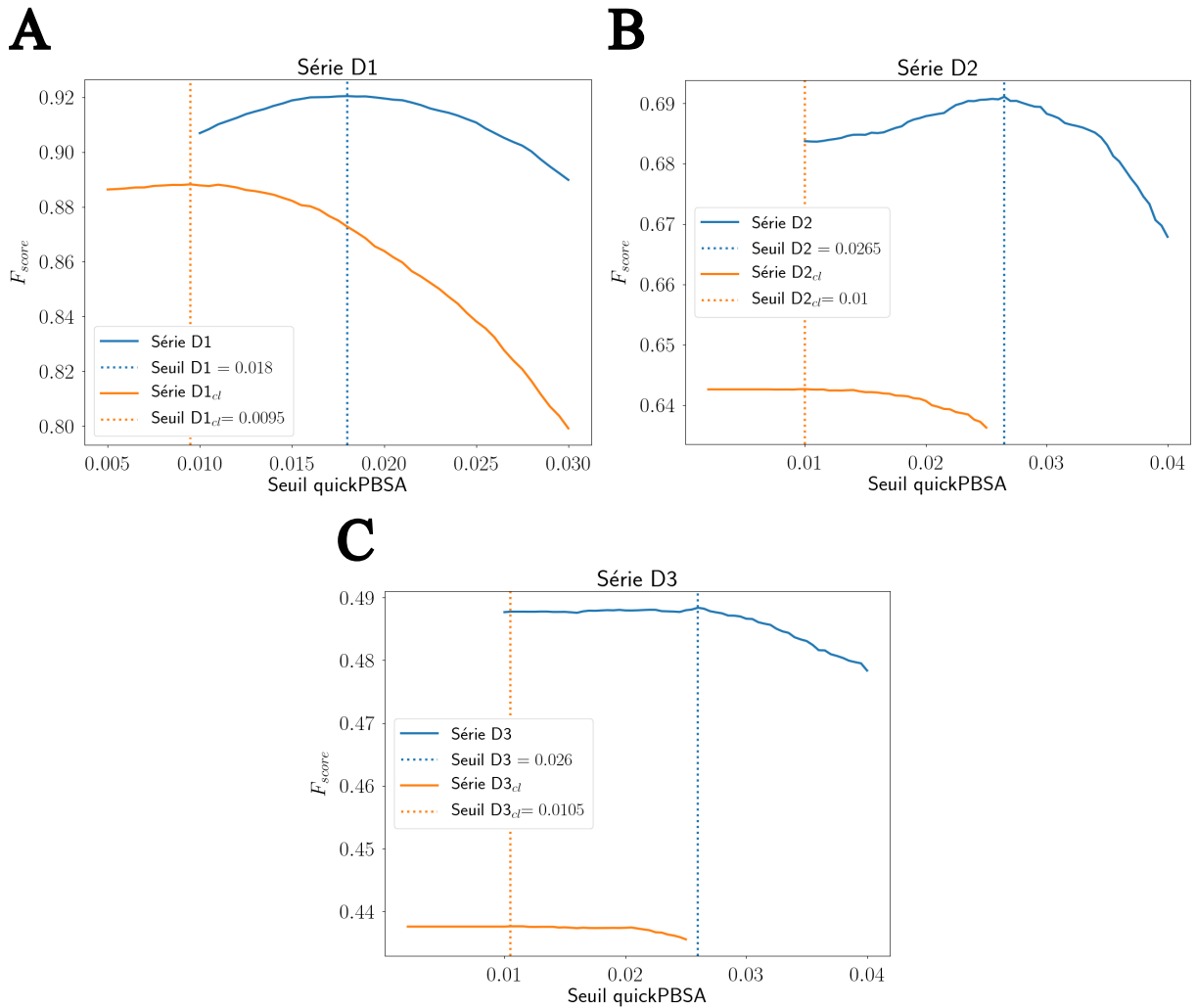
Au chapitre 4, les analyses faites avec quickPBSA ont été réalisées avec des seuils différents, dont je donne les valeurs dans le Tableau F.2.

	Seuil SANS clignotements	Seuil AVEC clignotements
D1	0,0180	0,0095
D2	0,0265	0,0100
D3	0,0260	0,0105

**TABLEAU F.2** – **Seuils utilisés avec quickPBSA.** Les seuils choisis sont ceux qui maximisent le  $F_{score}$  pour chaque série de données.

Ces seuils ont été choisis parce qu'ils maximisent le  $F_{score}$ . En effet, pour chaque série de données, quickPBSA a été appliqué aux 1600 signaux avec un seuil allant de 0,005 à 0,04 par pas de 0,0005. Ensuite, les valeurs de TP, FP et FN ont été sommées sur les 1500 données tests dans le but de calculer le *Recall*, la *Precision* puis d'en déduire le  $F_{score}$ . Le désavantage de quickPBSA ici est qu'il n'y a pas de phase d'entraînement et que chaque seuil a été utilisé sur l'ensemble des données tests. Cette méthode étant très lente, je n'ai pas pu affiner le choix du seuil.

Sur la Figure F.2, je trace le  $F_{score}$  en fonction des seuils testés pour chaque série de données, avec et sans clignotements. Nous pouvons voir que les valeurs de  $F_{score}$  évoluent très peu autour des maximums. Cela signifie que même si la recherche du meilleur seuil est améliorée, les performances de quickPBSA ne seront que très peu modifiées.



**FIGURE F.2** – Évolution du  $F_{score}$  en fonction du seuil quickPBSA. Séries A) D1, B) D2 et C) D3. Les valeurs de  $F_{score}$  ont été calculées à partir des 1500 signaux tests de chaque série. Les courbes bleues concernent les données sans clignotements et celles en orange avec clignotements.



## Annexe G

# Photométrie de masse

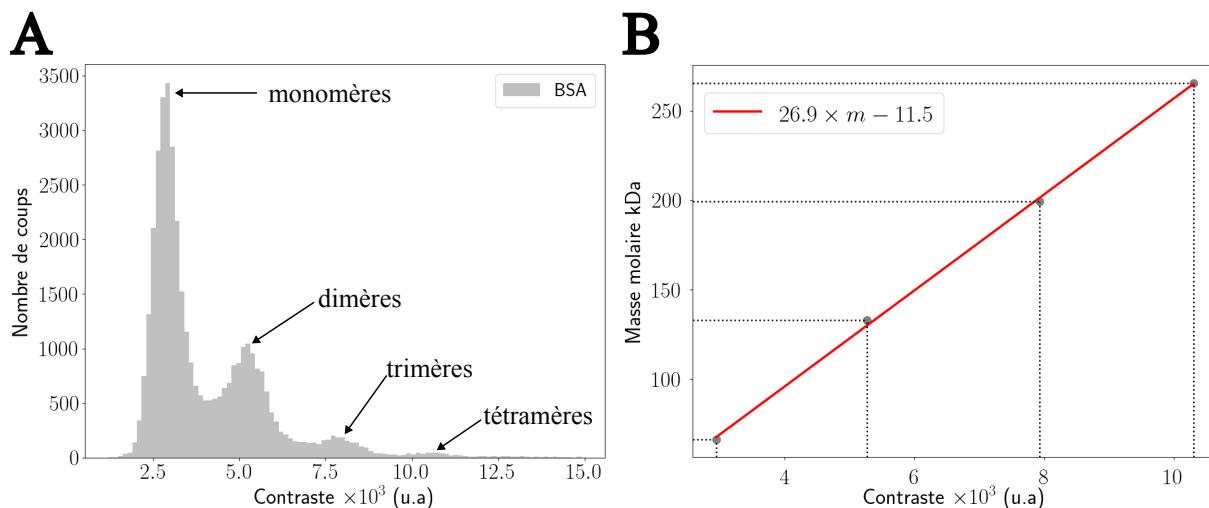
Cette annexe vient compléter la section 5.1 sur la photométrie de masse. Tous les résultats présentés ci-dessous ont été obtenus à partir d'expériences menées sur le modèle *TwoMP* de la compagnie *Refeyn*.

