



**HAL**  
open science

# étude structurale et fonctionnelle de la protéine a1 du bactériophage t5 : une dnase octamérique originale

Léo Zangelmi

► **To cite this version:**

Léo Zangelmi. étude structurale et fonctionnelle de la protéine a1 du bactériophage t5 : une dnase octamérique originale. Biochimie, Biologie Moléculaire. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLS505 . tel-04368850

**HAL Id: tel-04368850**

**<https://theses.hal.science/tel-04368850>**

Submitted on 2 Jan 2024

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

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

# Étude Structurale et Fonctionnelle de la Protéine A1 du Bactériophage T5 : une DNase octamérique originale

Thèse de doctorat de l'Université Paris-Saclay  
préparée à l'Université Paris-Sud

École doctorale n°569 Innovation thérapeutique : du fondamental à  
l'appliqué (ITFA)  
Spécialité de doctorat: Biochimie et Biologie Structurale

Thèse présentée et soutenue à Orsay, le 6 décembre 2018, par

**M. Léo Zangelmi**

Composition du Jury :

Stéphanie BURY-MONE Professeure, I2BC (CEA/CNRS/U. Paris-Sud)	Présidente
Laurent DEBARBIEUX Directeur de recherche, Institut Pasteur	Rapporteur
Wim BURMEISTER Professeur, IBS (CEA/CNRS/U. Grenoble-Alpes)	Rapporteur
François LECOINTE Chargé de recherche, Institut Micalis (INRA)	Examineur
Pascale BOULANGER Directrice de recherche, I2BC (CEA/CNRS/U. Paris-Sud)	Directrice de thèse

**Titre :** Etude structurale et fonctionnelle de la protéine A1 du bactériophage T5 : une DNase octamérique originale.

**Mots clés :** Bactériophage T5, DNase, Exonucléase, Structure, Cryomicroscopie électronique, SAXS

**Résumé :** Les bactériophages neutralisent les systèmes de défense et détournent les fonctions vitales de leur hôte pour favoriser leur multiplication. Les gènes de phages qui gouvernent cette prise de contrôle de l'hôte restent mal connus, pourtant leur caractérisation présente un intérêt majeur pour mettre à jour des fonctions bactériennes spécifiquement ciblées par les phages et pour concevoir de nouveaux agents antibactériens. Le phage T5 injecte son ADN dans la bactérie *Escherichia coli* en deux étapes. Seuls les gènes précoces codés par 8% du génome entrent dans la cellule et le transfert s'arrête. Leur expression induit la dégradation du chromosome de l'hôte et l'inactivation de ses systèmes de restriction et de réparation de l'ADN. Après quelques minutes, le reste de la molécule d'ADN est injecté, ce qui permet la production de nouveaux phages. Deux gènes précoces A1 et A2 ont été identifiés comme essentiels pour la reprise du transfert de l'ADN et A1 est également nécessaire pour induire la dégradation de l'ADN de l'hôte. A1 et A2 sont les deux seuls gènes connus pour être impliqués dans la régulation de ce système original d'infection, mais leur fonction n'a jamais été identifiée.

Ma thèse porte sur la caractérisation fonctionnelle et structurale des protéines A1 et A2. J'ai purifié A1 et démontré *in vitro* qu'elle avait une activité DNase dépendante du manganèse. Sa structure atomique a été résolue par cryomicroscopie électronique à 3.01 Å de résolution, montrant une organisation octamérique de symétrie D4 inédite pour une DNase.

Chaque monomère (61kDa) contient un domaine exonuclease dont le site actif lie deux ions  $Mn^{2+}$  et qui s'apparente au site catalytique des domaines exonucléases de la DNA polymerase II et des DNases associées aux systèmes de recombinaison homologue et de réparation de l'ADN comme Mre11. En construisant différents mutants de A1, j'ai identifié certains acides aminés essentiels pour l'activité catalytique et, par des expériences de complémentation fonctionnelle, j'ai montré que cette activité était indispensable pour l'infection. L'ensemble de ces résultats suggèrent que A1 est la DNase, jusqu'ici inconnue, responsable de la dégradation massive du génome de l'hôte au tout début de l'infection. Enfin, j'ai observé que la production de A1 pendant l'infection induit une forte activité recombinaison. De nombreux autres bactériophages qui n'appartiennent pas à la famille des T5virus produisent également une protéine similaire à A1 dont la fonction n'a jamais été identifiée. Ce travail est un premier pas vers la compréhension de son rôle dans le mécanisme général d'infection par les phages. Une deuxième partie de cette thèse porte sur la caractérisation structurale de A2. Des recherches de similarité indiquent la présence d'un domaine Helix-Turn-Helix typique des régulateurs transcriptionnels. J'ai purifié A2 et montré que cette protéine de 14 kDa est un dimère en solution. La caractérisation des propriétés biochimiques de A2 a permis de débiter l'étude de sa structure par RMN. Les résultats de ma thèse ont révélé la structure originale d'une DNase de bactériophage qui contrôle la dégradation du génome bactérien et la régulation du transport de l'ADN viral au début du cycle infectieux. Ces résultats soulèvent des questions intrigantes : comment l'ADN de T5 est-il protégé de l'activité DNase de A1 ? Comment A1 et A2 interagissent-elles lors des étapes de prise de contrôle de l'hôte ?



**Title :** Structural and functional study of bacteriophage T5 A1 protein: an original octameric DNase.

**Keywords :** Bacteriophage T5, DNase, Exonuclease, Structure, Cryo Electron Microscopy, SAXS

**Abstract :** Bacteriophages defeat bacterial defences and hijack host cell machineries to establish a favourable environment for their multiplication. Early-expressed viral genes that govern host takeover are highly diverse from one phage to another and most of them have no assigned function. They thus represent a pool of novel genes whose products potentially subvert bacterial cell vital functions and could help in designing new antibacterial strategies. T5 phage uses a unique 2-step mechanism to deliver its DNA into its host *Escherichia coli*. At the onset of the infection, only 8 % of the genome enter the cell before the transfer temporarily stops. Expression of the genes encoded by this DNA portion leads to host chromosome degradation and inactivation of host restriction and DNA mending systems. After a few minutes, T5 DNA transfer resumes, allowing further phage multiplication. A1 and A2 are early genes required for DNA transfer completion and A1 is also necessary to trigger host DNA degradation. A1 and A2 are the only two genes known to be involved in the regulation of this original infection system, but their function yet remains to be characterized. The objectives of this work were to characterize the function and structure of A1 and A2 proteins. I have purified the A1 protein and shown that it has a manganese-dependent DNase activity *in vitro*. Cryo Electron Microscopy at 3.01 Å resolution unravelled its structure, showing an octameric organization with a D4 symmetry, which is unprecedented for a DNase.

Each monomer (61 kDa) carries an exonuclease domain harbouring an active site with two Mn<sup>2+</sup> ions. This site is similar to those from the exonuclease domain of the DNA polymerase II and from DNases involved in DNA mending and recombination events like Mre11. I identified essential catalytic residues for the DNase activity and demonstrated that this activity is crucial for infection by engineering A1 mutant proteins and by doing functional complementation assays. Taken together, my results suggest that A1 could then be the elusive DNase responsible for the massive host genome degradation observed during T5 phage infection. Eventually, I uncovered a recombinase activity associated to A1 production during infection. Similar proteins to A1 with unknown functions are produced in several other bacteriophages outside of the T5virus family. This work is a first step towards understanding the role of this protein in the general mechanism of infection by bacteriophages. In a second part, I worked on the structural characterisation of A2 protein. Similarity searches revealed a helix-turn-helix domain typically found in transcriptional regulators. I purified and demonstrated the dimeric organisation of this 14-kDa protein in solution. This initial characterization of A2 has opened avenues for further NMR studies. During my Ph.D., I uncovered the structure of an original bacteriophage DNase that controls bacterial genome degradation and that regulates viral DNA transport at the beginning of the infectious cycle. These results open the intriguing question about the mechanism for T5 DNA protection from A1 DNase activity as well as about the interplay between A1 and A2 during the host takeover.







# Remerciements

Je souhaite remercier toutes les personnes qui m'ont soutenu dans cette thèse, que ce soit directement en m'aidant à faire des expériences ou à écrire ce manuscrit, ou indirectement en m'encourageant dans les moments difficiles.

Tout d'abord, je remercie ma directrice de thèse, **Pascale Boulanger**, pour son encadrement pendant ces trois années.

Je tiens à remercier toutes les personnes qui ont accepté de juger mon travail et je suis reconnaissant des commentaires précieux qui m'ont été faits pour améliorer ce manuscrit. Je remercie donc **Stéphanie Bury-Moné** pour avoir présidé ce jury, les rapporteurs, **Wim Burmeister** et **Laurent Debarbieux**, ainsi que **François Lecointe** en tant qu'examineur.

Un grand merci aux personnes qui m'ont aidé pour effectuer les expérimentations et analyser les résultats de mes travaux :

- **Madalena Renouard**, que je remercie chaleureusement pour m'avoir aidé à mettre en place les premiers tests de ma thèse et qui a obtenu les résultats préliminaires qui ont rendu ce travail possible.
- **Jean Lepault**, **Stéphane Roche** et **Rafael Navaza** qui ont résolu la structure de la protéine A1 par cryomicroscopie électronique. Par ailleurs, j'ai fait les essais de microscopie à coloration négative avec Jean Lepault, qui m'a également donné d'importants conseils pour tirer le meilleur de mon travail et de moi-même. Je le remercie donc tout particulièrement.
- **Inès Gallay** et **Herman van Tilbeurgh** qui ont résolu la structure par Cristallographie aux Rayons X. Je remercie également **Stéphane Plancqueel**, avec qui nous avons optimisé la production de cristaux de protéine dans la bonne humeur et **Pierre Legrand** qui a réalisé le tout premier modèle de A1 pendant les expériences de cristallographie au Synchrotron.
- **Dominique Durand** pour les expériences de SAXS et la gentillesse avec laquelle elle m'a aidé à comprendre les résultats et à rédiger cette partie de ma thèse.
- **Christophe Velours** pour les essais de SEC MALLS et pour m'avoir aidé à analyser les résultats de Dichroïsme Circulaire.



- **Marie-Agnès Petit** qui a évalué mes recherches à mi-thèse, avec qui j'ai pu avoir des discussions scientifiques enrichissantes et qui m'a fourni les souches bactériennes mutantes pour leur activité recombinase.
- **Sophie Zinn-Justin** qui a effectué les premiers essais de RMN sur la protéine A2.
- **Luis Ramirez** avec qui nous avons mis au point un programme Python pour analyser des tests de toxicité.

Je remercie toutes les personnes avec qui j'ai pu avoir des discussions constructives qui m'ont aidé à orienter ce projet, essentiellement à l'Institut de Biologie Intégrative de la Cellule (I2BC) et dans le département de Virologie dont nous faisons partie.

Je remercie les personnes de notre équipe Phage T5, qui m'ont toutes aidé à divers moments de ma thèse. Il s'agit donc de Pascale Boulanger, Madalena Renouard, Luis Ramirez, mais aussi **Ombeline Rossier** avec qui j'ai eu de nombreuses discussions et qui m'a aidé à faire les meilleurs choix pour la fin de cette thèse et **Emeline Vernhes** qui était présente au début de ma thèse et grâce à qui j'ai pu prendre mes marques au laboratoire.

Je remercie **Corentin Léger** et à nouveau **Luis Ramirez**, qui étaient en thèse en même temps que moi et avec qui nous nous sommes serré les coudes.

Mes remerciements vont enfin aux personnes qui m'ont soutenu dans ce travail de longue haleine. En particulier, je remercie de tout cœur **Eloïse Gomard** et mes parents, **Emilie Zangelmi** et **Alain Zangelmi**, qui m'ont encouragé tout au long de cette thèse et qui m'ont parfois aidé à résoudre des problèmes, simplement en m'écoutant les reformuler.

# Table des matières

<b>I. Introduction .....</b>	<b>1</b>
<b>1) Les bactériophages.....</b>	<b>2</b>
1.1 Généralités.....	2
1.2 Classification .....	3
1.3 Histoire des phages.....	4
1.4 Application et utilisation des phages .....	5
1.4.1 Thérapie.....	5
1.4.1.1 Phages et évolution .....	6
1.4.2 Utilisation des phages dans l'industrie Agroalimentaire .....	8
1.4.3 Biologie moléculaire .....	8
1.4.4 Biotechnologie.....	9
1.5 Cycle lytique et cycle lysogénique.....	10
1.6 Prise de contrôle des machineries de biosynthèse par les phages :.....	12
<b>2) Le bactériophage T5 .....</b>	<b>13</b>
2.1 Généralités.....	13
2.2 Le mécanisme original d'infection de T5 .....	14
2.2.1 L'adsorption.....	14
2.2.2 Transfert de l'ADN de T5 : mécanisme en deux étapes .....	14
2.2.2.1 Principe .....	14
2.2.2.2 Mécanisme d'injection.....	15
2.2.3 Etapes précoces de l'infection.....	16
2.2.4 La réplication et l'assemblage de particules virales .....	17
2.2.5 La lyse cellulaire.....	18
2.3 Les protéines précoces du phage T5.....	19
2.3.1 Carte génétique.....	19
2.3.2 La protéine A1 .....	20
2.3.3 La protéine A2 .....	21
2.3.4 Autres protéines précoces.....	22
<b>3) Sujet de thèse : Etude des protéines A1 et A2 du bactériophage T5.....</b>	<b>23</b>
3.1 Travaux préliminaires sur les protéines A1 et A2 .....	23
3.1.1 Etude bioinformatique .....	23
3.1.2 Travaux préliminaires sur la protéine A1 .....	24
3.1.2.1 L'activité DNase de la protéine A1.....	24
3.1.2.2 Etudes structurales .....	26
3.1.3 Travaux préliminaires sur la protéine A2 .....	26

3.2	Objectifs de la thèse.....	27
<b>II.</b>	<b>Résultats.....</b>	<b>29</b>
<b>1)</b>	<b>Caractérisation de la protéine virale A1 .....</b>	<b>30</b>
1.1	Études structurales .....	30
1.1.1	Mise au point des conditions de production et purification de A1.....	30
1.1.1.1	Toxicité de la production de A1 dans la bactérie.....	30
1.1.1.2	Optimisation des conditions de purification.....	34
1.1.2	Analyse de l'état oligomérique de A1 par SEC-MALLS .....	38
1.1.3	Etude structurale de A1 en solution par SAXS.....	40
1.1.4	Analyse de la structure secondaire de A1 par Dichroïsme Circulaire.....	45
1.1.5	Cristallographie aux rayons X .....	47
1.1.6	Cryo-microscopie électronique de la protéine A1.....	49
1.1.6.1	Structure du monomère .....	51
1.1.6.2	Domaine 402-455.....	52
1.1.6.3	Domaine 89-295 et Domaine actif (296-401 456-556) .....	52
1.1.6.4	Analyse des domaines d'interaction entre les monomères de A1 .....	60
1.1.6.5	Analyse de la structure et des propriétés de l'octamère.....	64
1.2	Etude de l'activité DNase de la protéine A1 purifiée .....	69
1.2.1	Dégradation d'oligonucléotides de 75 bases .....	69
1.2.1.1	ADN fluorescents simples brins et doubles brins.....	69
1.2.1.2	Comparaison de l'activité de A1 avec différentes nucléases .....	71
1.2.1.3	Absence d'activité DNase chez les mutants du site actif .....	73
1.2.1.4	Caractérisation des conditions optimales pour observer l'activité de A1 .....	74
1.2.1.5	Comparaison de l'activité de différents lots de A1 sur des oligonucléotides marqués ..	75
1.2.2	Tests à différentes températures .....	76
1.2.3	Tests d'activité endonucléase avec des oligonucléotides portant des liaisons phosphorothioate aux deux extrémités .....	77
1.2.3.1	Blocage des extrémités par 3 liaisons phosphorothioate .....	77
1.2.3.2	PTO 25-25-25 .....	78
1.2.4	Détermination du sens de la digestion avec les oligonucléotides PTO 15-25-35 et PTO 35-25-15 .....	79
1.2.5	Comparaison avec d'autres enzymes connues sur PTO 25-50 et 50-25.....	81
1.2.6	Activité sur des ADN partiellement hybridés .....	82
1.2.7	Tests sur des ADN double-brin circulaires et sur un ADN chromosomique .....	84
1.2.7.1	Test d'activité sur des ADN double-brin circulaires .....	84
1.2.8	Test d'activité sur un chromosome extrait de E. coli .....	86
1.2.9	Mise en évidence d'activités nickase et endonucléase de A1 sans cation divalent .....	86
1.2.10	Discussion sur l'activité de A1 in vivo .....	87

1.2.11	Altération de l'activité de A1 en préincubation avec du MnCl <sub>2</sub> .....	88
1.3	Microscopie électronique à coloration négative .....	89
1.3.1	Oligomérisation de A1 en présence de MnCl <sub>2</sub> .....	89
1.3.1.1	Variation de la concentration en A1 et en MnCl <sub>2</sub> .....	89
1.3.1.2	Dissociation de l'octamère en EDTA 10 mM.....	91
1.3.2	Incubations avec l'ADN.....	92
1.3.2.1	Comparaison de la vitesse d'oligomérisation avec ou sans l'ADN .....	92
1.3.2.2	Taille minimale d'ADN permettant l'oligomérisation .....	93
1.3.2.3	Observation de A1 pendant la dégradation de l'ADN.....	94
1.3.2.4	Incubation avec un ADN long.....	96
1.3.3	Comportement des mutants pour l'activité DNase.....	97
1.3.4	L'oligomérisation in vivo.....	100
1.4	Rôle de l'activité DNase de A1 dans l'infection : Tests de complémentation avec A1 .....	101
1.4.1	Dichroïsme circulaire .....	101
1.4.2	Test de complémentation .....	102
1.4.3	Mise en évidence d'une activité recombinaise importante .....	104
1.4.4	Recombinaison chez des mutants recA recT .....	106
1.4.5	Confirmation de l'activité recombinaise par séquençage : .....	107
<b>2)</b>	<b>Caractérisation de la protéine virale A2 .....</b>	<b>109</b>
2.1	Caractérisation fonctionnelle de A2 .....	109
2.1.1	Analyses in silico .....	109
2.1.1.1	HHPred .....	109
2.1.1.2	Prédiction de structure sous Phyre.....	110
2.1.2	Dichroïsme Circulaire .....	110
2.1.3	Analyses de SAXS.....	112
2.1.4	Essais de cristallogénèse .....	114
2.1.5	Étude structurale de A2 par RMN : spectre HSQC.....	114
<b>III.</b>	<b>Discussion et perspectives .....</b>	<b>117</b>
	<b>Discussion .....</b>	<b>118</b>
	Structure originale de la protéine A1. ....	118
	Spécificité de l'activité DNase de la protéine A1.....	120
	Assemblage et dissociation de l'octamère .....	123
	A1, une possible DNase-Recombinaise ?.....	124
	Caractérisation biochimique de la protéine A2 vers son étude structurale.....	125
	<b>Perspectives.....</b>	<b>127</b>
	Poursuite de l'étude structurale de la protéine A1 .....	127
	Exploration de l'activité DNase de A1 .....	127
	Complémentation fonctionnelle et recombinaison .....	128