### Calibration de l'appareil :

À la section, 5.1.1.1, j'insiste sur le fait que l'équivalence entre le contraste des interférences et la masse molaire des objets d'intérêt n'est possible que si l'appareil utilisé est calibré à partir d'objets de masse connue. Sur la Figure G.1, je montre le résultat de cette étape faite à partir de la BSA diluée dans le tampon B' (50 mM MES, 100 mM NaCl à pH = 6,1). Nous pouvons voir sur la Figure G.1.A la distribution des contrastes correspondant aux différentes particules de BSA détectées. Il est possible de repérer quatre maximums locaux qui décrivent les quatre premiers oligomères de la BSA : le monomère de 66 kDa, le dimère de 133 kDa, le trimère de 199 kDa et le tétramère de 266 kDa. En utilisant les valeurs de contraste de chacun de ces maximums ainsi que les masses molaires des oligomères correspondants, une relation linéaire est obtenue comme le montre la Figure G.1.B. **Les paramètres du modèle linéaire ont ensuite été utilisés par le logiciel d'analyse de l'appareil afin d'avoir accès aux histogrammes en masse molaire pour les échantillons que nous avons souhaités caractériser par la suite.**

Je précise ici qu'un cinquième point a été ajouté à la calibration : il s'agit de l'uréase dont la masse molaire est de 544 kDa. L'intérêt d'utiliser une protéine plus volumineuse est de faire en sorte que la calibration soit plus précise si les objets à caractériser ont une masse molaire importante, comme c'est le cas pour des capsides pleines de CCMV avec 4,6 MDa. Enfin, pour calibrer l'appareil et obtenir les meilleurs résultats possibles, Bertrand Raynal m'a conseillé d'utiliser le tampon dans lequel les particules d'intérêt sont étudiées au cours de l'expérience. Si ce n'est pas le cas, il est possible que les masses molaires déduites par l'appareil soient moins précises. Une fois la calibration effectuée, l'expérience peut démarrer.





**FIGURE G.1 – Photométrie de masse : calibration expérimentale de l'appareil *Refeyn TwoMP*.** A) Distribution des contrastes correspondant aux quatre premiers oligomères de BSA (monomères à tétramères). B) Relation linéaire entre le contraste mesuré expérimentalement et la masse molaire des oligomères de BSA. La droite rouge montre le résultat du modèle linéaire utilisé où  $m$  est la masse molaire des échantillons analysés.

## Déroulement d'une expérience de photométrie de masse :

À la section 5.1.1.2, je précise que les échantillons doivent être dilués avant de commencer une acquisition. En effet, au moment de l'explication du protocole d'utilisation générale de l'appareil, Bertrand Raynal m'a indiqué qu'il est nécessaire de déposer une goutte de 15 à 18  $\mu\text{L}$  de tampon sur la lamelle afin que le logiciel puisse trouver le plan d'observation. Dans un second temps, 2 à 5  $\mu\text{L}$  d'échantillon doivent être mélangés à la première goutte juste avant le début de l'acquisition pour obtenir un volume total de 20  $\mu\text{L}$ . En d'autres termes, un facteur de dilution pouvant varier de 4 à 10 doit être appliqué sur l'échantillon. Pour la majorité de nos acquisitions, les échantillons ont été dilués cinq fois, ce qui correspond à une concentration finale en ARN, et donc maximale en nucléocapsides, égale à 200 pM ou 1 nM selon la condition d'assemblage considérée. Ces valeurs doivent être comparées à la gamme de concentration détectée par l'appareil : 100 pM à 100 nM théoriquement mais plutôt entre 1 et 10 nM pour un fonctionnement optimal.

Cette étape de dilution avait posé un problème lors de la première journée d'expérience puisque toutes les conditions d'assemblage avaient été préparées à partir de 1 nM en ARN. La statistique en particules détectées ayant été vraiment faible, de l'ordre de quelques dizaines de points, nous n'avons pas pu observer des différences significatives entre les conditions expérimentales testées. En plus de ça, ce jour-là nous avons été limité au plus petit champ de vue proposé par l'appareil (environ 30  $\mu\text{m}^2$ ). Ainsi, pour les résultats présentés à la section 5.1.2, les concentrations en ARN ont, pour la plupart, été augmentées

à 5 nM et le champ de vue le plus grand ( $200 \mu\text{m}^2$ ) a pu être sélectionné. Enfin, chaque acquisition a duré 2 minutes.

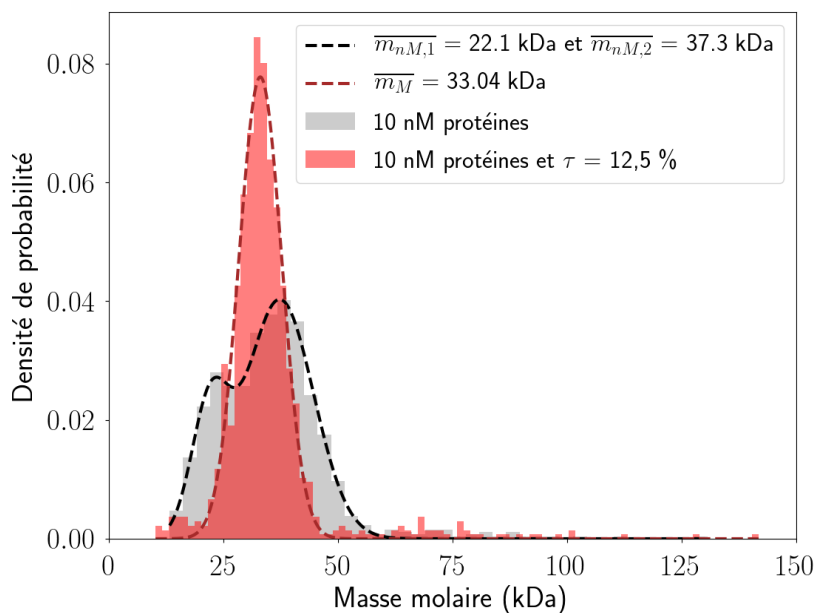
Remarques :

- Pour trouver le plan d'observation, un volume de 10  $\mu\text{L}$  de tampon peut être déposé sur la lamelle, ce qui limite le facteur de dilution à deux. Cette précision est mentionnée sur le mode d'emploi fourni par *Refeyn* mais nous a été déconseillée par Bertrand Raynal à cause d'une potentielle complication de la recherche du plan de focalisation.

- J'ai appris plus tard qu'il existe un mode appelé *Buffer free*, avec lequel la recherche du plan de focalisation peut être réalisée directement avec l'échantillon, ce qui élimine le problème lié à la dilution. Cependant, dans le cas d'une utilisation normale de l'appareil, la recherche du plan d'observation permet de voir si la zone de la lamelle utilisée ne contient pas de saletés. Si c'est le cas, il est possible de changer de zone pour en trouver une meilleure, sans que les particules commencent à se déposer sur la surface. En utilisant le mode *Buffer free*, s'il faut changer de zone, cela doit avoir lieu alors que l'échantillon a déjà été déposé sur la lamelle. Il ne faut pas perdre de temps au risque de saturer la lamelle et de ne plus observer de particules arriver sur la surface. C'est d'ailleurs pour cette raison que les acquisitions ne durent pas plus de deux minutes. En effet, j'ai pu remarquer qu'après une minute, le nombre d'objets qu'on voit se fixer à la surface est déjà moins élevé qu'au début.

## Caractérisation des dimères :

La Figure G.2 illustre les distributions obtenues lors de la caractérisation de deux échantillons : un premier contenant exclusivement des sous-unités non marquées (stockées dans le tampon Dim : 50 mM HEPES, 500 mM NaCl, 1 mM EDTA, 1 mM DTT et 0,5 mM PMSF à pH = 7,5) et un autre où environ 12,5 % des protéines sont marquées (stockées dans le tampon M : 100 mM HEPES et 5 mM MgCl<sub>2</sub> à pH = 8,2).



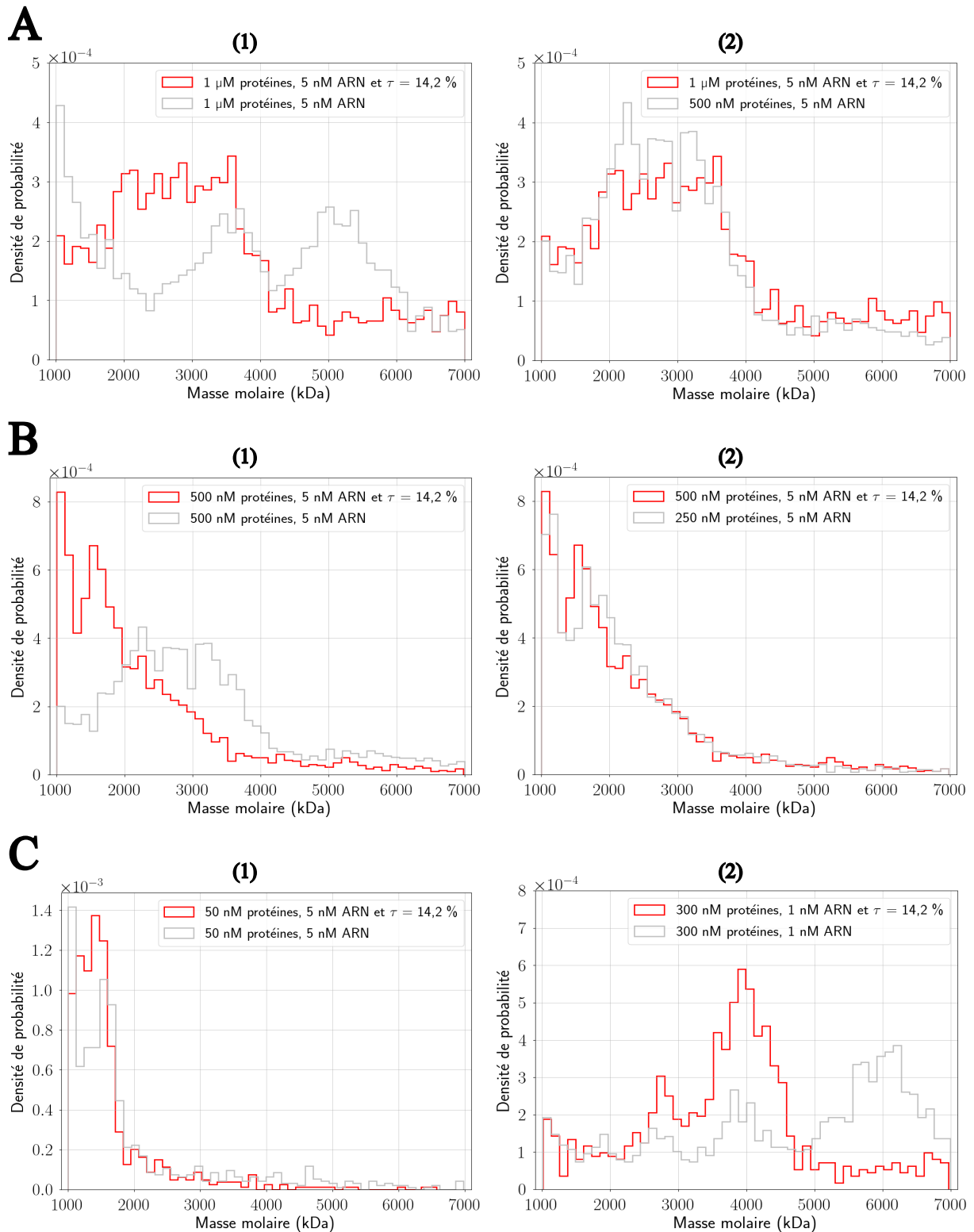
**FIGURE G.2 – Photométrie de masse : caractérisation de dimères marqués et non marqués.** Distribution des masses molaires de dimères marqués et non marqués respectivement représentées en rouge et gris clair. La courbe marron correspond à un ajustement gaussien dont la valeur moyenne est notée  $\overline{m_M}$  pour les dimères marqués. La courbe noire a été tracée à partir de deux distributions gaussiennes dont les valeurs moyennes sont données par  $\overline{m_{nM,1}}$  et  $\overline{m_{nM,2}}$  pour les dimères non marqués. Les protéines non marquées ont été diluées dans le tampon de fin de purification Dim alors que pour les espèces marquées, il s'agit du tampon de marquage M. Le taux de marquage est de 12,5 %.