Caractérisation fonctionnelle et structurale de A2 .....	129
<b>Conclusion .....</b>	<b>130</b>
<b>IV. Matériel et Méthodes.....</b>	<b>131</b>
<b>1) Bactéries et phages utilisés.....</b>	<b>132</b>
1.1 Souches bactériennes d' <i>Escherichia coli</i> utilisées.....	132
1.2 Phages utilisés.....	132
1.3 Mutations ambre et souches suppressives.....	132
<b>2) Clonages et mutagenèses .....</b>	<b>133</b>
2.1 Plasmides utilisés .....	133
2.2 Clonages conventionnels avec enzymes de restriction.....	133
2.3 Mutagenèse dirigée .....	134
<b>3) Production et purification des protéines A1 et A2 .....</b>	<b>134</b>
3.1 Production.....	134
3.2 Purification.....	135
3.2.1 Extraction des protéines.....	135
3.2.2 Purification par chromatographie d'affinité sur colonne de métal immobilisé (His trap) .....	135
3.2.3 Deuxième étape de purification de la protéine A1 par chromatographie d'affinité (colonne Héparine) .....	136
3.2.4 Purification par chromatographie d'échange de cations de la protéine A2.....	136
3.2.5 Chromatographie d'exclusion par la taille.....	136
3.2.6 Vérification des protéines obtenues par migration sur gel d'acrylamide .....	137
<b>4) Tests de toxicité .....</b>	<b>137</b>
<b>5) SEC-MALLS .....</b>	<b>138</b>
<b>6) Diffusion des rayons X aux petits angles (SAXS).....</b>	<b>138</b>
<b>7) Dichroïsme Circulaire (CD) .....</b>	<b>139</b>
<b>8) Microscopie électronique à coloration négative .....</b>	<b>139</b>
<b>9) Cryo-microscopie électronique .....</b>	<b>140</b>
<b>10) Cristallographie aux rayons X.....</b>	<b>140</b>
<b>11) Tests d'activité DNase .....</b>	<b>142</b>
11.1 ADN utilisés .....	142
11.2 Test du tampon de charge .....	144
11.3 Test d'activité.....	145
11.4 Migration sur gel et révélation.....	146
<b>12) Tests de complémentation .....</b>	<b>146</b>
<b>13) Résonance magnétique nucléaire (RMN).....</b>	<b>147</b>
<b>V. Références .....</b>	<b>149</b>

---

# I. Introduction

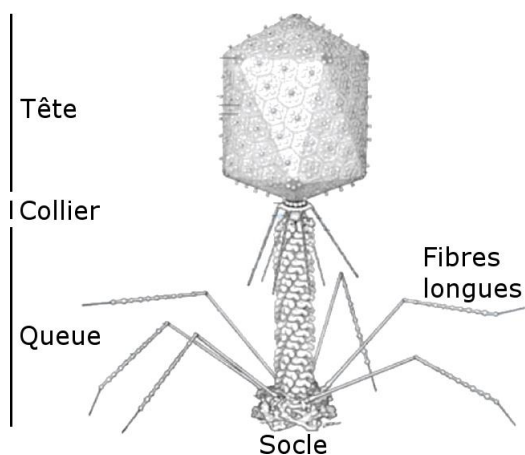
---

# 1) Les bactériophages

## 1.1 Généralités

Les bactériophages, ou phages, sont les virus qui infectent les bactéries. Ils sont présents dans chaque niche colonisée par des bactéries et ce sont les entités biologiques les plus abondantes sur Terre, avec  $10^{31}$  particules virales estimées <sup>1</sup>. Avec l'avènement d'outils de séquençage et d'outils bioinformatiques de plus en plus puissants et les études de séquençage à grande échelle d'environnements marins, les mesures sur la composition des écosystèmes microbiens sont de plus en plus précises <sup>2,3</sup>. Ces résultats montrent que les bactériophages sont responsables du renouvellement des populations bactériennes dans l'environnement. En effet, plus de 90 % de la biomasse dans la mer est constituée de microorganismes et les virus provoquent la lyse de 20 % de cette biomasse par jour, ce qui libère des composés organiques qui sont dissous dans l'eau et réutilisés par les producteurs primaires. Les virus altèrent donc les cycles géochimiques et structurent les populations et les communautés grâce à leurs propriétés lytiques <sup>4</sup>. Les virus jouent également un rôle majeur dans le transfert horizontal de gènes en emportant parfois par erreur des fragments de l'ADN de l'hôte par les mécanismes de transduction.

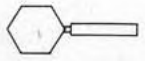
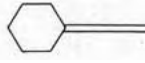
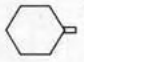







Plus de 96 % des phages connus possèdent un génome à ADN double-brin pouvant mesurer jusqu'à 500000 paires de bases (pb) et sont caudés <sup>5,6</sup>. Ils sont composés d'une tête généralement icosaédrique contenant le génome viral et d'une queue impliquée dans le transfert de l'ADN dans la bactérie hôte. Les phages sont des outils très utilisés en biologie moléculaire : par exemple, les phages  $\lambda$  et M13 sont utilisés comme vecteurs de clonage et de transfert de gènes. Cette capacité des phages à transférer des gènes fait d'eux des moteurs essentiels de l'évolution bactérienne <sup>7</sup>. Du fait de leurs propriétés lytiques, ils peuvent être utilisés en thérapie et constituer une solution alternative ou complémentaire aux antibiotiques.



**Figure 1. Structure typique d'un bactériophage caudé, le bactériophage T4.** Adapté de M. L. Yap, 2014 <sup>8</sup>.

## 1.2 Classification

La classification des phages permet la généralisation et la simplification des observations en étudiant les phages en tant que groupes. Il y a des centaines de milliers voire des millions d'espèces et la grande majorité des séquences connues n'ont pas d'homologie dans les bases de données. La classification des virus pose certains problèmes car les phages échangent des gènes en permanence et cela rend difficile l'établissement de groupes. On ne peut donc pas les classer à partir d'un seul critère. Les phages présentent des morphologies très diverses <sup>6</sup>. De plus en plus de génomes séquencés sont disponibles, ce qui permet de faire de la taxonomie basée sur ces génomes. Cependant, de nombreuses données sont encore manquantes et les systèmes de classification subissent actuellement une évolution très rapide <sup>9</sup>.

Forme	Ordre ou Famille	Acide nucléique, particularités, taille	Membre
	<i>Caudovirales</i>	ADN double-brin linéaire. Pas d'enveloppe	
	<i>Myoviridae</i>	Queue contractile	T4
	<i>Siphoviridae</i>	Queue longue, non contractile	T5
	<i>Podoviridae</i>	Queue courte	T7
	<i>Microviridae</i>	ADN simple-brin circulaire, 27 nm, 12 capsomères en forme de bouton	Φ X174
	<i>Corticoviridae</i>	ADN double-brin circulaire, capsid complexe, lipides, 63 nm	PM2
	<i>Tectiviridae</i>	ADN double-brin linéaire, vésicule lipidique interne, pseudo-queue, 60 nm	PRD1
	<i>Leviviridae</i>	ARN simple-brin linéaire, 23 nm, ressemble à poliovirus	MS2
	<i>Cystoviridae</i>	ARN double-brin linéaire, segmenté, enveloppe lipidique, 70-80 nm	f 6
	<i>Inoviridae</i>	ADN simple-brin circulaire, filaments ou bâtons, 85-1950 x 7 nm	fd
	<i>Plasmaviridae</i>	ADN double-brin circulaire, enveloppe lipidique, pas de capsid, 80 nm	MVL2

**Tableau 1. Récapitulatif des familles de bactériophages <sup>6</sup>. Les *Caudovirales* sont un ordre qui contient plusieurs familles, indiquées par des bordures plus épaisses.**



### 1.3 Histoire des phages

Les premières observations indirectes de phages ont été faites par Ernest Hanbury Hankin en 1896, qui a mis en évidence un principe biologique capable de détruire des cultures de bactéries responsables du choléra dans les eaux du Gange. Les phages ont été découverts indépendamment par Frederick Twort en 1915 et par Félix d'Hérelle en 1917. En 1915, Frederick Twort identifie l'apparition de colonies transparentes sur des cultures de *Micrococcus*, dont le phénotype se transmet à d'autres colonies et conduit à leur mort. Deux ans plus tard, Félix d'Hérelle isole un microbe invisible capable de provoquer la lyse de bactéries *Shigella*. Il le définit alors comme un "bactériophage obligatoire" car ce microbe semble manger les bactéries.

**Figure 2. Félix d'Hérelle (1873-1949), initiateur de la recherche sur les bactériophages à l'institut Pasteur.**

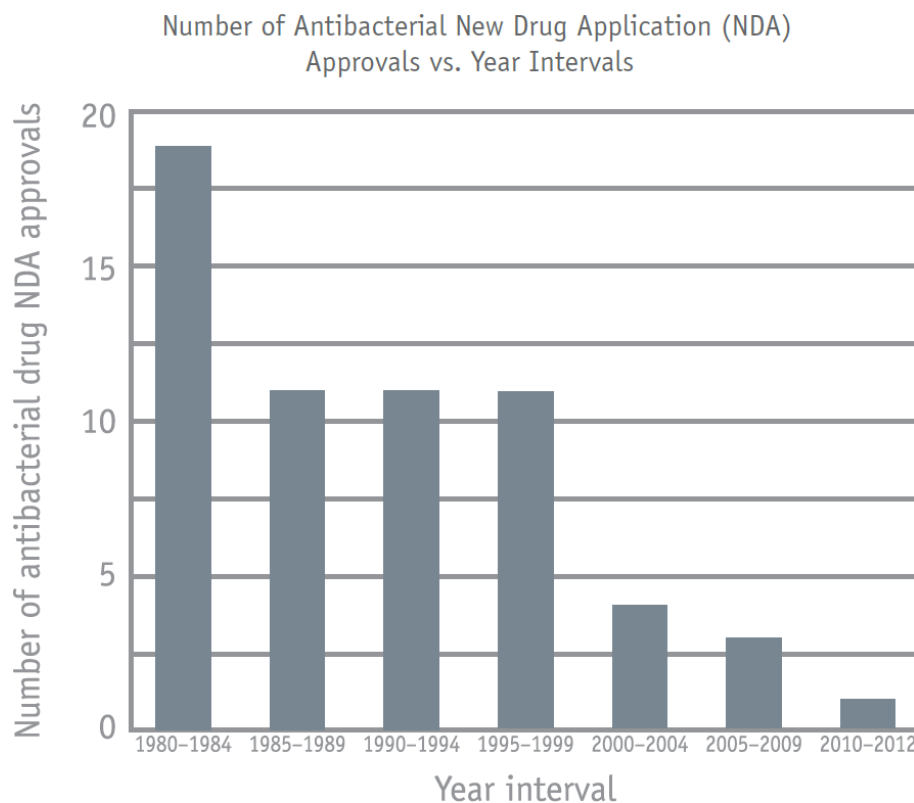


La nature virale des phages fut longtemps discutée jusqu'à leur observation en microscopie électronique dans les années 1940 par Helmut Ruska <sup>10</sup>. La morphologie de 4 types de bactériophages fut étudiée plus en détail deux ans plus tard. Les expériences ont confirmé que les particules sphériques munies d'une queue fine étaient des bactériophages, qui s'adsorbent sur des bactéries sensibles après seulement quelques minutes d'incubation <sup>11, 12</sup>.

## 1.4 Application et utilisation des phages

### 1.4.1 Thérapie

La résistance aux antibiotiques et le déficit de nouveaux traitements anti-infectieux constituant aujourd'hui une grave menace pesant sur la santé mondiale et la sécurité alimentaire <sup>13</sup>, l'utilisation des phages en thérapie et en agroalimentaire éveille de plus en plus d'intérêt. En effet, d'après une étude de 2013 du Centers for Disease Control and Prevention (CDC), il y a 2 millions d'infections multi-résistantes qui causent 23000 morts par an aux Etats-Unis <sup>14</sup> alors que de moins en moins de nouveaux antibiotiques sont développés (**Figure 3**). Pourtant, le désintérêt des industries pharmaceutiques pour le développement de nouveaux antibiotiques continue d'augmenter à cause de perspectives économiques insuffisantes pour les industriels <sup>15</sup>. Trouver de nouvelles alternatives aux antibiotiques devient donc vital.



**Figure 3. Diminution au cours du temps du nombre de nouveaux antibiotiques acceptés sur le marché <sup>14</sup>.** Les intervalles de 1980 à 2009 sont de 5 ans. L'intervalle de 2010 à 2012 est de 3 ans.

Les phages furent utilisés comme agents antibactériens dès 1917. Chaque type de bactérie peut être ciblé par au moins un bactériophage, y compris les bactéries pathogènes multi-résistantes aux antibiotiques. Les phages ont un spectre d'hôte étroit, leur utilisation n'est donc pas néfaste pour le microbiome existant. En revanche, la grande spécificité des bactériophages empêche

leur utilisation contre des bactéries contre lesquelles ils ne sont pas dirigés. Les développements actuels en biologie de synthèse permettent de contourner ce problème en développant des phages hybrides possédant des caractéristiques précises : un phage particulièrement infectieux peut être modifié avec les fibres de queue d'un autre phage assurant la spécificité de reconnaissance d'un hôte <sup>16,17</sup>. Les phages ne peuvent pas infecter les cellules eucaryotes, ce qui diminue les risques d'effets secondaires et ils sont particulièrement adaptés pour l'élimination des biofilms composés d'un nombre restreint d'espèces bactériennes <sup>18</sup>. La thérapie par les phages est complémentaire à l'utilisation des antibiotiques, avec lesquels une synergie d'action peut être observée <sup>18,19</sup>.

L'utilisation des phages n'est cependant pas encore très répandue pour de multiples raisons. A la fin des années 1930, le conseil sur la pharmacie et la chimie de l'association médicale Américaine a conclu que l'efficacité de la thérapie par les phages était ambiguë et que plus de recherche était nécessaire <sup>20</sup>. Avec l'essor des antibiotiques et la seconde guerre mondiale, cette déclaration a conduit à un déclin dans l'intérêt de la thérapie par les phages. Des essais de traitements ont continué dans les pays de l'Est, mais les essais comprennent rarement des groupes contrôles et ne répondent pas aux standards actuels de la recherche clinique. Les résultats de traitements par phagothérapie sur les animaux n'apparaissent à nouveau dans la littérature en anglais que dans les années 1980. Des articles décrivent l'utilisation de bactériophages ciblant des souches de *Escherichia coli* toxiques pour l'homme sur des souris <sup>21</sup>, des veaux, des porcelets et des agneaux <sup>22</sup>. Dans ces expériences, ces traitements ont été plus efficaces que les antibiotiques pour soigner les infections. Les expériences pour soigner les humains ont recommencé dans les années 2000, avec le premier essai contrôlé aux Etats-Unis publié en 2009 <sup>23</sup>.

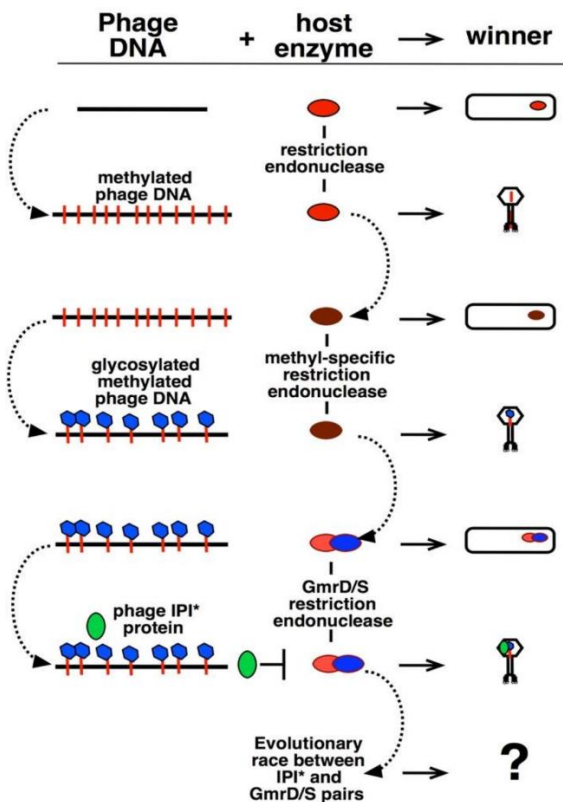
L'utilisation des phages en thérapie a certains désavantages : une utilisation responsable des phages est nécessaire pour éviter que les bactéries deviennent résistantes, notamment en utilisant des cocktails de phages renouvelés régulièrement <sup>24</sup>. Les échanges génétiques sont particulièrement courants chez les phages dont l'ADN peut s'intégrer au sein du chromosome bactérien, il faut donc exclure les phages tempérés et utiliser des virus à cycle strictement lytique qui n'ont pas cette capacité.

#### 1.4.1.1 Phages et évolution

Il est important de contrôler l'utilisation des phages pour ne pas accélérer l'évolution des bactéries pathogènes. En effet, les phages sont les moteurs de l'évolution bactérienne à travers

de nombreux mécanismes <sup>7,25</sup>. En provoquant la lyse des bactéries, ils libèrent de l'ADN que les bactéries compétentes peuvent acquérir et intégrer dans leur génome par recombinaison homologue. La sélection d'hôtes mutants résistants à l'infection par les phages peut conduire à des réarrangements génomiques importants dans la bactérie <sup>26</sup>. La transduction généralisée ou spécialisée est responsable du transport de fragments d'ADN entre les bactéries. Les bactéries ont détourné cette fonction : certains ADN phage-like non répliatifs font de la transduction généralisée de manière constitutive. Les phages tempérés peuvent aussi participer à l'évolution bactérienne par la conversion lysogénique : l'apport de gènes bénéfiques pour l'hôte. Des prophages génétiquement proches peuvent être présents sur le même génome et déclencher des événements de recombinaison conduisant à des inversions ou des délétions <sup>27</sup>. Enfin, les bactéries peuvent utiliser des éléments de phage tels que les pyocines ou le système de sécrétion de type VI pour interagir avec des organismes eucaryotes.

Les phages et les bactéries co-évoluent à travers un mécanisme appelé la "course à l'armement", où un processus de sélection constant conduit au développement de défenses contre l'infection ou au contournement de ces défenses (**Figure 4**).