Nous pouvons voir que dans les deux cas, le ou les pics de distribution sont légèrement sous-estimés par rapport à la masse des dimères qui est de 40,6 kDa. En ce qui concerne les sous-unités non marquées, il est difficile d'expliquer le pic observé aux alentours de 22,1 kDa. Le plus cohérent serait qu'il corresponde aux monomères de protéine mais l'appareil ne permet pas, en théorie, de détecter des objets dont la masse est inférieure à 30 kDa. Par ailleurs, le monomère est supposé être moins stable en solution que le dimère.

## Comparaison des particules assemblées avec ou sans marquage :

Cette dernière partie vient compléter ce qui a été mentionné à la fin de la section 5.1.2.2. En effet, je mets en valeur que même si des particules similaires à des nucléo-capsides peuvent être formées avec marquage, la population globale d'objets formés est différente de celle issue d'un assemblage sans protéines marquées.

Sur les Figures G.3.A et B, nous voyons que les distributions tracées pour 1  $\mu\text{M}$  et 500 nM en protéines, avec marquage, correspondent presque parfaitement à celles obtenues respectivement pour 500 nM et 250 nM sans marquage. Ce décalage exact dans ces deux cas est assez difficile à expliquer puisque d'un côté les masses des objets sont réparties de manière très similaire à 50 nM en protéines avec ou sans marquage (Figure G.3.C.1). De l'autre, pour le rapport en masse  $\frac{\text{protéines}}{\text{ARN}}$  le plus élevé, avec 300 nM en protéines et 1 nM en ARN, les pics situés à 3,9 MDa sont quasiment identiques, comme le montre la Figure G.3.C.2. En revanche, en absence de fluorophores, la population d'objets plus gros que les virions est beaucoup plus déterminée. Le marquage semble donc limiter l'apparition de ce type de particules comme cela est mis en avant à la section 5.1.2.3.



**FIGURE G.3 – Photométrie de masse : illustration de la différence de masse des objets assemblés avec et sans marquage.** A) 1  $\mu\text{M}$  en protéines avec marquage et (1) 1  $\mu\text{M}$  ou (2) 500 nM en protéines sans marquage. B) 500 nM en protéines avec marquage et (1) 500 nM ou (2) 250 nM en protéines sans marquage. C) (1) 50 nM en protéines avec et sans marquage, (2) 300 nM en protéines avec et sans marquage. Le taux de marquage est de 14,2 %.

## Annexe H

# Dynamique d'assemblage à l'équilibre

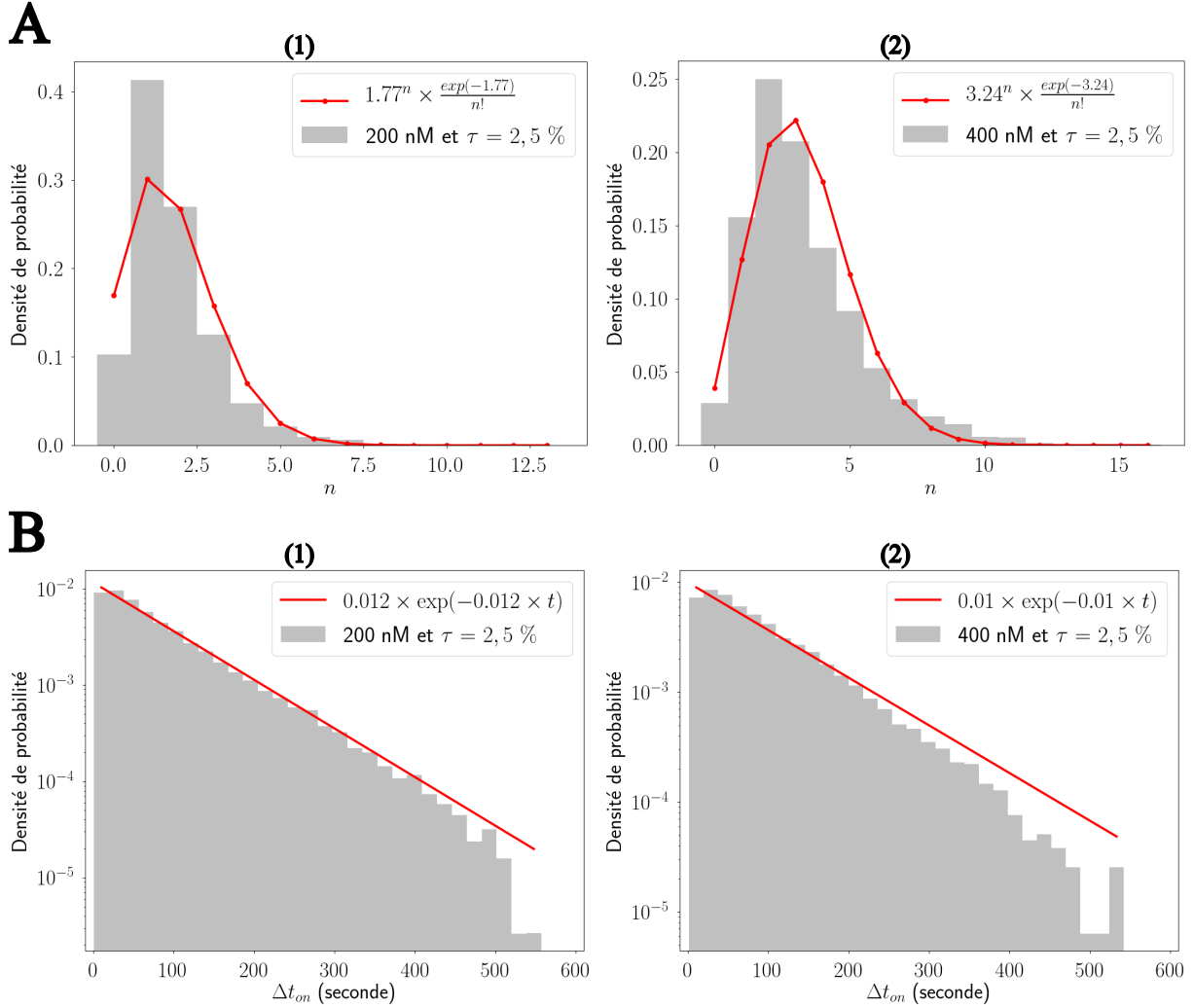
Dans la section 5.4, je présente les résultats que j'ai pu obtenir suite aux expériences de dynamique d'assemblage à l'équilibre. Pour compléter certaines informations données dans cette partie, je rajoute dans cette annexe des distributions supplémentaires pour  $n$  et  $\Delta t_{on}$ . Il est également possible de retrouver un tableau résumant les valeurs des paramètres d'assemblage mesurées expérimentalement pour chaque lot de protéines.