**Figure 4. Exemple de "course à l'armement" chez le phage T4.** Adapté de Nechaev et Severinov (2008) <sup>25</sup>. Chaque interaction génétique (infection) est montrée au même niveau horizontal, avec le résultat de chaque interaction listée à droite. L'hôte gagne (winner) si le système de restriction a pu dégrader l'ADN du phage et le phage gagne s'il est résistant au système de restriction de l'hôte. IPI\* : internal protein I mature qui contre spécifiquement le système de défense GmrS/GmrD.

En plus de l'apport de gènes bénéfiques par transduction, les phages peuvent altérer la régulation des gènes de l'hôte pendant l'infection pour lui conférer un meilleur fitness, ce qui permet indirectement un meilleur développement des phages <sup>28</sup>.

#### 1.4.2 Utilisation des phages dans l'industrie Agroalimentaire

L'utilisation de plusieurs souches de phages dans l'agroalimentaire a été autorisée par la FDA (Food and Drug Administration) aux Etats-Unis en 2006 <sup>29-32</sup>. Des modèles mathématiques permettent de prévoir l'efficacité de l'utilisation des phages en fonction de facteurs environnementaux tels que la température ou le pH et leur précision est encore en train d'être améliorée <sup>33</sup>. Par ailleurs, des gènes d'endolysines phagiques ont été clonés dans des vecteurs d'expression, transfectés et exprimés dans des plantes pour les protéger des infections. Ces expériences ont obtenu un franc succès, sans toxicité pour les plantes <sup>34</sup>.

#### 1.4.3 Biologie moléculaire

La découverte des systèmes de restriction/modification à travers l'observation d'une diminution de la progénie virale après une lyse par les phages sur certaines souches bactériennes <sup>35,36</sup> a participé au développement de la biologie moléculaire. Ils ont permis le développement des techniques de clonage de gènes, grâce aux coupures dans l'ADN <sup>37</sup> et à la liaison avec l'ADN ligase de T4 <sup>38</sup>. Par ailleurs, l'étude des phages a augmenté les possibilités de ces systèmes : des études sur l'encapsulation de l'ADN du phage lambda ont conduit au développement des cosmides qui sont utilisés pour délivrer par transduction de larges portions d'ADN dans les bactéries <sup>39</sup>.

Des réponses à des questions centrales en biologie ont été obtenues grâce aux phages. Des études ont démontré que les mutations préexistent en absence de sélection car il existe une variation aléatoire du nombre de bactéries survivantes à l'infection par un phage dans des cultures indépendantes <sup>40</sup>. Pour montrer que l'ADN est le support de l'information génétique, des phages avec de l'ADN et des protéines marqués radioactivement ont été utilisés. Seul le marquage de l'ADN a été retrouvé dans les bactéries et a provoqué la production de nouveaux virions <sup>41</sup>. Par la suite, la nature en triplet du code génétique et les promoteurs ont été découverts grâce à des tests de recombinaison chez des phages T4 mutants <sup>42</sup>. Des études sur le phage lambda ont révélé l'existence d'activateurs et de répresseurs de la transcription se liant à l'ADN et des anti-terminateurs transcriptionnels <sup>43</sup>.

Récemment, nous assistons au développement des outils basés sur le système CRISPR-Cas, qui est responsable de l'immunité adaptative des bactéries et des archées <sup>44,45</sup>. Ce système est composé de répétitions courtes d'ADN séparées par des séquences (les "spacers") acquises par la bactérie lors d'infections précédentes auxquelles elle a survécu. Les protéines Cas se servent

des "spacers" pour reconnaître et détruire l'ADN indésirable par un clivage endonucléolytique. Cas9 est une protéine de ce système utilisant un ARN traqueur pour reconnaître une région sur laquelle faire une coupure double-brin. En utilisant un ARN spécifique, elle permet d'éditer le génome avec précision chez de nombreux organismes. Par ailleurs, une protéine Cas9 déficiente pour l'activité nucléase peut être utilisée pour reconnaître une région opératrice de l'ADN et agir comme un répresseur ou un activateur de la transcription <sup>46</sup>.

#### 1.4.4 *Biotechnologie*

Les phages ont également contribué à de nombreux développements en biotechnologie. Les outils de séquençage ont été mis au point à l'aide de la T7 DNA polymérase <sup>47</sup> et les génomes de phage ont été les premiers génomes d'ARN, d'ADN simple brin et d'ADN double brin à avoir été séquencés <sup>48-50</sup>.

Une autre technique plus récente est la technique de "phage display" découverte par George Smith. Elle consiste en l'amplification de phages filamenteux présentant une protéine particulière capable d'interagir avec une cible d'intérêt <sup>51</sup>. Cette technique a des applications dans la découverte de médicaments et la science environnementale <sup>52</sup>. Elle a été développée afin de sélectionner des protéines se liant à une cible spécifique par criblage. La protéine d'enveloppe des phages est associée à la protéine à tester. Les phages qui se lient à la cible sont alors isolés, amplifiés et séquencés. Cette méthode est plus simple, moins chère et plus rapide que les méthodes de séquençage de protéine comme la spectrométrie de masse <sup>53,54</sup>. Cette technique a eu un tel impact dans le développement de nouveaux traitements médicaux que les chercheurs l'ayant découverte et mise au point font partie des colauréats du prix Nobel de chimie 2018 récompensant les méthodes d'évolution dirigée en biochimie. George Smith a reçu ce prix pour avoir développé la méthode de phage display et Gregory Winter pour avoir utilisé le phage display pour provoquer l'évolution dirigée d'anticorps dans le but de développer de nouveaux traitements.

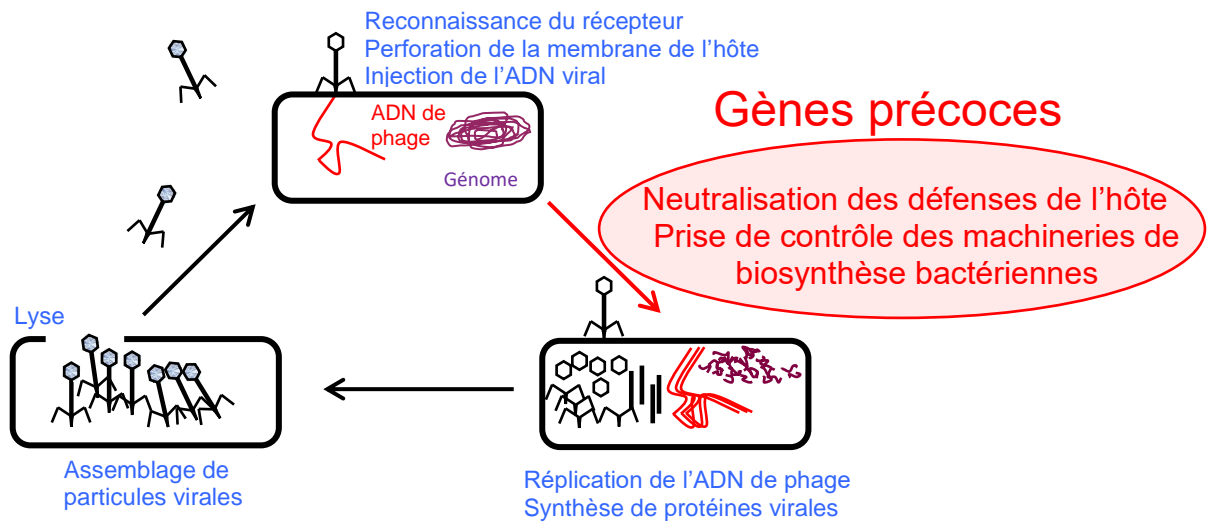
Les phages ont aussi eu un rôle crucial dans le développement de la biologie de synthèse. Le génome du phage phiX174 a été le premier à être synthétisé artificiellement <sup>55</sup>. Les intégrases et les recombinases des phages sont utilisées en biotechnologie pour contrôler les événements d'intégration, d'excision et d'inversion <sup>56</sup>. Il est possible de muter des gènes de manière conditionnelle grâce au système Cre-loxP de recombinaison à site spécifique du phage P1. Ces altérations génétiques peuvent être contrôlées de manière à affecter un gène dans un tissu spécifique ou à un stade précis du développement, dans des génomes eucaryotes et procaryotes <sup>57</sup>. Ces applications s'étendent dans d'autres domaines comme l'informatique : l'ARN polymérase de T7 peut être utilisée pour miniaturiser des circuits synthétiques <sup>58</sup>.

## 1.5 Cycle lytique et cycle lysogénique

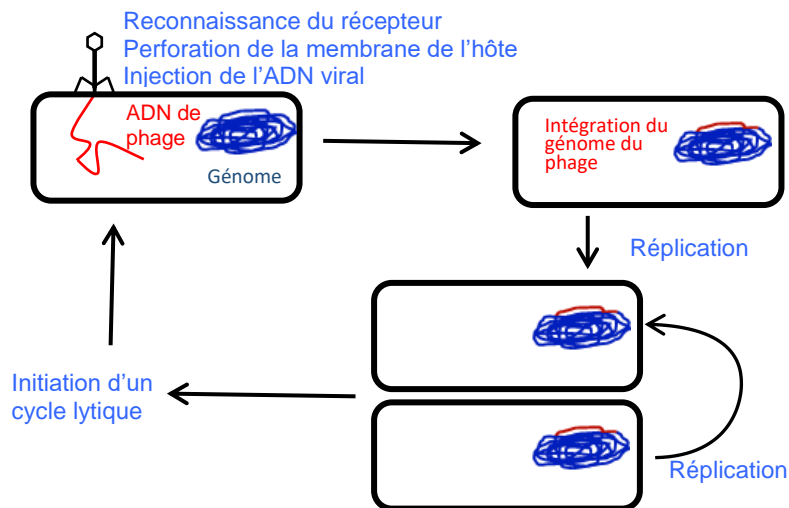
Les études d'interaction phage/hôte ont révélé deux stratégies distinctes pour le cycle phagique :

- Les phages lytiques produisent de nouvelles particules virales qui sont relâchées dans le milieu extracellulaire par la lyse de l'hôte ou par bourgeonnement dans le cas des phages filamenteux, ce qui permet une nouvelle transmission horizontale<sup>59</sup>. Les phages doivent prendre le contrôle des machineries de biosynthèse de leur hôte et de ses ressources au profit de la réplication de leur génome, de la synthèse de leurs protéines et de l'assemblage de nouvelles particules virales. Cette prise de contrôle s'effectue au début du cycle infectieux grâce aux gènes précoces (**Figure 5**).
- Les phages tempérés peuvent déclencher un cycle lytique ou lysogéniser l'hôte. C'est-à-dire que le phage devient un épisome : son génome s'intègre au génome de l'hôte pour persister en tant que prophage dormant. Il sera alors répliqué de manière stable pour assurer sa transmission verticale. Dans cet état, les gènes impliqués dans le développement lytique sont réprimés par un répresseur du phage souvent associé à un mécanisme d'interrupteur (**Figure 6**).

Un troisième état plus atypique a été mis en évidence grâce à des études de cellules isolées en microscopie. Dans cet état, une cellule portant un phage donne naissance à des cellules sans phage temporairement immunisées contre l'infection. Ce mécanisme pourrait permettre aux phages d'éviter de détruire la population de l'hôte pour pouvoir continuer à se multiplier<sup>59</sup>.



**Figure 5: Cycle lytique des bactériophages.**



**Figure 6. Cycle lysogénique des bactériophages.** L'initiation du cycle lytique a lieu lorsque le répresseur du phage n'est plus exprimé, par exemple lorsque la bactérie est en condition de stress.



## 1.6 Prise de contrôle des machineries de biosynthèse par les phages :

Dès le démarrage du cycle lytique qui se produit immédiatement après l'injection du génome viral chez les phages lytiques ou après l'excision du génome viral chez les phages tempérés, les gènes précoces sont exprimés pour permettre la prise de contrôle de l'hôte. La plupart des phages à ADN double-brin codent pour les éléments de leur propre machinerie de réplication et ils détournent rarement les enzymes de réplication de l'hôte. En revanche, l'ARN polymérase est l'enzyme la plus régulée de la cellule et elle est très couramment ciblée par les phages pendant l'infection pour permettre la transcription de leurs gènes <sup>25</sup>. De nombreux phages comme T5 codent pour leurs propres ARN de transfert (ARNt). Le séquençage de plusieurs centaines de génomes de phages a permis des études bioinformatiques : les ARNt conservés dans de nombreux génomes de phages correspondent aux codons abondants chez ces phages et rares chez l'hôte, ce qui permet une traduction plus rapide et efficace des protéines de phage <sup>60</sup>. Ce n'est cependant pas le cas pour tous les phages : T4 et plusieurs phages T4-like utilisent des ARNt qui reconnaissent des codons riches en GC, qui sont pourtant rares dans des génomes de phages T4-like riches en AT <sup>61</sup>.

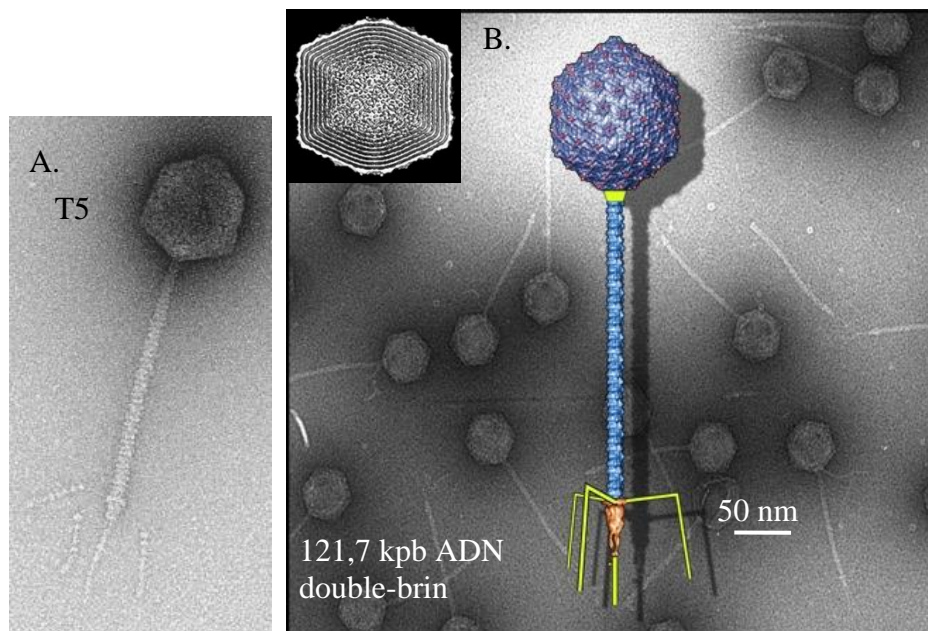
L'exploration des mécanismes moléculaires impliqués dans la prise de contrôle des fonctions cellulaires au profit du phage par les protéines précoces est cruciale pour identifier des cibles bactériennes contre lesquelles développer de nouveaux agents antibactériens. Bien que de nombreux phages soient régulièrement séquencés, de nombreux gènes inconnus ou dont la fonction est putative ne sont pas caractérisés. Les protéines codées par ces gènes présentent souvent peu d'homologies avec les protéines connues. Même chez des bactériophages très étudiés comme les phages T4 ou T5, presque la moitié des gènes prédits n'ont pas de fonction assignée <sup>61-63</sup>. Étant donné le faible taux d'homologie entre les séquences de phages, de nombreuses fonctions antibactériennes inconnues pourraient être mises à jour : les protéines précoces des bactériophages responsables de la prise de contrôle de l'hôte constituent donc un réservoir de fonctions antibactériennes inexplorées.

Cette faible homologie de séquence constitue un obstacle à l'identification de la fonction des gènes précoces. De plus, les gènes précoces sont souvent répartis sans ordre apparent sur le génome du phage : leur identification a été possible dans quelques rares cas grâce à des études de transcriptomique qui ont identifié les différents gènes de phages exprimés au cours des différentes phases du cycle infectieux <sup>64,65</sup>. Nous allons voir comment le mécanisme d'infection original du bactériophage T5 en fait un modèle pour l'étude des gènes précoces.

## 2) Le bactériophage T5

### 2.1 Généralités

Le bactériophage T5 est étudié depuis les années 1940. C'est un phage strictement lytique infectant *Escherichia coli*, une souche commensale de la flore intestinale dont certaines sous-espèces sont connues pour causer des infections multi-résistantes aux antibiotiques <sup>66</sup>. Il appartient à l'ordre des *Caudovirales* et à la famille des *Siphoviridae* (**Tableau 1**). Il possède une longue queue flexible et non-contractile attachée à une capsidie icosaédrique contenant un ADN double-brin de 121,7 kpb (**Figure 7**). L'injection du génome dans l'hôte suit un mécanisme unique en deux étapes.



**Figure 7. Particule virale du bactériophage T5.** **A.** Phage T5 vu en microscopie électronique à coloration négative. **B.** Modèle structural de la particule virale avec reconstruction à basse résolution (25 Å) de la capsidie et de la queue du phage T5, obtenue à partir de données de cryo-microscopie électronique <sup>67</sup>.

## 2.2 Le mécanisme original d'infection de T5

### 2.2.1 L'adsorption

La reconnaissance de l'hôte est initiée par la liaison réversible des longues fibres latérales de queue du phage au O-antigène du lipopolysaccharide. Cette interaction permet un rapprochement du phage avec la surface de la bactérie et conduit à la fixation irréversible de la protéine pb5, localisée à l'extrémité de la fibre centrale, au récepteur FhuA <sup>68</sup>. FhuA est une protéine de la membrane externe de *E. coli* de 78,9 kDa qui catalyse le transport du ferrichrome et de l'albomycine à travers la membrane. Elle est également le récepteur des bactériophages T1, phi80 et de la colicine M. Cette interaction induit un changement de conformation de l'extrémité de la fibre droite conduisant à la perforation de l'enveloppe bactérienne, à l'ouverture de la capsid et à la libération de l'ADN viral qui est transféré dans le cytoplasme bactérien <sup>63</sup>. De récents travaux sur la structure de la queue de T5 suggèrent que cette transmission du signal depuis la queue jusqu'à la capsid est médiée par la protéine vernier pb2, ou TMP (Tape measure protein), longue protéine localisée à l'intérieur du tube caudal et qui est éjectée après interaction de T5 avec FhuA <sup>69</sup>. FhuA est utilisée comme récepteur, mais l'ADN ne la traverse pas <sup>70</sup>.