### Caractérisation des distributions de $n$ et $\Delta t_{on}$ :

Les deux distributions présentées à la section 5.4.3.1 pour une concentration totale en protéines de 50 nM nous permettent d'observer l'hétérogénéité du processus d'auto-assemblage des capsides de CCMV. Dans un second temps, il est possible de calculer des valeurs moyennes afin d'en déduire des comportements globaux, comme lorsque je présente l'évolution de  $\langle n_{tot} \rangle$  et  $\frac{k_{on}^{tot}}{C}$  en fonction de la concentration en protéines. Sur la Figure H.1, je montre les distributions de  $n$  et  $\Delta t_{on}$  pour 200 et 400 nM en protéines issues du lot 1.

Comme pour une concentration de 50 nM, nous pouvons voir que les distributions observées pour  $n$  suivent bien l'allure attendue de la loi de Poisson. Dans les deux cas illustrées par les Figures H.1.A.1-2, nous obtenons un nombre moyen de sous-unités marquées  $\langle n \rangle$  par ARN de 1,8 pour 200 nM et 3,2 pour 400 nM. En ce qui concerne la constante d'association  $k_{on}$ , nous pouvons voir qu'elle vaut environ  $0,012 \text{ s}^{-1}$  à 200 nM contre  $0,010 \text{ s}^{-1}$  à 400 nM. Cela correspond à une durée moyenne entre deux événements d'accroche en dimères marqués d'environ 83 et 100 secondes. Par ailleurs, nous pouvons voir que, comme à 50 nM, les ajustements exponentiels tracés en rouge traduisent bien l'allure de la distribution des  $\Delta t_{on}$ , sauf pour les plus grandes valeurs temporelles. À la différence de ce qui a été observé pour 50 nM, nous observons une surestimation de l'ajustement par rapport à l'allure de la distribution. Il pourrait être intéressant de voir si nous retrouvons ce comportement en simulation, en réduisant la valeur rentrée pour  $k_{on}$  (qui

est de  $0,015 \text{ s}^{-1}$  pour les résultats détaillés à la section 4.3).



**FIGURE H.1 – Dynamique à l'équilibre : distributions de  $n$  et  $\Delta t_{on}$  pour 200 et 400 nM en protéines.** Histogrammes A) du nombre de sous-unités marquées  $n$  estimées en début d'acquisition et B) de la durée entre deux événements d'accroche consécutifs  $\Delta t_{on}$ . Les figures (1) illustrent les résultats à 200 nM en protéines contre 400 nM pour celles numérotées (2). Les ajustements par une loi de Poisson et exponentielle sont tracés en rouge. Le taux de marquage est de 2,5 %.

## Résumé des paramètres d'assemblage mesurés expérimentalement :

Sur la Figure 5.18 de la section 5.4.3.2, je montre les évolutions des différents paramètres d'assemblage auxquels nous nous intéressons. Pour compléter ces informations, le Tableau H.1 regroupe les valeurs moyennes et écarts-types de  $\langle n \rangle$ ,  $k_{on}$  et  $\tau_u$  pour les différents lots de protéines étudiés. Je précise que les taux de marquage sont d'environ 10 % lorsque la concentration en protéines est de 10 et 50 nM (9,4 % pour le lot 1), contre 2,5 % pour 200 et 400 nM. Pour tous les paramètres d'assemblage, les valeurs ont été moyennées sur les résultats obtenus à partir de chaque cellule micro-fluidique. Pour toutes les

conditions expérimentales, sauf celle à 400 nM du lot 1, plusieurs journées d'expériences ont été combinées.

	$C_{\text{prot}}$ (nM)	$N_{\text{cellules}}$	$\langle n \rangle$	$k_{\text{on}}$ ( $s^{-1}$ )	$\tau_u$ (s)
<b>Lot 1</b>	50	9	$2,6 \pm 0,6$	$0,016 \pm 0,004$	$168 \pm 62$
	200	8	$1,8 \pm 0,2$	$0,012 \pm 0,001$	$153 \pm 25$
	400	3	$3,2 \pm 0,1$	$0,010 \pm 3 \times 10^{-4}$	$319 \pm 20$
<b>Lot 2</b>	10	8	$0,9 \pm 0,2$	$0,030 \pm 0,006$	$31 \pm 9$
	50	10	$1,9 \pm 0,3$	$0,018 \pm 0,006$	$120 \pm 56$
	200	11	$1,1 \pm 0,2$	$0,020 \pm 0,006$	$61 \pm 24$
	400	13	$1,1 \pm 0,2$	$0,017 \pm 0,005$	$68 \pm 22$
<b>Lot 3</b>	10	8	$1,9 \pm 0,9$	$0,023 \pm 0,006$	$90 \pm 40$
	50	10	$2,4 \pm 1,2$	$0,018 \pm 0,004$	$128 \pm 53$
	200	11	$1,4 \pm 0,1$	$0,016 \pm 0,002$	$85 \pm 12$
	400	13	$1,5 \pm 0,2$	$0,014 \pm 0,002$	$108 \pm 21$

**TABLEAU H.1** – Dynamique à l'équilibre : résumé des valeurs des paramètres d'assemblage pour les trois lots de protéines.  $N_{\text{cellules}}$  indique le nombre de cellules micro-fluidiques qui ont été prises en compte dans l'analyse et  $C_{\text{prot}}$  correspond à la concentration totale en protéines.





# Bibliographie

- J. A. Abels, F. Moreno-Herrero, T. van der Heijden, C. Dekker, and N. H. Dekker. Single-Molecule Measurements of the Persistence Length of Double-Stranded RNA. *Biophysical Journal*, 88(4), 2005. ISSN 0006-3495. doi : 10.1529/biophysj.104.052811.
- K. W. Adolph and P. J. G. Butler. Assembly of a spherical plant virus. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 276(943), 1976. doi : 10.1098/rstb.1976.0102.
- M. Adrian, J. Dubochet, S. D. Fuller, and J. R. Harris. Cryo-negative staining. *Micron*, 29(2), 1998. doi : [https://doi.org/10.1016/S0968-4328\(97\)00068-1](https://doi.org/10.1016/S0968-4328(97)00068-1).
- C. E. Aitken, R. A. Marshall, and J. D. Puglisi. An Oxygen Scavenging System for Improvement of Dye Stability in Single-Molecule Fluorescence Experiments. *Biophysical Journal*, 94(5), 2008. doi : 10.1529/biophysj.107.117689.
- A. Ali and M. J. Roossinck. Rapid and efficient purification of *Cowpea chlorotic mottle virus* by sucrose cushion ultracentrifugation. *Journal of Virological Methods*, 141(1), 2007. doi : 10.1016/j.jviromet.2006.11.038.
- D. Andres, Y. Roske, C. Doering, U. Heinemann, R. Seckler, and S. Barbirz. Tail morphology controls DNA release in two Salmonella phages with one lipopolysaccharide receptor recognition system. *Molecular Microbiology*, 83(6), 2012. doi : 10.1111/j.1365-2958.2012.08006.x.
- D. Baltimore. Expression of animal virus genomes. *Bacteriological reviews*, 35(3), 1971. doi : 10.1128/br.35.3.235-241.1971.
- J. B. Bancroft. The self-assembly of spherical plant viruses. *Advances in Virus Research*, 16, 1970. doi : 10.1016/s0065-3527(08)60022-6.
- J. B. Bancroft and E. Hiebert. Formation of an infectious nucleoprotein from protein and nucleic acid isolated from a small spherical virus. *Virology*, 32(2), 1967. doi : 10.1016/0042-6822(67)90284-X.
- N. Barbier. *Étude des erreurs programmées du ribosome mammifère par microscopie de fluorescence en molécule unique*. PhD thesis, Université Paris-Saclay, 2017.

- A. Borodavka, R. Tuma, and P. G. Stockley. Evidence that viral RNAs have evolved for efficient, two-stage packaging. *Proceedings of the National Academy of Sciences*, 109(39), 2012. doi : 10.1073/pnas.1204357109.
- B. Brandenburg, L. Y. Lee, M. Lakadamyali, M. J. Rust, X. Zhuang, and J. M. Hogle. Imaging Poliovirus Entry in Live Cells. *PLOS Biology*, 5(7), 2007. doi : 10.1371/journal.pbio.0050183.
- R. F. Bruinsma and W. S. Klug. Physics of viral shells. *Annual Review of Condensed Matter Physics*, 6(1), 2015. doi : 10.1146/annurev-conmatphys-031214-014325.
- R. F. Bruinsma, G. J. Wuite, and W. H. Roos. Physics of viral dynamics. *Nature Research*, 3(2), 2021. doi : 10.1038/s42254-020-00267-1.
- P. Buzon, S. Maity, P. Christodoulis, M. J. Wiertsema, S. Dunkelbarger, C. Kim, G. J. Wuite, A. Zlotnick, and W. H. Roos. Virus self-assembly proceeds through contact-rich energy minima. *Science Advances*, 7(45), 2021. doi : 10.1126/sciadv.abg0811.
- R. D. Cadena-Nava, M. Comas-Garcia, R. F. Garmann, A. L. N. Rao, C. M. Knobler, and W. M. Gelbart. Self-Assembly of Viral Capsid Protein and RNA Molecules of Different Sizes : Requirement for a Specific High Protein/RNA Mass Ratio. *Journal of Virology*, 86(6), 2012. doi : 10.1128/jvi.06566-11.
- R. D. Cadena-Nava, M. Comas-Garcia, R. F. Garmann, A. L. N. Rao, C. M. Knobler, and W. M. Gelbart. Correction for Cadena-Nava et al., “Self-Assembly of Viral Capsid Protein and RNA Molecules of Different Sizes : Requirement for a Specific High Protein/RNA Mass Ratio”. *Journal of Virology*, 93(4), 2019. doi : 10.1128/jvi.01939-18.
- G. L. Casini, D. Graham, D. Heine, R. L. Garcea, and D. T. Wu. In vitro papillomavirus capsid assembly analyzed by light scattering. *Virology*, 325(2), 2004. doi : 10.1016/j.virol.2004.04.034.
- D. L. D. Caspar and A. Klug. Physical principles in the construction of regular viruses. *Cold Spring Harbor Symposia on Quantitative Biology*, 27, 1962. doi : 10.1101/SQB.1962.027.001.005.
- M. Chevreuil. *Phénomènes dynamiques d’auto-assemblage et de désassemblage dans des virus icosaédriques*. PhD thesis, Université Paris-Saclay, 2020.
- M. Chevreuil, D. Law-Hine, J. Chen, S. Bressanelli, S. Combet, D. Constantin, J. Degrouard, J. Möller, M. Zeghal, and G. Tresset. Nonequilibrium self-assembly dynamics of icosahedral viral capsids packaging genome or polyelectrolyte. *Nature Communications*, 9(1), 2018. doi : 10.1038/s41467-018-05426-8.

- M. Chevreuil, L. Lecoq, S. Wang, L. Gargowitsch, N. Nhiri, E. Jacquet, T. Zinn, S. Fieulaine, S. Bressanelli, and G. Tresset. Nonsymmetrical Dynamics of the HBV Capsid Assembly and Disassembly Evidenced by Their Transient Species. *The Journal of Physical Chemistry B*, 124(45), 2020. doi : 10.1021/acs.jpcc.0c05024.
- D. Cole, G. Young, A. Weigel, A. Sebesta, and P. Kukura. Label-Free Single-Molecule Imaging with Numerical-Aperture-Shaped Interferometric Scattering Microscopy. *ACS Photonics*, 4(2), 2017. doi : 10.1021/acsp Photonics.6b00912.
- M. Comas-Garcia, R. D. Cadena-Nava, A. L. N. Rao, C. M. Knobler, and W. M. Gelbart. In Vitro Quantification of the Relative Packaging Efficiencies of Single-Stranded RNA Molecules by Viral Capsid Protein. *Journal of Virology*, 86(22), 2012. doi : 10.1128/jvi.01695-12.
- A. Dance. Beyond coronavirus : the virus discoveries transforming biology. *Nature*, 595(7865), 2021. doi : 10.1038/d41586-021-01749-7.
- D. Dulin, S. Barland, X. Hachair, and F. Pedaci. Efficient Illumination for Microsecond Tracking Microscopy. *PLOS ONE*, 9(9), 2014. ISSN 1932-6203. doi : 10.1371/journal.pone.0107335.
- A. L. Duran-Meza, M. V. Villagrana-Escareño, J. Ruiz-García, C. M. Knobler, and W. M. Gelbart. Controlling the surface charge of simple viruses. *PLOS ONE*, 16(9), 2021. doi : 10.1371/journal.pone.0255820.
- E. C. Dykeman, P. G. Stockley, and R. Twarock. Packaging Signals in Two Single-Stranded RNA Viruses Imply a Conserved Assembly Mechanism and Geometry of the Packaged Genome. *Journal of Molecular Biology*, 425(17), 2013. doi : 10.1016/j.jmb.2013.06.005.
- S. W. Englander, D. B. Calhoun, and J. J. Englander. Biochemistry without oxygen. *Analytical Biochemistry*, 161(2), 1987. doi : 10.1016/0003-2697(87)90454-4.
- S. Faez, Y. Lahini, S. Weidlich, R. F. Garmann, K. Wondraczek, M. Zeisberger, M. A. Schmidt, M. Orrit, and V. N. Manoharan. Fast, Label-Free Tracking of Single Viruses and Weakly Scattering Nanoparticles in a Nanofluidic Optical Fiber. *ACS Nano*, 9(12), 2015. doi : 10.1021/acsnano.5b05646.
- H. Fraenkel-Conrat and R. C. Williams. Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components. *Proceedings of the National Academy of Sciences of the United States of America*, 41(10), 1955. doi : 10.1073/pnas.41.10.690.
- S. D. Fuerstenau and W. H. Benner. Molecular weight determination of megadalton DNA electrospray ions using charge detection time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, 9(15), 1995. doi : 10.1002/rcm.1290091513.