### 2.2.2 Transfert de l'ADN de T5 : mécanisme en deux étapes

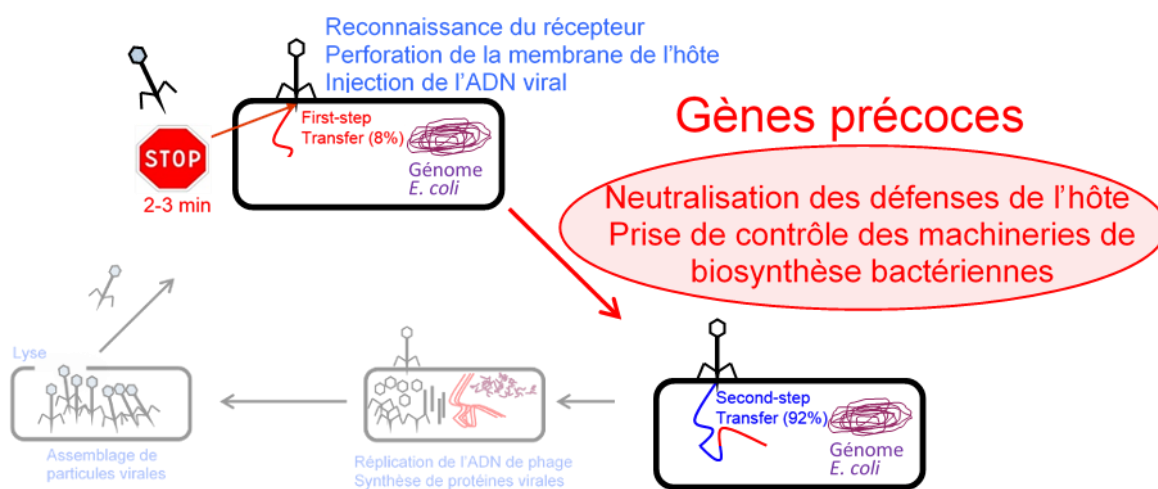
#### 2.2.2.1 Principe

Le phage T5 est le seul phage connu qui injecte son ADN en deux étapes distinctes. Les phages homologues à T5 utilisent probablement le même système d'injection de l'ADN mais seules les études sur le phage homologue BF23 l'ont prouvé <sup>71</sup>. BF23 est un phage qui n'a pas été entièrement séquencé mais qui est génétiquement très proche de T5 <sup>72</sup>. Chez T5, 8 % du génome correspondant à 10,2 kpb sont d'abord injectés dans la bactérie (FST, First Step Transfer) sans que la synthèse de protéines ne soit nécessaire <sup>73</sup>.

Environ 4 minutes après l'addition des phages T5 à une suspension de *E. coli*, la moitié de l'ADN de l'hôte est dépolymérisé et des mononucléotides sont excrétés à partir de 10 minutes d'infection. L'ADN du phage ne semble pas être dégradé <sup>74-76</sup>. L'expression des gènes précoces est essentielle pour la dégradation massive du génome de l'hôte et le transport de l'ADN restant après 3 à 5 minutes (SST, Second Step Transfer). Les gènes plus tardifs permettent la réplication du génome et l'assemblage des virions <sup>73</sup>.

Comme le phage T5 injecte son ADN en deux étapes, il est possible d'éliminer l'ADN qui n'est pas rentré à l'aide d'un "Blender" qui coupe les phages liés à l'hôte pour étudier les étapes

précoces de l'infection séparément des autres étapes. Les premiers essais sur le transfert de l'ADN de T5 avaient permis de mettre en évidence grâce à un marquage radioactif et à la découpe des phages au "Blender" qu'environ 10 % de l'ADN marqué de T5 était injecté dans les bactéries dans un premier temps <sup>77</sup>. Deux hypothèses principales ont alors émergé : soit 10 % des phages transféraient tout leur ADN, soit tous les phages transféraient 10 % de leur ADN. L'utilisation du "Blender" a permis à Lanni, McCorquodale et Wilson de trancher entre les deux hypothèses : l'ADN partiellement transféré extrait présentait un poids moléculaire correspondant à 8 % de celui de l'ADN intact. Les phages avaient donc injecté 8 % de leur génome avant d'être séparés de l'hôte. Ces expériences ont également montré que l'ADN du FST du phage T5 seul n'est pas suffisant pour permettre l'infection productive par le phage <sup>78</sup>.



**Figure 8. Cycle infectieux du bactériophage T5.**

#### 2.2.2.2 Mécanisme d'injection

La protéine de liaison au récepteur pb5 est présente en un seul exemplaire sur le phage et elle est composée d'un seul domaine compact. La liaison de pb5 à FhuA conduit à la rigidification de la structure tertiaire du complexe <sup>68</sup>. Aucun réarrangement majeur n'est observé chez les deux partenaires. L'information de cette liaison doit pourtant être transmise au reste du phage pour déclencher l'injection de l'ADN : le changement de conformation de pb5 nécessaire à la transmission de l'information doit donc être de faible amplitude <sup>79</sup>.

Lorsque le phage T5 est mis en présence de protéine FhuA purifiée ou de liposomes portant la protéine FhuA, l'injection complète du génome est déclenchée sans diminution significative de la vitesse du transfert <sup>70,80,81</sup>. Il existe donc un facteur de l'hôte indispensable au transfert en deux étapes *in vivo*. L'étude du mécanisme d'arrêt et de reprise du transfert du génome du phage T5 est essentielle pour comprendre les avantages sélectifs associés à cette stratégie d'infection qui a été conservée par ce bactériophage au cours de l'évolution.

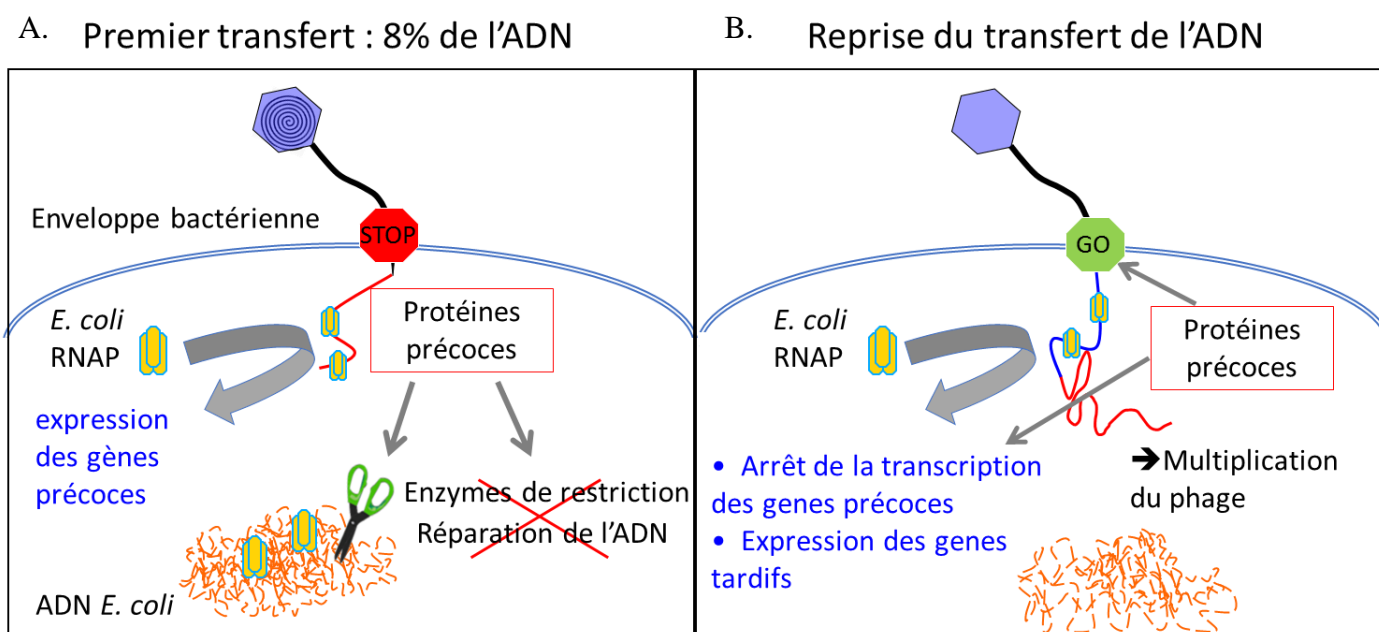
### 2.2.3 Etapes précoces de l'infection

Dès la fin de l'injection de l'ADN du FST, les gènes précoces du phage sont exprimés. L'ADN du FST ne possède aucun site de restriction EcoRI alors qu'il existe 6 sites EcoRI sur le reste de l'ADN. Pourtant, le bactériophage T5 peut se développer sur un hôte exprimant l'enzyme EcoRI, il doit donc exister un système codé par les gènes précoces pour inactiver ces systèmes de défense <sup>82</sup>. Ce système inconnu ne permet pas la protection de l'ADN du FST car des phages mutants portant un site EcoRI dans l'ADN du FST ne peuvent pas se développer sur un hôte exprimant l'enzyme EcoRI. Un exemple de système d'inhibition est la protéine de protection de restriction Ocr des phages T3 et T7 qui imite l'ADN et se fixe irréversiblement sur certaines enzymes de restriction <sup>83</sup>. Les phages utilisent souvent un ADN modifié (méthylation ou utilisation d'une base modifiée comme l'hydroxyméthylcytosine chez T4) pour ne plus être reconnu par l'enzyme <sup>84</sup>. Cependant, le phage T5 utilise un ADN non-modifié.

Des travaux comparant l'inhibition des DNases dépendantes à l'ATP de *E. coli* par des phages à ADN double-brin suggèrent que la nucléase RecB est rapidement inhibée par le bactériophage T5 <sup>85</sup>. Cette inhibition n'est pas observée en présence de Chloramphénicol qui empêche la synthèse de protéines et empêche donc le SST chez T5. Il existe donc un mécanisme associé aux gènes du FST responsable de l'inhibition de l'activité nucléase de RecB. De plus, l'ARN polymérase bactérienne est rapidement détournée lors de l'infection car T5 ne produit pas sa propre ARN polymérase et il peut se développer en présence de rifampicine (inhibiteur de l'ARN polymérase) sur un hôte dont l'ARN polymérase résiste à la rifampicine alors qu'il ne peut pas le faire sur un hôte sensible à la rifampicine <sup>86,87</sup>. Ce détournement des machineries de biosynthèse permet d'établir un environnement favorable au développement viral. Ce remaniement du métabolisme bactérien précède les étapes de réplication de l'ADN, de synthèse des protéines virales et d'assemblage des virions (**Figure 9**).

La dégradation de l'ADN de l'hôte <sup>88</sup> étant massive et rapide lors de l'infection par T5, l'expression des gènes de l'hôte est fortement inhibée en quelques minutes.

Des études sur les systèmes CRISPR-Cas ont révélé que l'infection par le bactériophage T5 a lieu sans interférence sur des bactéries portant des "spacers" ciblant l'ADN du SST. Cet ADN n'est pourtant pas résistant au clivage endonucléolytique car il est coupé lorsqu'il est directement transformé dans une bactérie possédant des "spacers" ciblant l'ADN du phage. Les systèmes CRISPR ciblant le phage T5 ne peuvent être efficaces *in vivo* que s'ils ciblent l'ADN du FST <sup>45</sup>, ce qui indique qu'il existe un mécanisme d'inhibition du système CRISPR qui protège l'ADN du SST.



**Figure 9. Etapes précoces de l'infection par le bactériophage T5. A.** Arrêt du transfert après l'injection de 8% du génome et prise de contrôle de l'hôte. **B.** Injection de l'ADN restant et transition vers les étapes de réplication du phage.

Les étapes de prise de contrôle de l'hôte par le bactériophage T5 sont encore peu étudiées et mal comprises. Il n'y a pas eu d'évolution majeure dans la compréhension de ces mécanismes depuis plus de 40 ans <sup>72,89</sup>. Pourtant, leur étude permettrait une meilleure compréhension des étapes précoces de l'infection des bactériophages dans l'objectif de développer de nouveaux agents antibactériens.

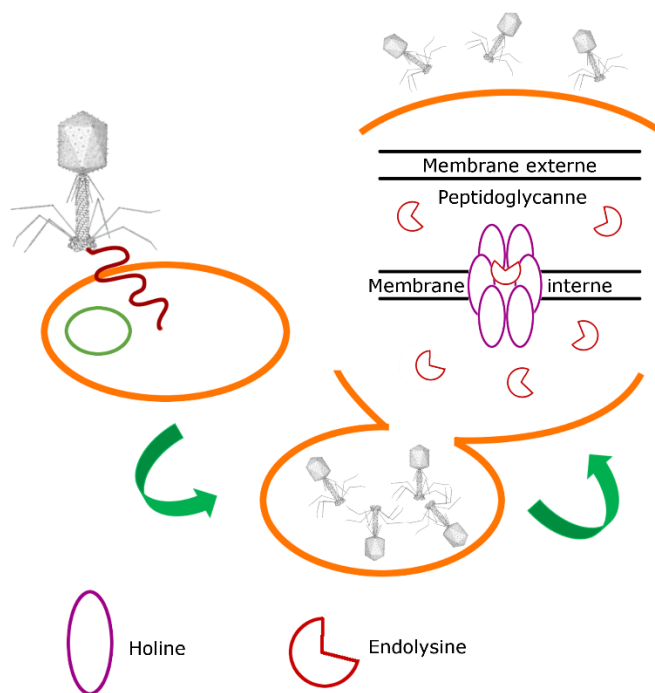
#### 2.2.4 La réplication et l'assemblage de particules virales

Le phage T5 synthétise ses nucléotides *de novo* <sup>74</sup>, contrairement à d'autres phages comme T3, T4 ou T7 qui dégradent l'ADN lentement et réincorporent les nucléosides produits dans l'ADN des nouveaux virions <sup>90,91</sup>.

Le phage T5 utilise sa propre ADN polymérase <sup>86</sup>. La réplication de l'ADN conduit à la production de concatémères contenant plusieurs copies du génome liées en série. Les 10 kpb des gènes du FST du phage T5 sont dupliqués à l'extrémité du génome. *In vitro*, l'ADN polymérase de T5 est capable de répliquer à la fois l'ADN circulaire par « rolling circle » et l'ADN linéaire en structures branchées en utilisant des interruptions simple brin comme point de départ, l'ADN en amont de l'interruption servant d'amorce <sup>92</sup>. Des molécules d'ADN de phage linéaires et circulaires ont été observées par microscopie électronique à différentes étapes du cycle de réplication du phage T5 <sup>93</sup>. Il n'est donc pas possible de distinguer entre ces deux méthodes de réplication de l'ADN.

### 2.2.5 La lyse cellulaire

Le phage T5 utilise probablement un système canonique pour provoquer la lyse bactérienne avec des protéines de la famille des holines et des endolysines. L'holine est une protéine codée par le phage qui s'accumule sans dégâts dans la membrane cytoplasmique jusqu'à ce qu'elle soit activée à un moment spécifique pour former des trous pouvant mesurer jusqu'à 1  $\mu\text{m}$  dans la membrane de l'hôte. Ceux-ci permettent à l'endolysine également produite par le phage de s'échapper du cytoplasme pour dégrader le peptidoglycane (**Figure 10**). Les holines contrôlent le moment où la lyse a lieu, de manière programmée directement dans leur structure primaire et indépendamment des étapes de réplication du phage. C'est l'accrétion rapide de ces protéines sous forme de radeaux sur la membrane qui provoque un effondrement de la membrane et la formation des trous <sup>94</sup>.



**Figure 10. Schéma de la lyse bactérienne par les holines et les endolysines.** Adapté de V. Jarábková, 2015 <sup>95</sup>.

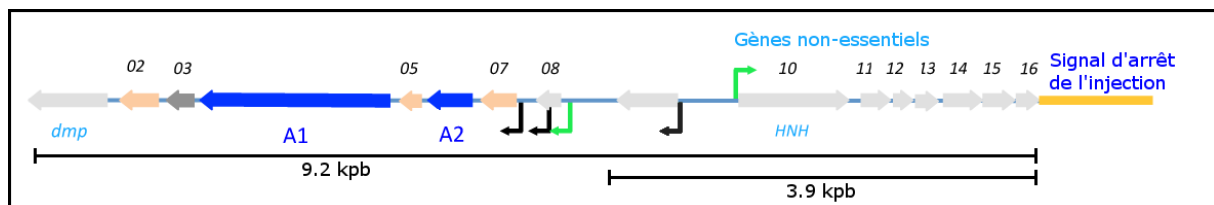
Le temps de formation des trous par les holines peut être modifié en mutant les holines ou en modifiant la polarisation de la membrane. Altérer le niveau de transcription ou de traduction du gène de l'holine a aussi un effet sur la temporisation de la lyse. Le seul système chez lequel une modification dynamique de la temporisation de la lyse a été observé est chez le bactériophage T4. Ce phage utilise une anti-holine qui peut empêcher le déclenchement de la lyse par l'holine lorsqu'il y a trop peu d'hôtes disponibles pour les virions après la progénie. L'anti-holine est inactivée et dégradée tant qu'il n'y a pas de surinfection par d'autres phages T4, signe d'une pénurie d'hôtes <sup>96,97</sup>. Le prolongement du cycle viral permet de produire plus de virions avant la lyse.

## 2.3 Les protéines précoces du phage T5

### 2.3.1 Carte génétique

Le séquençage du phage T5 indique qu'il a un génome de 121752 pb, avec deux répétitions terminales directes de 10139 pb. Il possède 168 cadres ouverts de lecture pour lesquels 70 gènes ont une fonction prédite. 24 gènes codants pour des ARNt peuvent être délétés sans conséquence pour la réplication du phage et on trouve 47 promoteurs et 27 terminateurs transcriptionnels Rho-indépendants. T5 possède également certains des promoteurs procaryotes les plus puissants connus pour l'expression de ses gènes précoces<sup>63,86</sup>.

Le séquençage du phage T5 a révélé 16 ORFs dans l'ADN du FST. Les ORF01 à 09 sont transcrites de droite à gauche et les ORF10 à 16 de gauche à droite (**Figure 11**).



**Figure 11** : Carte génétique des gènes précoces du phage T5. L'ORF 16 se termine à la base 9157, le FST a donc une longueur de 9,2 kpb. Les promoteurs identifiés comme très forts sont indiqués par des flèches vertes.

Parmi les gènes précoces, seuls les gènes *A1*, *A2* et *dmp* ont été caractérisés grâce à l'isolement de mutants ambre pour chacun de ces 3 gènes : l'intégration d'un codon stop ambre dans la séquence du gène conduit à la production d'une protéine tronquée qui perd généralement sa fonction, modifiant le comportement des phages mutants. Certaines bactéries utilisent un ARNt reconnaissant les codons ambre, ce qui permet la production d'une protéine entière et l'infection par ces phages mutants. Les protéines *A1* et *A2* sont requises pour déclencher la reprise du transfert de l'ADN restant du phage<sup>88</sup>. *dmp* n'est pas un gène essentiel pour l'infection, il intervient dans l'hydrolyse du groupement phosphate des dNMP 5'<sup>98</sup>. L'arrêt de la croissance de l'hôte sauvage est observé lors de l'infection par un phage mutant ambre pour le gène *A1* ou *A2*. Cette observation suggère qu'il existe des gènes toxiques parmi les autres gènes précoces inconnus. Bien que seuls 8 % de l'ADN du phage soient transférés au début de l'infection, les gènes codés par cet ADN sont particulièrement efficaces pour altérer les fonctions de l'hôte.