- R. F. Garmann, M. Comas-Garcia, A. Gopal, C. M. Knobler, and W. M. Gelbart. The Assembly Pathway of an Icosahedral Single-Stranded RNA Virus Depends on the Strength of Inter-Subunit Attractions. *Journal of Molecular Biology*, 426(5), 2014. doi : 10.1016/j.jmb.2013.10.017.
- R. F. Garmann, R. Sportsman, C. Beren, V. N. Manoharan, C. M. Knobler, and W. M. Gelbart. A Simple RNA-DNA Scaffold Templates the Assembly of Monofunctional Virus-Like Particles. *Journal of the American Chemical Society*, 137(24), 2015. doi : 10.1021/jacs.5b03770.
- R. F. Garmann, A. M. Goldfain, and V. N. Manoharan. Measurements of the self-assembly kinetics of individual viral capsids around their RNA genome. *Proceedings of the National Academy of Sciences*, 116(45), 2019. doi : 10.1073/pnas.1909223116.
- R. F. Garmann, A. M. Goldfain, C. R. Tanimoto, C. E. Beren, F. F. Vasquez, D. A. Villarreal, C. M. Knobler, W. M. Gelbart, and V. N. Manoharan. Single-particle studies of the effects of RNA–protein interactions on the self-assembly of RNA virus particles. *Proceedings of the National Academy of Sciences*, 119(39), 2022. doi : 10.1073/pnas.2206292119.
- A. M. Goldfain. *Using Interferometric Scattering Microscopy to Study the Dynamics of Viruses*. PhD thesis, Harvard University, 2018.
- I. González-Domínguez, E. Puente-Massaguer, L. Cervera, and F. Gòdia. Quantification of the HIV-1 virus-like particle production process by super-resolution imaging : From VLP budding to nanoparticle analysis. *Biotechnology and Bioengineering*, 117(7), 2020. doi : 10.1002/bit.27345.
- S. E. Green and A. Reid. *Viruses Throughout Life and Time : Friends, Foes, Change Agents*. American Society for Microbiology, 2013.
- M. F. Hagan, O. M. Elrad, and R. L. Jack. Mechanisms of kinetic trapping in self-assembly and phase transformation. *The Journal of Chemical Physics*, 135(10), 2011. doi : 10.1063/1.3635775.
- N.-P. Huang, R. Michel, J. Voros, M. Textor, R. Hofer, A. Rossi, D. L. Elbert, J. A. Hubbell, and N. D. Spencer. Poly(l-lysine)-g-poly(ethylene glycol) Layers on Metal Oxide Surfaces : Surface-Analytical Characterization and Resistance to Serum and Fibrinogen Adsorption. *Langmuir*, 17(2), 2001. doi : 10.1021/la000736+.
- N.-P. Huang, J. Voros, S. M. De Paul, M. Textor, and N. D. Spencer. Biotin-Derivatized Poly(l-lysine)-g-poly(ethylene glycol) : A Novel Polymeric Interface for Bioaffinity Sensing. *Langmuir*, 18(1), 2002. doi : 10.1021/la010913m.

- C. G. Hubner, A. Renn, I. Renge, and U. P. Wild. Direct observation of the triplet lifetime quenching of single dye molecules by molecular oxygen. *The Journal of Chemical Physics*, 115(21), 2001. doi : 10.1063/1.1421382.
- J. Hummert, K. Yserentant, T. Fink, J. Euchner, Y. X. Ho, S. A. Tashev, and D.-P. Herten. Photobleaching step analysis for robust determination of protein complex stoichiometries. *Molecular Biology of the Cell*, 32(21), 2021. doi : 10.1091/mbc.E20-09-0568.
- V. Jacobsen, P. Stoller, C. Brunner, V. Vogel, and V. Sandoghdar. Interferometric optical detection and tracking of very small gold nanoparticles at a water-glass interface. *Opt. Express, OE*, 14(1), 2006. doi : 10.1364/OPEX.14.000405.
- S. Kaneshashi, K. Ishizu, M. Kawano, S. Han, S. Tomita, H. Watanabe, K. Kataoka, and H. Handa. Simian virus 40 VP1 capsid protein forms polymorphic assemblies in vitro. *Journal of General Virology*, 84(7), 2003. doi : 10.1099/vir.0.19067-0.
- A. Klug. The tobacco mosaic virus particle : structure and assembly. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 354(1383), 1999. doi : 10.1098/rstb.1999.0404.
- P. Kukura, H. Ewers, C. Müller, A. Renn, A. Helenius, and V. Sandoghdar. High-speed nanoscopic tracking of the position and orientation of a single virus. *Nature Methods*, 6(12), 2009. doi : 10.1038/nmeth.1395.
- R. Kumar, B. Yang, J. Barton, M. Stejfova, A. Schafer, M. Koenig, P. Knittel, P. Cigler, and M. Hirtz. Diamond Surfaces with Clickable Antifouling Polymer Coating for Microarray-Based Biosensing. *Advanced Materials Interfaces*, 9(33), 2022. doi : 10.1002/admi.202201453.
- K. C. Kwok, K. M. Yeung, and N. H. Cheung. Adsorption Kinetics of Bovine Serum Albumin on Fused Silica : Population Heterogeneities Revealed by Single-Molecule Fluorescence Microscopy. *Langmuir*, 23(4), 2007. doi : 10.1021/la061779e.
- A. Le Gall. *Pince optique et microscopie de fluorescence pour l'étude de la synthèse des protéines en molécule unique*. PhD thesis, Paris 11, 2011.
- C. M. Ling, P. P. Hung, and L. R. Overby. Specificity in self-assembly of bacteriophages Q $\beta$  and MS2. *Biochemistry*, 8(11), 1969. doi : 10.1021/bi00839a036.
- C. A. Lutomski, N. A. Lykтей, E. E. Pierson, Z. Zhao, A. Zlotnick, and M. F. Jarrold. Multiple Pathways in Capsid Assembly. *Journal of the American Chemical Society*, 140(17), 2018. doi : 10.1021/jacs.8b01804.
- M. Medrano, M. A. Fuertes, A. Valbuena, P. J. P. Carrillo, A. Rodriguez-Huete, and M. G. Mateu. Imaging and Quantitation of a Succession of Transient Intermediates Reveal

- the Reversible Self-Assembly Pathway of a Simple Icosahedral Virus Capsid. *Journal of the American Chemical Society*, 138(47), 2016. doi : 10.1021/jacs.6b07663.
- J.-P. Michel, M. Gingery, and L. Lavelle. Efficient purification of bromoviruses by ultrafiltration. *Journal of Virological Methods*, 122(2), 2004. doi : 10.1016/j.jviromet.2004.09.005.
- R. Michel, S. Pasche, M. Textor, and D. G. Castner. Influence of PEG Architecture on Protein Adsorption and Conformation. *Langmuir*, 21(26), 2005. doi : 10.1021/la051726h.
- J. Ortega Arroyo, D. Cole, and P. Kukura. Interferometric scattering microscopy and its combination with single-molecule fluorescence imaging. *Nature Protocol*, 11(4), 2016. doi : 10.1038/nprot.2016.022.
- S. Panahandeh, S. Li, L. Marichal, R. Leite Rubim, G. Tresset, and R. Zandi. How a Virus Circumvents Energy Barriers to Form Symmetric Shells. *ACS Nano*, 14(3), 2020. ISSN 1936-0851. doi : 10.1021/acsnano.9b08354.
- J. H. Park, J. A. Jackman, A. R. Ferhan, G. J. Ma, B. K. Yoon, and N.-J. Cho. Temperature-Induced Denaturation of BSA Protein Molecules for Improved Surface Passivation Coatings. *ACS Applied Materials and Interfaces*, 10(38), 2018. doi : 10.1021/acсами.8b13749.
- S. Pasche, S. M. De Paul, J. Voros, N. D. Spencer, and M. Textor. Poly(l-lysine)-graft-poly(ethylene glycol) Assembled Monolayers on Niobium Oxide Surfaces : A Quantitative Study of the Influence of Polymer Interfacial Architecture on Resistance to Protein Adsorption by ToF-SIMS and in Situ OWLS. *Langmuir*, 19(22), 2003. doi : 10.1021/la034111y.
- J. D. Perlmutter, M. R. Perkett, and M. F. Hagan. Pathways for Virus Assembly around Nucleic Acids. *Journal of Molecular Biology*, 426(18), 2014. doi : 10.1016/j.jmb.2014.07.004.
- S. S. Perry, X. Yan, F. T. Limpoco, S. Lee, M. Müller, and N. D. Spencer. Tribological Properties of Poly(l-lysine)-graft-poly(ethylene glycol) Films : Influence of Polymer Architecture and Adsorbed Conformation. *ACS Applied Materials and Interfaces*, 1(6), 2009. doi : 10.1021/am900101m.
- J. Z. Porterfield and A. Zlotnick. A simple and general method for determining the protein and nucleic acid content of viruses by UV absorbance. *Virology*, 407(2), 2010. doi : 10.1016/j.virol.2010.08.015.
- O. Rolfsson, S. Middleton, I. W. Manfield, S. J. White, B. Fan, R. Vaughan, N. A. Ranson, E. Dykeman, R. Twarock, J. Ford, C. Cheng Kao, and P. G. Stockley. Direct Evidence

- for Packaging Signal-Mediated Assembly of Bacteriophage MS2. *Journal of Molecular Biology*, 428(2, Part B), 2016. doi : 10.1016/j.jmb.2015.11.014.
- J. A. Speir, S. Munshi, G. Wang, T. S. Baker, and J. E. Johnson. Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure*, 3(1), 1995. doi : 10.1016/S0969-2126(01)00135-6.
- M. Steinitz. Quantitation of the Blocking Effect of Tween 20 and Bovine Serum Albumin in ELISA Microwells. *Analytical Biochemistry*, 282(2), 2000. doi : 10.1006/abio.2000.4602.
- G. Tresset, J. Chen, M. Chevreuil, N. Nhiri, E. Jacquet, and Y. Lansac. Two-Dimensional Phase Transition of Viral Capsid Gives Insights into Subunit Interactions. *Physical Review Applied*, 7(1), 2017. doi : 10.1103/PhysRevApplied.7.014005.
- C. Truong, L. Oudre, and N. Vayatis. Selective review of offline change point detection methods. *Signal Processing*, 167, 2020. doi : 10.1016/j.sigpro.2019.107299.
- H. Tsukamoto, M. Kawano, T. Inoue, T. Enomoto, R. Takahashi, N. Yokoyama, N. Yamamoto, T. Imai, K. Kataoka, Y. Yamaguchi, and H. Handa. Evidence that SV40 VP1–DNA interactions contribute to the assembly of 40-nm spherical viral particles. *Genes to Cells*, 12(11), 2007. doi : 10.1111/j.1365-2443.2007.01134.x.
- M. van de Waterbeemd, J. Snijder, I. B. Tsvetkova, B. G. Dragnea, J. J. Cornelissen, and A. J. R. Heck. Examining the Heterogeneous Genome Content of Multipartite Viruses BMV and CCMV by Native Mass Spectrometry. *Journal of the American Society for Mass Spectrometry*, 27(6), 2016. doi : 10.1007/s13361-016-1348-6.
- V. Vienne. *Étude de la cinétique de scanning lors de l'initiation de la traduction par microscopie de fluorescence en molécule unique*. PhD thesis, Université Paris-Saclay, 2023.
- J. Vogelsang, R. Kasper, C. Steinhauer, B. Person, M. Heilemann, M. Sauer, and P. Tinnefeld. A Reducing and Oxidizing System Minimizes Photobleaching and Blinking of Fluorescent Dyes. *Angewandte Chemie International Edition*, 47(29), 2008. doi : 10.1002/anie.200801518.
- J. Vogelsang, T. Cordes, C. Forthmann, C. Steinhauer, and P. Tinnefeld. Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. *Proceedings of the National Academy of Sciences*, 106(20), 2009. doi : 10.1073/pnas.0811875106.
- M. Yamashita and J. B. Fenn. Electrospray ion source. Another variation on the free-jet theme. *The Journal of Physical Chemistry*, 88(20), 1984. doi : 10.1021/j150664a002.



- K. M. Yeung, Z. J. Lu, and N. H. Cheung. Adsorption of bovine serum albumin on fused silica : Elucidation of protein–protein interactions by single-molecule fluorescence microscopy. *Colloids and Surfaces B : Biointerfaces*, 69(2), 2009. doi : 10.1016/j.colsurfb.2008.11.020.
- G. Young, N. Hundt, D. Cole, A. Fineberg, J. Andrecka, A. Tyler, A. Olerinyova, A. Ansari, E. G. Marklund, M. P. Collier, S. A. Chandler, O. Tkachenko, J. Allen, M. Crispin, N. Billington, Y. Takagi, J. R. Sellers, C. Eichmann, P. Selenko, L. Frey, R. Riek, M. R. Galpin, W. B. Struwe, J. L. P. Benesch, and P. Kukura. Quantitative mass imaging of single biological macromolecules. *Science*, 360(6387), 2018. doi : 10.1126/science.aar5839.
- G. Zhao, J. R. Perilla, E. L. Yufenyuy, X. Meng, B. Chen, J. Ning, J. Ahn, A. M. Gronenborn, K. Schulten, C. Aiken, and P. Zhang. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, 497(7451), 2013. doi : 10.1038/nature12162.
- A. Zlotnick, J. M. Johnson, P. W. Wingfield, S. J. Stahl, and D. Endres. A Theoretical Model Successfully Identifies Features of Hepatitis B Virus Capsid Assembly. *Biochemistry*, 38(44), 1999. doi : 10.1021/bi991611a.