Plusieurs phages mutants pour des gènes précoces ont été isolés sans effet sur leurs propriétés infectieuses chez BF23 qui est très similaire au phage T5. Les ORF 09 à 16 ont été effacés chez BF23 sans affecter la viabilité en laboratoire ni modifier la position d'arrêt d'injection de



l'ADN. Il existe donc une séquence spécifique sur l'ADN responsable de l'arrêt du transfert, nommée signal d'arrêt de l'injection (ISS : injection stop signal) <sup>99</sup>. Par ailleurs, une région de 3,9 kpb des gènes précoces qui inclut les ORF 09 à 16 chez T5 ne permet pas d'observer de transcription et de traduction détectable dans des systèmes de production acellulaire (région représentée sur la **Figure 11**) <sup>100</sup>. Les gènes du FST ne sont donc pas tous nécessaires pour prendre le contrôle de la bactérie. Le groupe minimal de gènes nécessaires aux étapes précoces de l'infection par le phage T5 reste encore à identifier. C'est parmi ces gènes que l'on s'attend à identifier des fonctions toxiques qui pourraient se révéler de nouveaux antibactériens.

### 2.3.2 La protéine A1

A1 a une masse moléculaire de 61,7 kDa et contient 556 acides aminés. Elle est une protéine clé de l'infection, impliquée dans de nombreux aspects de la prise de contrôle de l'hôte. Lors de l'infection avec un phage mutant pour le gène A1, la seconde partie de l'ADN du phage n'est pas transférée, le chromosome de l'hôte reste intact et la synthèse des protéines de l'hôte et du FST continue d'avoir lieu pendant plus d'une heure après l'infection. La protéine A1 est donc essentielle pour la reprise du transfert de l'ADN après les étapes précoces et elle est impliquée dans la dégradation massive du chromosome de l'hôte <sup>73,101-103</sup>. L'arrêt de la synthèse des protéines de l'hôte est probablement lié à la dégradation du chromosome dont A1 est responsable.

Des travaux de McCorquodale en 1981 indiquaient que A1 peut être co-immunoprécipitée avec des protéines liées à l'ADN comme l'ARN polymérase. Les auteurs proposaient que la protéine A1 peut interagir sur l'ARN polymérase pour rediriger son activité vers la transcription des gènes injectés plus tard pendant l'infection, mais pensaient qu'elle n'avait pas d'activité nucléase <sup>72,101</sup>. A la place, elle activerait un facteur de l'hôte responsable de la dégradation de l'ADN. Toutefois, la possibilité que la protéine A1 puisse empêcher la transcription en dégradant l'ADN a été envisagée en 1968 <sup>73, 103</sup>.

Les séquençages ont révélé deux possibilités de codon start de la protéine A1. Dans ce manuscrit, je distingue la version longue lorsque la séquence de la protéine utilisée commence à la première Méthionine possible et la version courte lorsqu'elle commence à la seconde Méthionine.

## Séquence de la Protéine A1

MVISAEKQTVILKMAADFNFYGKRLRATKLEVCDDISKMVDYDTTKHSTAICDWLEANKPAKP  
KAAKVAKAIKNDERPEAAGIVSSTVEQWEVKQGKRFIITSIQNNTFPHKNFLASLEQYAQFI  
GADLLVSKFIYNKNGFQNGEGADGIRYDSAFDKYICNKNVFLNNRRFAFMAEINVLPADYD  
LSGFAETATALNLEGLAIGAANKITAESVPALKGEIVRRMYSTGTATLKNYIQQKAGQKAEAL  
HNF GALIVEFDEDEGEFFVRQLETMDESGVFYDLNTCATPAGCYETTGHVGLQYGDIAEKL  
DEECAASWGHGDTYGLVDILKPKYQFVHDVHDFTSRNHHNRASGVFLAKQYAAGRDKVLDD  
LIDTGRVLESMERDFSQTIIVESNHDLALSRLDNRNANIKDDPANAELYHRLNAAIYGAIA  
EKDDTFNVLDYALRKVAGCEFNALFLTTDQSFKIAGIECGVHGHNGINGSRGNPKQFKKLGK  
LNTGHTHTASIIYGGVYTAGVTGSLDMGYNVGASSWTQTHIITYANGQRTLIDFKNGKFFA

La version longue (début de la séquence en noir) aurait un nombre d'acides aminés de 556 et une Masse moléculaire de 61463,23 Da alors que la version courte (début de la séquence en bleu) aurait un nombre d'acides aminés de 543 et une Masse moléculaire de 60021,44 Da.

La protéine A1 que j'ai étudiée dans ma thèse est l'isoforme courte. Elle a été produite en clonant le gène dans un plasmide pCDF et elle est fusionnée à son extrémité N-terminale à une étiquette 6-His et à un site de clivage reconnu par l'Enterokinase.

### 2.3.3 La protéine A2

A2 est une protéine de 135 acides aminés et d'une masse moléculaire de 14,3 kDa. Elle est capable d'interagir avec l'ADN, mais sa spécificité d'interaction et son affinité pour l'ADN n'ont pas été déterminées<sup>104,105</sup>. Lors de l'infection avec un phage mutant pour le gène A2, l'ADN du SST n'est pas transféré. L'organisation de la protéine A2 en dimère a été démontrée par filtration sur gel et par réticulation de la protéine purifiée et migration sur gel SDS-Page<sup>104</sup>. Il a été montré que la protéine A2 interagit avec l'enveloppe bactérienne et pourrait perturber l'organisation de sa structure en interagissant avec la protéine membranaire OmpF. Le profil de migration d'enveloppes bactériennes fractionnées est altéré après infection avec un phage sauvage ou mutant pour A1 alors qu'il ne change pas par rapport à une souche non-infectée après infection avec un phage mutant pour A2. Ces résultats suggéraient que A2 pourrait agir au niveau de l'enveloppe bactérienne et participer au « déblocage » du transfert<sup>106</sup>. Cependant, cette hypothèse s'appuie sur des approches biochimiques peu fiables pour caractériser une interaction stable avec les membranes et il est nécessaire de la réévaluer.

L'interaction entre les protéines A1 et A2 a longtemps été recherchée. Le marquage radioactif et l'extraction de protéines totales après infection par le phage T5 ont été utilisés pour identifier les protéines produites avant le développement des techniques de séquençage<sup>88</sup>. Ces méthodes

étaient très imprécises et en particulier, elles ne permettaient pas d'identifier de manière certaine la composition protéique des complexes identifiés sur gel d'acrylamide non dénaturants.

#### 2.3.4 *Autres protéines précoces*

Parmi les autres protéines précoces, seule la deoxyribonucleotidase 5' monophosphatase (dmp) a un rôle connu. Cette enzyme (polypeptide de 244 acides aminés) hydrolyse le groupement phosphate des dNMP 5', mais son activité n'est pas essentielle à la croissance du phage T5<sup>76,98</sup>. Une bactérie infectée par un phage mutant pour dmp n'excrète pas tous les produits finaux de dégradation de l'ADN et réincorpore 30% des thymines de l'hôte alors qu'un phage sauvage n'en réincorpore pas<sup>98</sup>. En comparaison, chez T7, 90% des nucléotides de l'hôte sont réincorporés dans le génome du phage, ce qui est une étape essentielle de la réplication. Chez T4, la dégradation de l'ADN de l'hôte n'est pas essentielle, mais 20% des nucléotides sont réincorporés. Chez ces phages, la dégradation et la synthèse de l'ADN ont lieu en même temps, contrairement à ce qui a été observé pour le phage T5.

La mutation de dmp n'a presque aucun effet significatif sur la croissance du phage T5, seul un retard dans l'expression de certaines enzymes et la production de l'ADN est observé<sup>98</sup>.

Nous possédons des mutants de A1 et de A2 qui stoppent le transfert de l'ADN de T5 à la première étape. Ces mutants permettent d'étudier l'effet des gènes précoces sur la bactérie indépendamment des étapes plus tardives. Dans ce contexte, le phage T5 est un modèle privilégié pour étudier les étapes précoces de l'infection dans lesquelles on trouve le plus grand nombre de protéines toxiques impliquées dans la prise de contrôle de l'hôte.

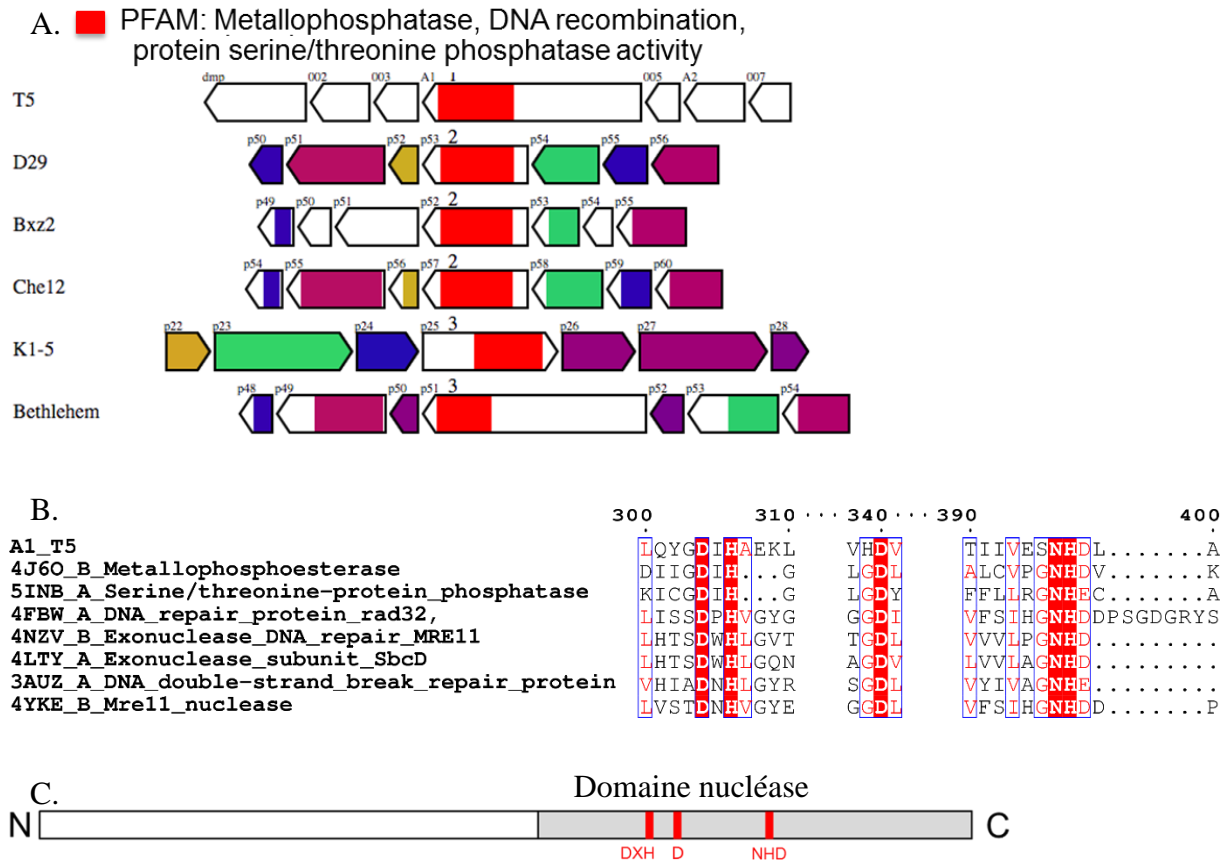
### 3) Sujet de thèse : Etude des protéines A1 et A2 du bactériophage T5

#### 3.1 Travaux préliminaires sur les protéines A1 et A2

##### 3.1.1 Etude bioinformatique

Jusqu'à très récemment, aucune homologie de séquence significative pour les gènes A1 ou A2 ne pouvait être mise en évidence *via* les logiciels d'alignement BLAST et HH-Pred. Seules des protéines A1 des phages apparentés au phage T5 (*T5viruses*) étaient retrouvées. L'outil Phagonaute <sup>107</sup> compare les génomes de phages entre eux. Il a permis d'identifier une homologie avec des métallophosphatases dans la région C-terminale de la protéine A1 (**Figure 12 A**). Les métallophosphatases sont des DNases utilisant des cations divalents pour exercer leur activité. Parmi les métallophosphatases connues, on retrouve Mre11 dont la structure a été résolue. Cette enzyme est représentative d'une grande famille de DNases dépendantes au manganèse impliquées dans la recombinaison homologue et la réparation de l'ADN <sup>108</sup>. Les protéines de la famille de Mre11 sont présentes dans de nombreux organismes. On trouve par exemple la protéine gp47 chez le phage T4 qui a un rôle similaire <sup>109</sup>. Mre11 apparaît dans un alignement de séquence avec la protéine A1 obtenu sur le serveur HHPred <sup>110</sup> (**Figure 12 B**). Ces résultats suggèrent que la protéine A1 pourrait être une DNase. Une représentation schématique de la protéine A1 basée sur l'alignement de séquence permet de voir que le domaine C-terminal possède plusieurs résidus potentiellement impliqués dans une activité DNase (**Figure 12 C**). En revanche, le rôle du domaine N-terminal est inconnu et n'a pas pu être déterminé à l'aide des logiciels de prédiction d'homologie.

Les activités identifiées pour la protéine A2 avec ces logiciels avant ma thèse présentaient un taux de fiabilité faible qui ne permettait pas de tirer de conclusions. La protéine LexA repressor qui est un régulateur transcriptionnel avait déjà été identifiée à ce moment-là.



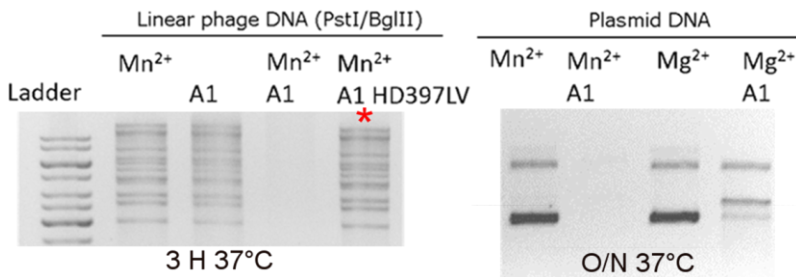
**Figure 12** : Extraits d'alignements de la protéine A1 du phage T5 obtenus avec Phagonaute (A) et HHpred (B). C : Schéma positionnant le domaine nucléase putatif de la protéine A1.

### 3.1.2 Travaux préliminaires sur la protéine A1

#### 3.1.2.1 L'activité DNase de la protéine A1

La toxicité de la protéine A1 est un obstacle à son expression. Des vecteurs d'expression pouvant être réprimés par l'ajout de glucose ont été utilisés pour cloner la protéine A1. L'obtention de la protéine A1 purifiée avec ou sans étiquette a permis des études fonctionnelles préliminaires et la production d'anticorps polyclonaux dirigés contre elle.

Des tests préliminaires d'activité DNase ont été faits sur la protéine A1. Différents ions ont été utilisés et les ions  $Mn^{2+}$  ont été les plus efficaces pour provoquer la disparition de l'ADN par la protéine A1 (**Figure 13**). Ces résultats sont une forte indication que la protéine A1 possède une activité DNase, bien qu'aucune activité nucléase associée à la protéine A1 n'ait été décrite dans la littérature <sup>88</sup>.



**Figure 13 : Tests d'activité préliminaires sur la protéine A1.** 0,4  $\mu M$  A1, 1 mM  $MnCl_2$  ou  $MgCl_2$ . A1 HD397LV est un mutant pour des acides aminés conservés chez les métallophosphatases similaires à A1.

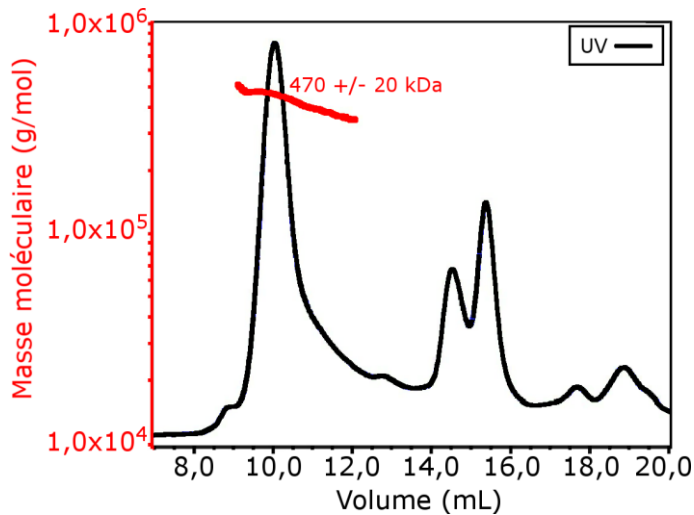
L'activité a aussi été testée avec un mutant de la protéine A1 pour l'Histidine 397 et l'Aspartate 398 qui sont conservés chez les métallophosphatases d'après les prédictions bioinformatiques. Ce mutant n'a pas fait disparaître l'ADN dans les mêmes conditions que la protéine sauvage. A ce stade, nous ne savons pas si la structure de ce mutant était conservée par rapport à celle de la protéine sauvage. L'objectif était alors d'utiliser ce mutant pour produire la protéine en évitant les problèmes de toxicité liés à la version sauvage et permettre des études structurales.

Cette activité DNase de la protéine A1 soulève de nombreuses questions : Est-elle la DNase longtemps restée inconnue responsable de la dégradation massive du chromosome de l'hôte ? Comment l'ADN du phage est-il protégé de la dégradation par la protéine A1 ? Cette activité est-elle liée à la reprise du transfert de l'ADN du phage ? Le domaine N-terminal de la protéine A1 reste inconnu : s'agit-il d'un domaine indépendant, joue-t-il un rôle dans son activité DNase ou est-il impliqué dans d'autres étapes de l'infection ?

Des protéines similaires à A1 sont présentes dans de nombreux autres phages d'après les recherches de similarité par BLAST. Par exemple, le gène 126 du phage phiCbK de *Caulobacter crescentus* code pour une protéine ayant 36% d'identité avec la protéine A1 sur 85% de sa séquence en incluant le domaine DNase <sup>111</sup>. Ces autres bactériophages n'ont pas tous une protéine homologue à A2 et ne partagent donc probablement pas avec T5 le processus d'injection en deux étapes. L'activité des protéines A1-like est-elle proche chez ces autres bactériophages ? Elucider le fonctionnement de la protéine A1 chez le phage T5 pourrait permettre de comprendre le rôle de ces protéines similaires à A1 dans un contexte beaucoup plus large.

### 3.1.2.2 Etudes structurales

Des études biochimiques préliminaires ont révélé que la protéine A1 purifiée et concentrée peut adopter une organisation oligomérique. La concentration de la protéine sur centricon provoque son oligomérisation, visible en SEC MALLS (**Figure 14**). Une masse moléculaire de 470 +/- 20 kDa avait été observée, suggérant que la protéine A1 forme un octamère. Cette organisation est surprenante : en effet, de nombreuses nucléases connues forment des dimères. L'étude de cette organisation particulière pourrait révéler un nouveau type d'activité DNase inconnu.



**Figure 14.** Analyse de l'état oligomérique de A1 en solution par SEC-MALLS.

### 3.1.3 Travaux préliminaires sur la protéine A2

Des tests de production de la protéine A2 ont révélé qu'elle est très soluble et peu toxique. Il est donc facile de la produire en grandes quantités, ce qui devrait faciliter les études structurales. Des anticorps polyclonaux contre la protéine A2 purifiée sans étiquette ont été obtenus.

## 3.2 Objectifs de la thèse

Comprendre le fonctionnement et l'activité propres de chaque protéine virale est essentiel pour mieux comprendre le fonctionnement des phages et identifier des mécanismes à cibler pour proposer de nouveaux traitements antibactériens.

Dès les premières minutes de l'infection par T5 et l'injection du fragment FST, l'ADN de l'hôte est massivement dégradé. La DNase responsable de cette activité n'avait pas été identifiée avant ma thèse bien que cette activité soit connue depuis 1958<sup>74</sup>. Des études préliminaires sur la protéine A1 ont révélé une activité DNase qui pourrait expliquer son rôle dans la dégradation massive du chromosome de l'hôte dans les étapes précoces de l'infection. Toutefois, la spécificité d'action de cette DNase n'avait pas été étudiée et il était nécessaire de déterminer sa structure pour comprendre son mode d'action. L'existence d'un équilibre entre une forme monomérique et une forme oligomérique avait été mise en évidence, mais l'état oligomérique qui n'était connu chez aucune autre DNase devait être caractérisé. De plus, A1 étant une protéine toxique, il fallait optimiser ses conditions de production et de purification pour effectuer les études structurales.

Le principal objectif de mon travail de thèse a été d'élucider la structure et l'activité de la protéine A1. Pour cela mon travail s'est axé autour de deux aspects. Tout d'abord la caractérisation de la structure de la protéine A1 par diverses approches : SEC MALLS, SAXS, Dichroïsme Circulaire, Cristallographie aux rayons X et cryo-microscopie électronique. Pour étudier l'équilibre de formation de l'oligomère, nous avons observé la protéine par microscopie électronique à coloration négative après incubation dans différentes conditions. La cryo-microscopie électronique nous a permis de résoudre la structure de la protéine A1 à 3,0 Å de résolution, ce qui a permis sa comparaison avec d'autres structures connues pour mieux étudier sa fonction.

Le second aspect de l'étude a porté sur la caractérisation de l'activité DNase de la protéine A1. Pour cela, j'ai effectué une étude fonctionnelle de la protéine en observant son activité sur différents substrats. La construction de mutants de la protéine a permis l'identification de certains résidus essentiels à son activité : Aspartate 304, Histidine 397 et Asparagine 398. Ces mutants m'ont également permis de faire des tests de complémentation afin de tester l'importance de cette activité DNase lors de l'infection virale. J'ai pu observer que l'activité DNase de la protéine A1 est indispensable pour l'infection par le bactériophage T5.



Le second objectif de mon travail de thèse était la caractérisation de la protéine A2 qui est un partenaire essentiel de A1 pour permettre la deuxième étape de l'injection du génome. Nous avons mené des études préliminaires de structure en SAXS qui ont révélé son organisation dimérique et fait des essais de cristallogénèse. Des essais préliminaires de RMN avec l'obtention d'un spectre HSQC ont révélé que l'étude de la structure de la protéine A2 par cette technique est possible.

Des protéines similaires à A1 sont présentes dans de nombreux bactériophages. Ce travail est un premier pas vers la compréhension de sa fonction dans des conditions variées.

---

## II. Résultats

---

## 1) Caractérisation de la protéine virale A1

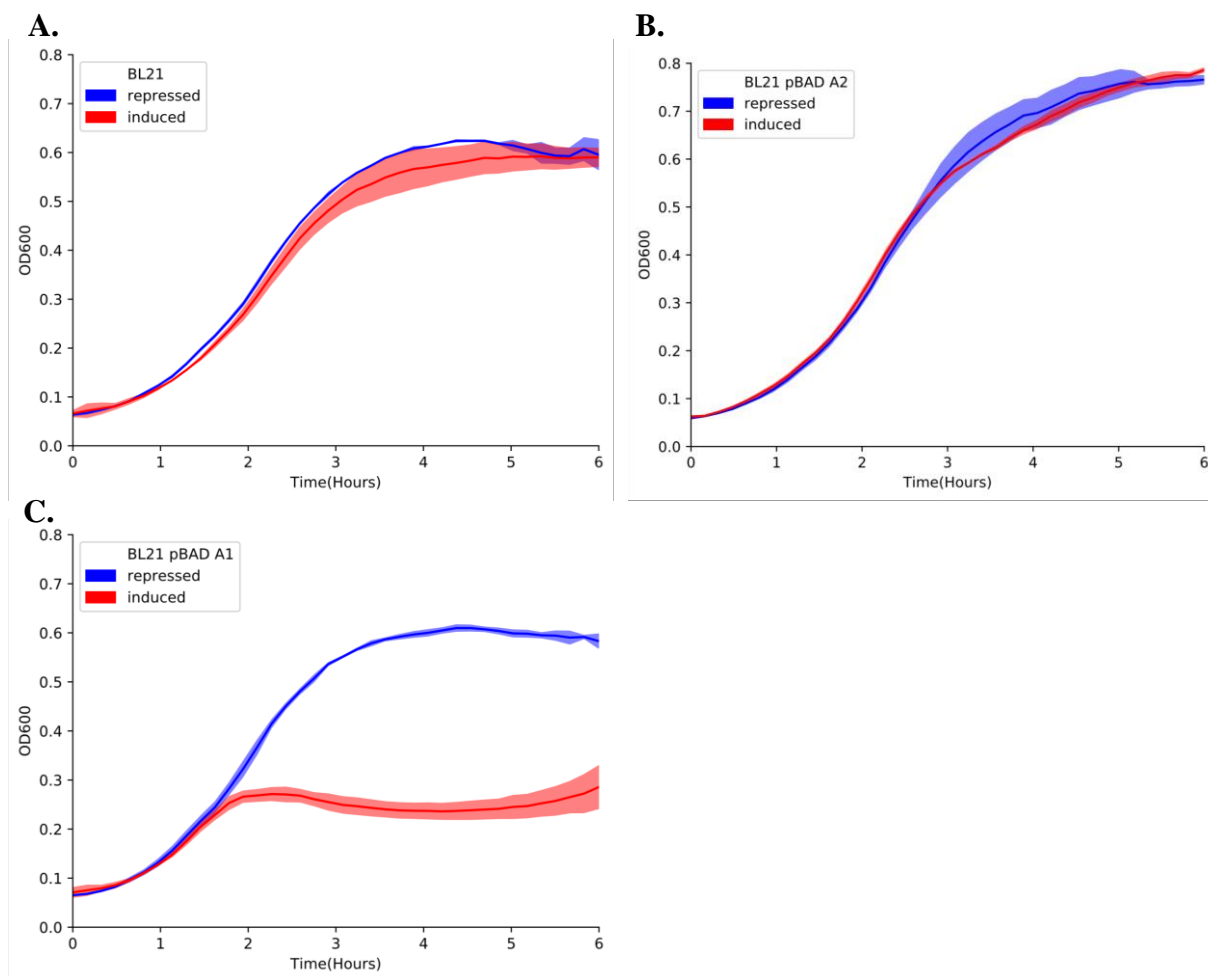
### 1.1 Études structurales

#### 1.1.1 Mise au point des conditions de production et purification de A1

##### 1.1.1.1 Toxicité de la production de A1 dans la bactérie

Pour pouvoir caractériser les propriétés biochimiques de la protéine A1 et étudier sa structure, j'ai optimisé les conditions de production et de purification de cette protéine. Le premier obstacle à la production de A1 en grandes quantités est sa toxicité. Pour la tester, j'ai utilisé des bactéries contenant un plasmide où le gène viral *A1* était cloné sous le contrôle d'un promoteur inductible par l'arabinose. J'ai déterminé la croissance de cultures effectuées sur 6 heures en induisant ou non l'expression de *A1* dans des bactéries *E. coli*. Les contrôles comprennent des bactéries sans plasmide ou contenant un plasmide portant le gène d'une autre protéine précoce de T5, la protéine A2, qui n'affecte pas la croissance bactérienne de manière significative et que l'on peut donc considérer comme non toxique.

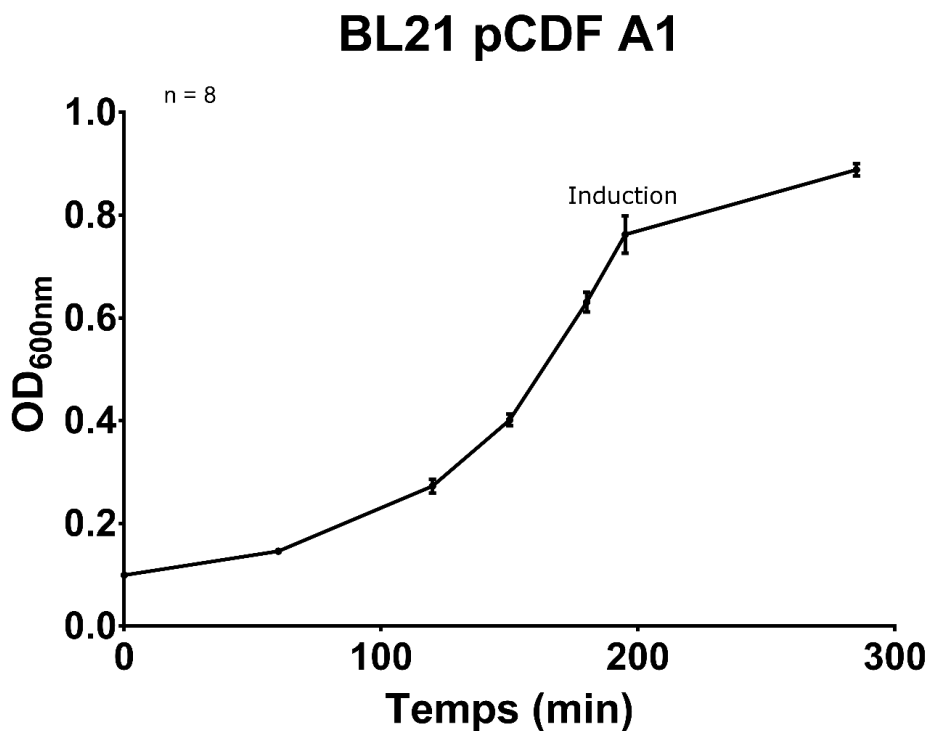
La croissance bactérienne ne varie pas en présence ou en absence d'inducteur pour les témoins sans plasmide ou avec une protéine non toxique (**Figure 15** A et B). L'inducteur n'est donc pas toxique pour les bactéries et la surexpression d'un gène non toxique comme A2 chez le témoin n'altère pas la croissance des bactéries. Au contraire, les bactéries exprimant *A1* montrent une croissance inhibée dès deux heures après l'addition de l'arabinose (**Figure 15** C). A1 est donc bien une protéine toxique, ce que l'on pouvait attendre étant donné son rôle présumé dans la dégradation de l'ADN bactérien.



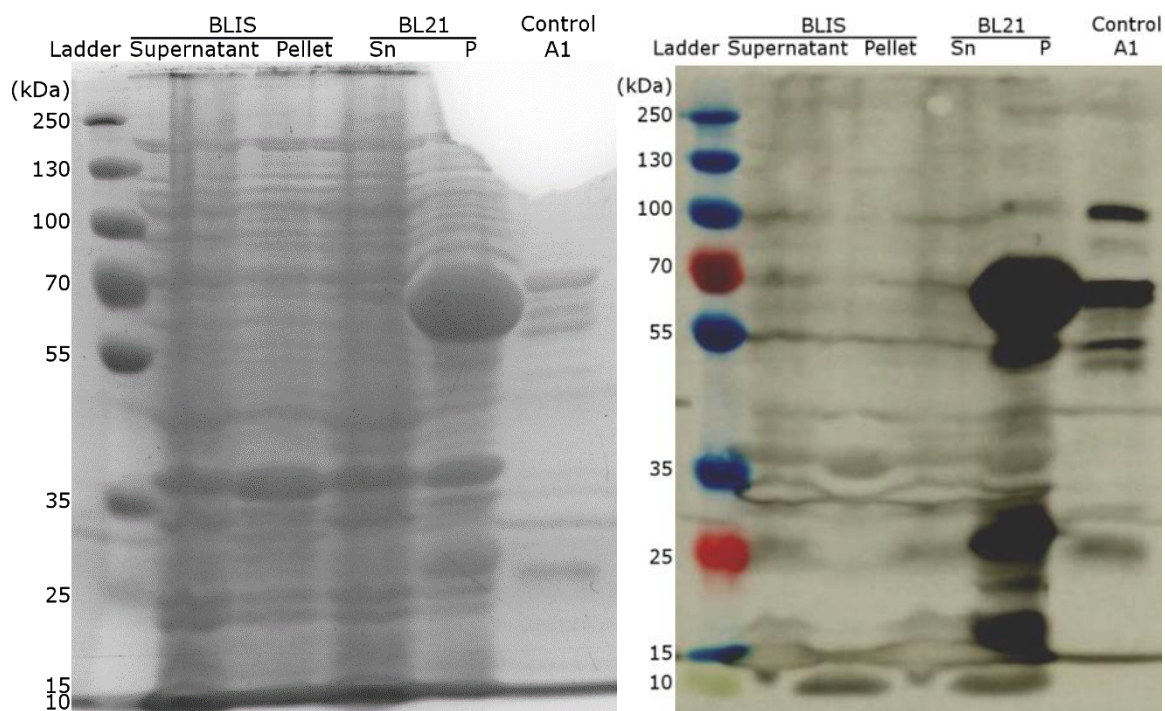
**Figure 15. Évaluation de la toxicité de A1.** Le promoteur contrôlant l'expression du gène cloné dans pBAD24 est induit par l'arabinose ou réprimé par le glucose. Les bactéries en phase exponentielle sont diluées dans le milieu de culture contenant de l'arabinose ou du glucose à  $t = 0$  min et cultivées dans les puits d'une microplaque à 37°C. La densité optique des cultures est mesurée au cours du temps ( $n = 5$ ). **A.** Témoin sans plasmide. **B.** Témoin produisant la protéine virale A2 non toxique. **C.** Impact de la production de la protéine virale A1 sur la croissance de *E. coli*.

Le plasmide pBAD24 n'est pas adapté pour produire A1 car il ne permet pas d'obtenir des quantités suffisantes de protéines. J'ai donc utilisé un plasmide d'expression pCDF (pCDF-2 Ek\_LIC, Novagen) qui possède un promoteur T7 inductible par l'IPTG, dans une souche BL21 ( $\lambda$  DE3) chez laquelle l'expression de la polymérase du phage T7 peut être réprimée par l'addition de glucose dans le milieu de culture <sup>112</sup>.

Les cultures bactériennes sont faites en milieu LB à 37°C et pour produire A1 malgré sa toxicité, l'induction a lieu seulement lorsque la densité optique est supérieure à 0,8. La biomasse obtenue permet alors de produire suffisamment de protéines pendant 1h30 d'induction. L'induction de l'expression de A1 dans ces conditions de culture est très toxique, comme le montre le fort ralentissement de la croissance bactérienne après induction (**Figure 16**). Parmi les deux isoformes possibles de A1 (556 et 543 acides aminés), nous avons choisi d'exprimer la version plus courte car nous n'avons pas pu obtenir l'isoforme longue sous forme soluble (**Figure 17**), contrairement à l'isoforme courte. J'ai démontré que cette isoforme courte de A1 est active *in vivo*, en effectuant des tests de complémentation fonctionnelle (voir paragraphe 1.4.2 et **Figure 84**). L'isoforme courte de A1 ainsi exprimée a pu être extraite et purifiée dans les conditions décrites au paragraphe suivant.



**Figure 16.** Croissance des bactéries BL21(DE3) transformées avec pCDF A1 lors de la production de la protéine A1 induite par l'IPTG (0,4 mM). Les points représentent la moyenne et l'écart-type standard de la densité optique de huit cultures. La croissance des bactéries non induites par l'IPTG n'a pas été suivie systématiquement en parallèle de ces cultures, mais j'ai mesuré une DO<sub>600nm</sub> comprise entre 2 et 3 après le même temps de culture sans IPTG.

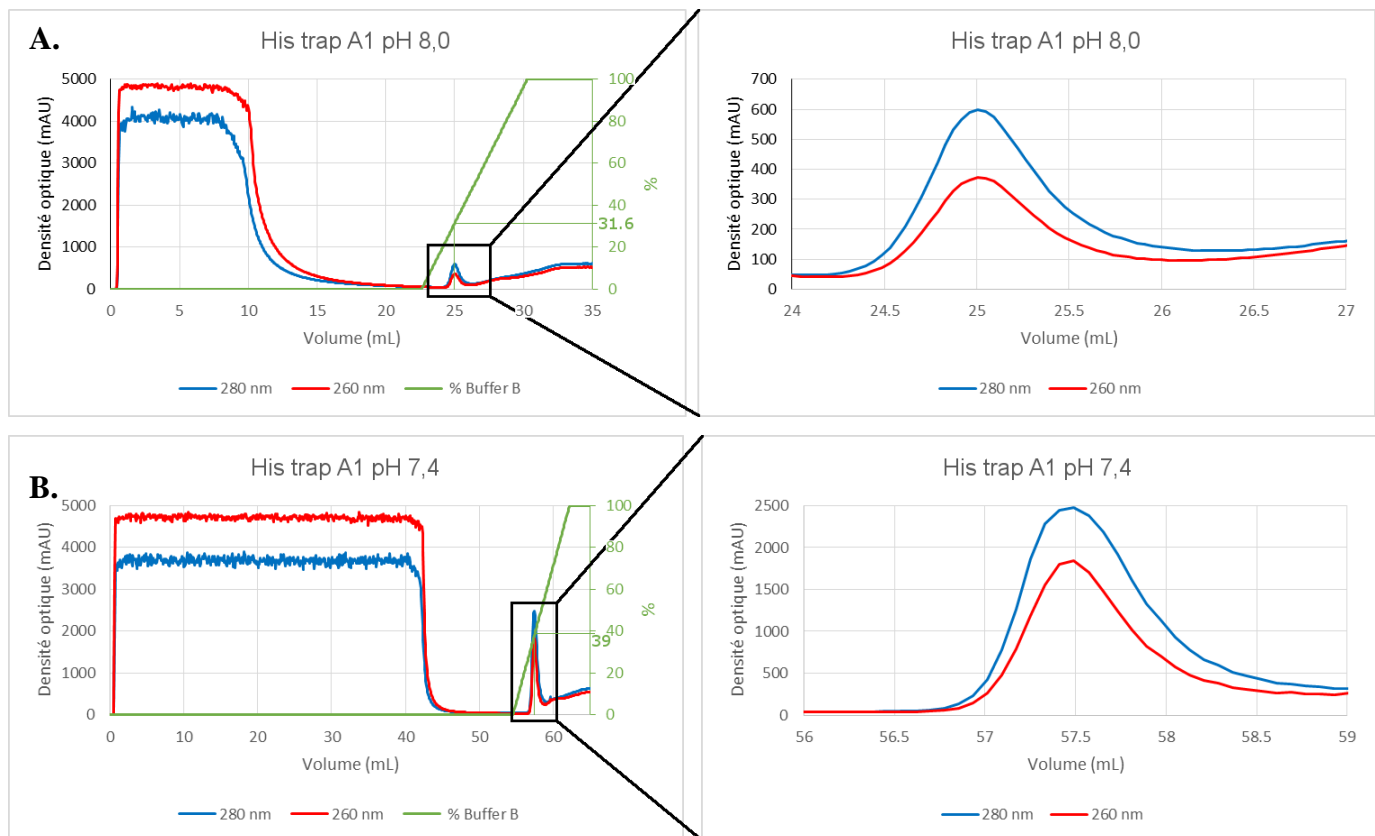


**Figure 17. Production de l'isoforme longue de A1 à partir du plasmide pET28b.** Coloration du gel SDS-PAGE au bleu de Coomassie et révélation du Western Blot avec des anticorps anti-A1. BLIS et BL21 sont deux souches bactériennes exprimant l'ARN polymérase de T7 sous contrôle d'un promoteur LacUV5, nécessaire pour l'expression du gène cloné dans le plasmide pET28b. La souche BLIS contient en plus le plasmide pLysS (voir Matériel et Méthodes). L'extrait total bactérien après la lyse des bactéries est centrifugé pour séparer le surnageant (supernatant, Sn) du culot (pellet, P) contenant les bactéries non lysées et les corps d'inclusion.

### 1.1.1.2 Optimisation des conditions de purification

Grâce aux analyses bioinformatiques qui prédisent une fonction DNase pour A1, nous pouvions prédire que cette protéine aurait une forte probabilité de se lier à l'ADN. Afin de minimiser la co-purification d'ADN avec A1, j'ai effectué la première étape de purification par affinité sur colonne de Nickel en présence de 1M de NaCl, ce qui diminue les interactions entre l'ADN et les protéines par compétition.

J'ai également testé l'effet du pH sur l'efficacité de la purification : un premier essai à pH 8,0 n'a permis de récupérer que très peu de protéine (**Figure 18 A**). Les protéines non retenues sur la colonne ont été diluées avec un tampon Tris à pH 7,4 et réinjectées sur la même colonne à ce pH. Un bien meilleur rendement a été obtenu, ce qui montre que la purification à pH 8,0 n'est pas adaptée par rapport à la purification à pH 7,4 (**Figure 18 B**).



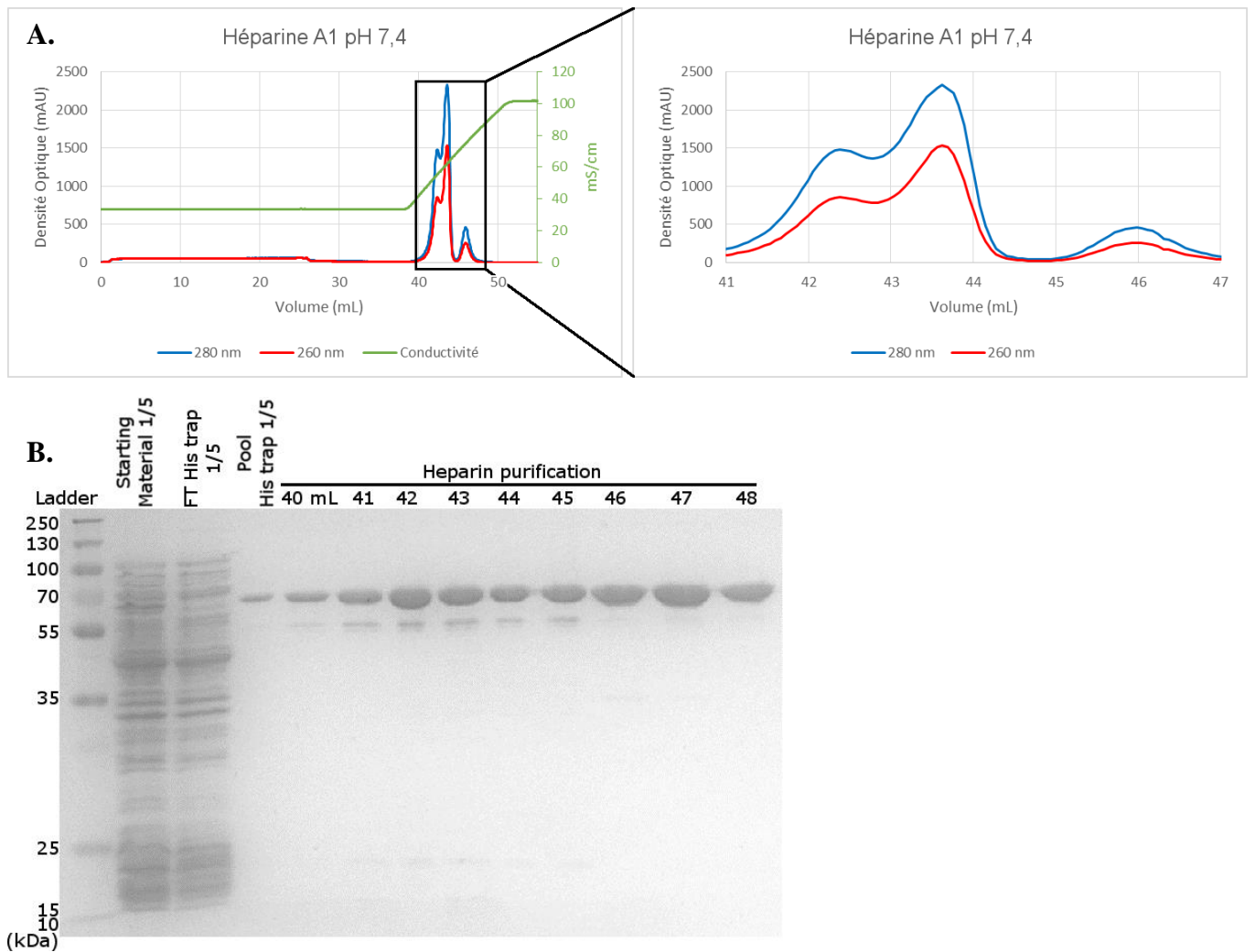
**Figure 18.** Chromatogrammes de la purification de A1 sur colonne His Trap FF. **A.** Injection d'un extrait de protéines cytoplasmiques de bactéries BL21(DE3)-pCDF A1. Tampon d'éluion : Tris 50 mM pH 8,0 ; NaCl 1 M ; Imidazole 1 M à pH 8,0. **B.** Réinjection sur la même colonne des fractions contenant les protéines non retenues à pH 8,0 (0 à 15 mL, dans le panneau A). Tampon d'éluion : Tris 50 mM pH 7,4 ; NaCl 1 M ; Imidazole 1 M.

Bien que la plupart des contaminants aient été éliminés à l'issue de cette première étape (voir la 4<sup>ème</sup> piste du gel SDS-PAGE de la **Figure 20 B**), nous avons choisi de compléter la purification avec une colonne d'héparine car c'est un support couramment utilisé pour la purification des protéines interagissant avec l'ADN. La résine des colonnes d'héparine contient des groupes sulfates chargés négativement, capables d'interagir avec des groupements chargés positivement telles que les chaînes latérales des lysines et des arginines fréquemment impliquées dans la liaison des acides nucléiques<sup>113,114</sup>. De nouveau, j'ai comparé l'effet du pH lors de cette purification sur colonne d'héparine. Les essais à différents pH présentés sur la Figure 19 n'ont pas été faits le même jour, mais les résultats présentés ont été reproduits plus de 3 fois pour chaque test. L'essai de purification à pH 8,0 présenté ici fait suite aux essais de purification en His Trap à pH 8,0 exposés ci-dessus.

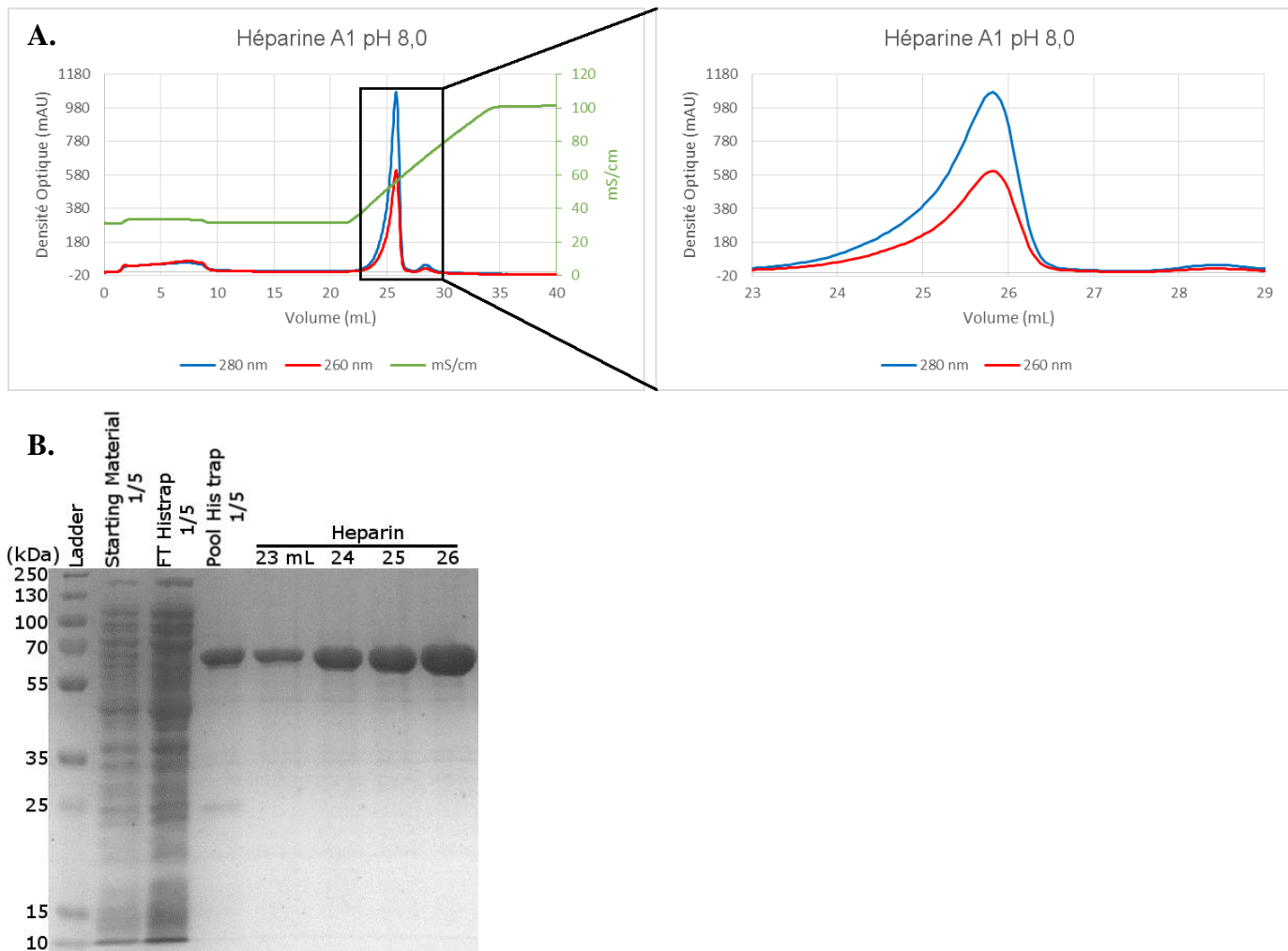
Si la purification sur héparine est faite à pH 7,4, de nombreux pics sont visibles (**Figure 19 A** et B). A1 est présente dans chacun de ces pics d'après l'analyse faite par électrophorèse sur gel SDS-PAGE (**Figure 19 B**). Une protéine d'environ 55 kDa est également présente. Il s'agit probablement d'une forme tronquée de A1 ou d'un contaminant. A pH 8,0 un pic majeur est obtenu, qui correspond à l'élution de A1. Cette fois, la protéine est pure, sans contaminant ni dégradation partielle de la protéine sur l'ensemble du pic (**Figure 20**). La purification de A1 en héparine est donc plus efficace et de meilleure qualité à pH 8,0.

En conclusion, cette optimisation du protocole de purification m'a permis d'obtenir des quantités suffisantes de protéine pour faire des études structurales. On obtient environ 2 mg de protéine purifiée pour 3 litres de culture.





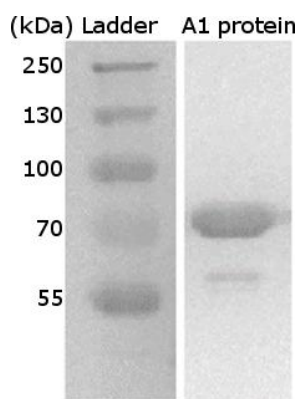
**Figure 19. Purification de A1 sur colonne Héparine à pH 7,4.** **A.** Chromatogramme de la purification. **B.** Analyse des fractions sur gel SDS-PAGE coloré au bleu de Coomassie. "Starting Material" : Protéines contenues dans le surnageant obtenu après cassage du culot bactérien et ultracentrifugation, qui sont injectées sur la colonne His Trap FF. FT = "Flow-Through" : protéines non retenues sur la colonne His Trap FF (fraction de 0 à 5 mL déposée sur gel). "Pool" : fractions du pic élué de la colonne His Trap (facteur de dilution : 1/5). "Heparin" : fractions des pics élués de la colonne Héparine.



**Figure 20. Purification de A1 sur colonne Héparine à pH 8,0.** Cette expérience fait suite à la purification en colonne His Trap à pH 7,4 présentée en Figure 18 B. **A.** Chromatogramme de purification. **B.** Analyse des fractions sur gel SDS-PAGE coloré au bleu de Coomassie. "Starting Material" : Protéines obtenues dans le surnageant après extraction et ultracentrifugation, qui sont injectées sur la colonne His trap FF. FT = "Flow-Through" : protéines non retenues sur la colonne His trap FF (fraction de 0 à 5 mL déposée sur gel). "Pool" : fractions du pic élué de la colonne His Trap à pH 7,4 (facteur de dilution : 1/5). "Heparin" : fractions du pic élué de la colonne Héparine.

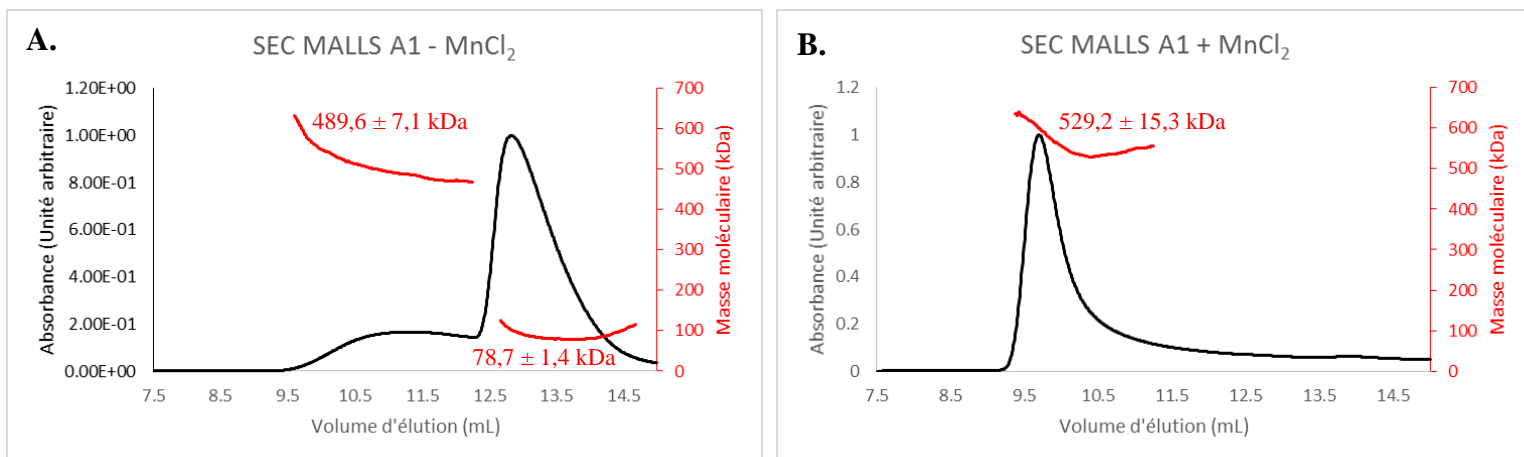
### 1.1.2 Analyse de l'état oligomérique de A1 par SEC-MALLS

Afin d'élucider l'organisation de la protéine A1 en solution, nous avons étudié son état oligomérique par SEC-MALLS. Nous avons fait le choix d'étudier l'influence du manganèse sur l'état d'oligomérisation de A1 pour deux raisons : i) les prédictions issues des analyses bioinformatiques indiquent une ressemblance avec la protéine Mre11 qui est une DNase dépendant des ions  $Mn^{2+}$ . ii) les premiers tests d'activité faits avec A1 confirment une activité DNase en présence de  $MnCl_2$  (voir introduction). La protéine purifiée sur colonne d'héparine (**Figure 21**) a été incubée sur la nuit en présence ou en absence de  $MnCl_2$ . Le lendemain, elle a été injectée sur une colonne de filtration sur gel puis directement analysée en sortie de colonne (**Figure 22**).



**Figure 21 : Protéine A1 utilisée pour le SEC MALLS.**

NB : L'échantillon de A1 utilisé provient d'une purification sur colonne Héparine effectuée à pH 7,4, alors que l'optimisation à pH 8,0 n'avait pas encore été faite, ce qui explique la présence de la bande à 55kDa.



**Figure 22 : Analyse de l'état oligomérique de A1 en solution par SEC-MALLS.** La protéine a été incubée pendant une nuit en absence (A) ou en présence (B) de 1 mM  $MnCl_2$ . En noir, absorbance, en rouge, masse moléculaire.

Sans  $\text{MnCl}_2$ , l'intégration de la courbe révèle que la protéine est majoritairement monomérique avec 77 % de monomères pour 23 % d'oligomères. Le premier pic correspond à une espèce protéique de masse moléculaire de  $489,6 \pm 7,1$  kDa. Le deuxième pic correspond à une espèce protéique d'une masse moléculaire de  $78,7 \pm 1,4$  kDa. La masse moléculaire théorique de A1 est de 61,7 kDa. Cette faible précision est due à la contamination par le signal de diffusion du premier pic qui incrémente la masse moléculaire mesurée.

Avec du  $\text{MnCl}_2$ , un seul pic est observé. La protéine est oligomérique avec une masse moléculaire estimée à  $529,2 \pm 15,3$  kDa.

Ces expériences montrent que A1 est en équilibre entre un état monomérique et un état oligomérique dont la masse apparente indique qu'elle pourrait être un octamère. Cet équilibre est fortement déplacé vers la formation de l'oligomère en présence de 1 mM de  $\text{MnCl}_2$ . Pour confirmer ces résultats, nous avons fait des analyses en diffusion des rayons X aux petits angles (SAXS).

### 1.1.3 Etude structurale de A1 en solution par SAXS

Le SAXS est une technique utilisée sur des échantillons en solution pour obtenir des informations structurales à basse résolution <sup>115</sup>. Avant l'obtention de cristaux et de clichés de cryo-microscopie électronique, le SAXS nous a donné de précieux indices sur la transition monomère/octamère en présence de MnCl<sub>2</sub>. Nous avons fait les expériences de SAXS en présence et absence de MnCl<sub>2</sub> afin de construire un modèle à basse résolution de l'enveloppe de A1 en solution.

Les expériences ont été faites sur la ligne BM29 de l'ESRF (European Synchrotron Radiation Facility). 100 µL de la protéine A1 sont injectés sur une colonne de filtration sur gel à une concentration d'environ 2 mg/mL en présence ou en absence de MnCl<sub>2</sub>. L'analyse est effectuée directement en sortie de colonne dans une cellule UV puis l'acquisition de l'intensité de diffusion aux petits angles est réalisée toutes les secondes au cours de l'élution sur une cellule SAXS (**Figure 23** A et B). L'ensemble des courbes accumulées au cours de l'élution est analysé à l'aide du module HPLC du programme US-SOMO <sup>116</sup>. De chaque courbe sont extraits les paramètres caractéristiques de l'objet diffusant : rayon de giration, extension maximale et masse, ce qui donne accès à l'état d'oligomérisation. La Figure 23 B montre dans les deux cas (avec et sans MnCl<sub>2</sub>) un large plateau pour la masse au voisinage du maximum du pic d'élution. On observe principalement une forme monomérique en absence de MnCl<sub>2</sub> estimée à 60 kDa et une forme oligomérique en présence de MnCl<sub>2</sub> estimée à 480 kDa. Cette valeur indique également que l'oligomère formé par A1 pourrait donc être un octamère. Ce résultat confirme que A1 s'oligomérisent en présence de MnCl<sub>2</sub>. Sans MnCl<sub>2</sub>, une troisième forme minoritaire apparaît et pourrait être un dimère de A1.

Les courbes mesurées successivement le long des plateaux sont comparées à l'aide du programme Cormap <sup>117</sup> et les courbes identiques sont moyennées, pour donner les courbes finales à analyser (**Figure 23** C et D). En l'absence de MnCl<sub>2</sub>, les frames 2060 à 2100 sont ainsi moyennés, et de 1800 à 1880 en présence de MnCl<sub>2</sub>.

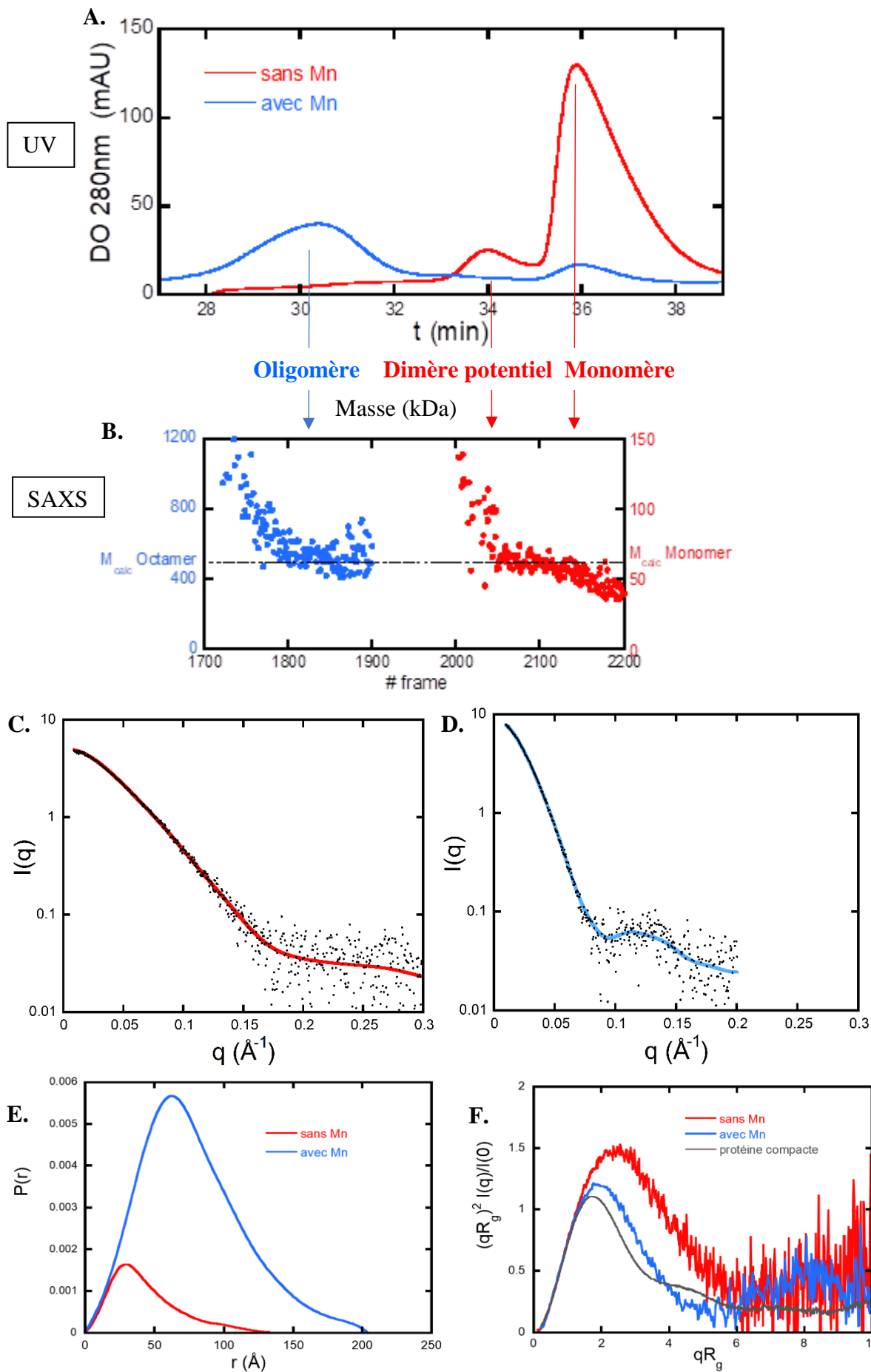
L'analyse de la courbe de SAXS en l'absence de MnCl<sub>2</sub> montre que l'objet diffusant est un monomère, avec un rayon de giration  $R_g$  de 35,3 Å et une extension maximale  $D_{max}$  de 136 Å. Ces valeurs sont très supérieures à celles attendues pour une protéine globulaire de N=557 résidus :  $R_g \sim 3(N)^{1/3} \sim 25$  Å et  $D_{max} = 2R_g \sqrt{5/3} \sim 65$  Å, et indiquent que la protéine est allongée, ce qui est confirmé par l'allure de la fonction de distribution des distances  $P(r)$  (courbe très asymétrique) et par l'allure de la courbe de diffusion portée en représentation de Kratky sans dimension (**Figure 23** E et F). En modélisant la courbe de diffusion  $I(q)$  à l'aide du programme

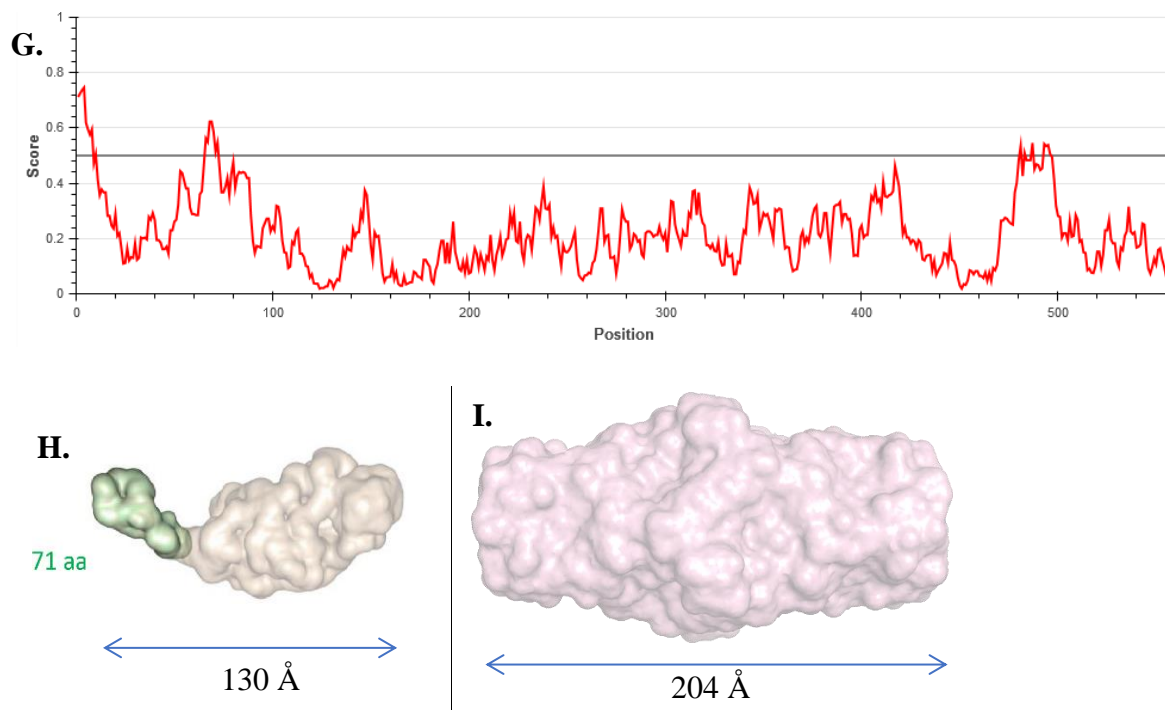
GASBOR<sup>118</sup>, qui représente la protéine comme une chaîne de pseudo-résidus, on peut obtenir un modèle d'enveloppe du monomère de A1 en solution, qui montre un grand domaine et un petit domaine, contenant environ 70 acides aminés, qui forment une partie séparée du reste de la protéine par un linker qui pourrait être flexible (**Figure 23 H**). L'examen de la séquence de la protéine laisse supposer que ce petit domaine correspond au domaine N-terminal car on trouve une région flexible d'après le serveur IUPred<sup>119</sup> entre les acides aminés 60 et 80 (**Figure 23 G**).

En présence de MnCl<sub>2</sub>, l'analyse de la courbe de SAXS montre que la masse de l'objet diffusant est multipliée par un facteur de l'ordre de 8. Le rayon de giration et l'extension maximale sont alors de 60,7 Å et 204 Å respectivement. Comme précédemment la modélisation *ab initio* de la courbe de SAXS a été conduite à l'aide du programme GASBOR, mais en imposant cette fois la symétrie P42 (**Figure 23 I**).

Les principales caractéristiques des courbes de SAXS détaillées ci-dessus sont résumées dans le **Tableau 7** page 139 donné dans la section Matériel et Méthodes.

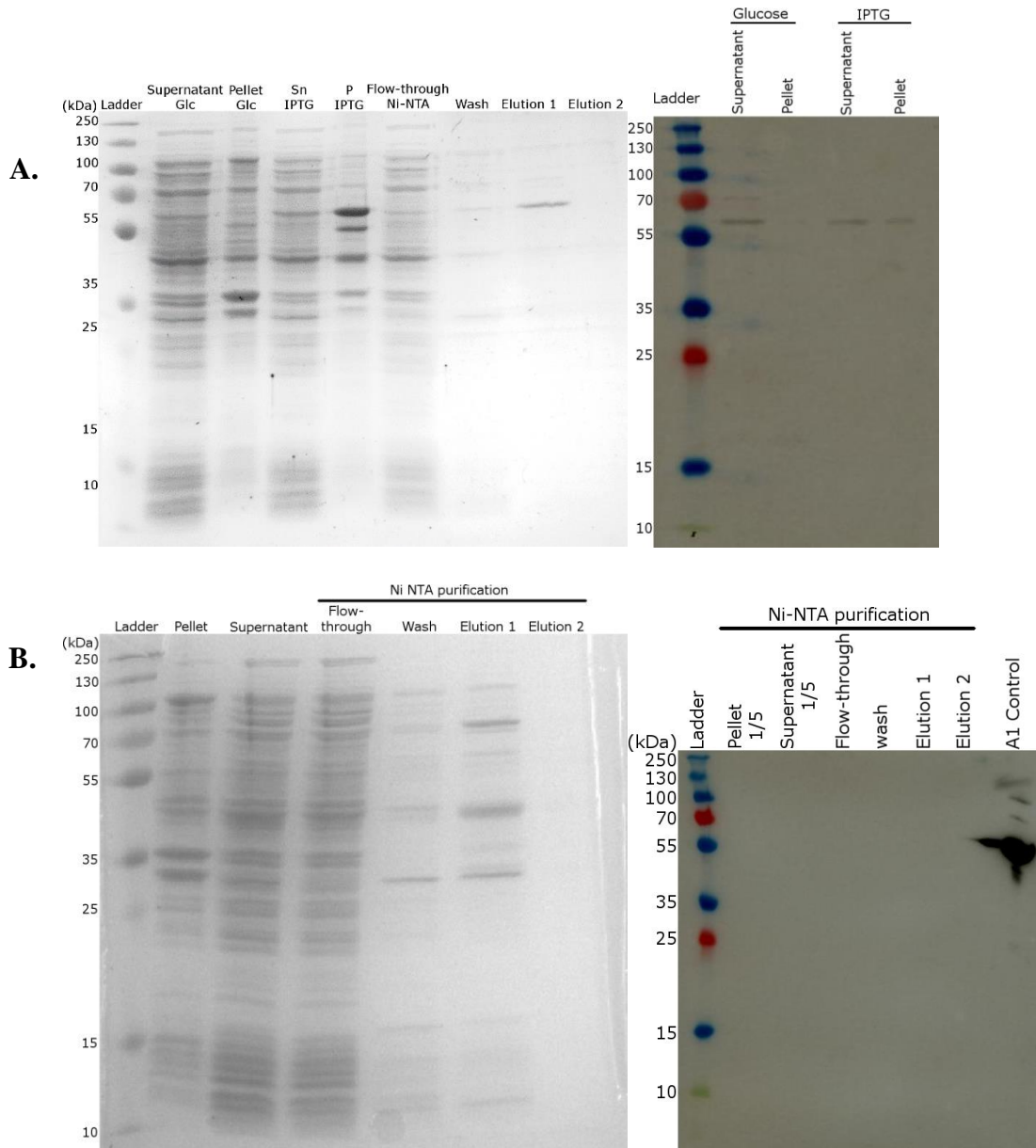
Afin de tester le rôle du domaine N-terminal dans la fonction de A1, j'ai fait un essai de production d'une version tronquée des 71 acides aminés du côté N-terminal. La construction a été faite dans un plasmide pET28 avec une étiquette Histidine et un site TEV fusionnés du côté N-terminal. J'ai utilisé des billes de Ni-NTA pour purifier rapidement la protéine et pouvoir plus facilement l'identifier. Les protéines obtenues ont été révélées par bleu de Coomassie et Western Blot avec un anticorps anti-A1. J'ai obtenu très peu de protéine soluble lors d'une production à 37°C (**Figure 24 A**). Une bande très fine à la masse attendue de 61 kDa est visible sur le gel SDS-PAGE pour la première élution. Une bande est également visible en Western Blot, mais elle apparaît après induction et après répression. La production est minimale : cette protéine tronquée est partiellement soluble et pourrait être purifiable sur colonne de Nickel. Elle pourrait être protéolysée rapidement au cours sa synthèse. J'ai aussi fait un essai de production à 16°C dans lequel je n'ai pas obtenu de protéine, bien que les anticorps aient bien reconnu le contrôle positif avec une protéine A1 entière (**Figure 24 B**). La construction d'un nouveau vecteur et/ou l'utilisation d'autres conditions de surexpression devra être envisagée. Cette question sera réexaminée plus loin dans ce manuscrit dans la discussion sur la structure de A1.





**Figure 23.** Analyse de l'état oligomérique de A1 en solution par SAXS. **A.** Densité optique à 280 nm en sortie de filtration sur gel. **B.** Analyse en cellule SAXS en sortie de filtration sur gel. **C.** Ajustement des données de SAXS moyennées en l'absence de  $MnCl_2$ , en utilisant les frames 2060 à 2100. Les points noirs représentent les points expérimentaux et la courbe bleue représente l'ajustement avec le programme Gasbor. **D.** Ajustement des données de SAXS moyennées en présence de  $MnCl_2$ , en utilisant les frames 1800 à 1880. Les points noirs représentent les points expérimentaux et la courbe bleue représente le fit avec le programme Gasbor (symétrie P42). **E.** Fonction de distribution des distances. **F.** Représentation de Kratky. **G.** Prédiction de désordre avec le serveur IUPred. Un score supérieur à 0,5 indique une région flexible. **H.** Enveloppe typique du monomère en solution obtenue à l'aide du programme Gasbor. Le fragment indiqué en vert a une longueur de 71 acides aminés. **I.** Enveloppe typique de l'octamère en solution obtenue à l'aide du programme Gasbor avec la symétrie P42.





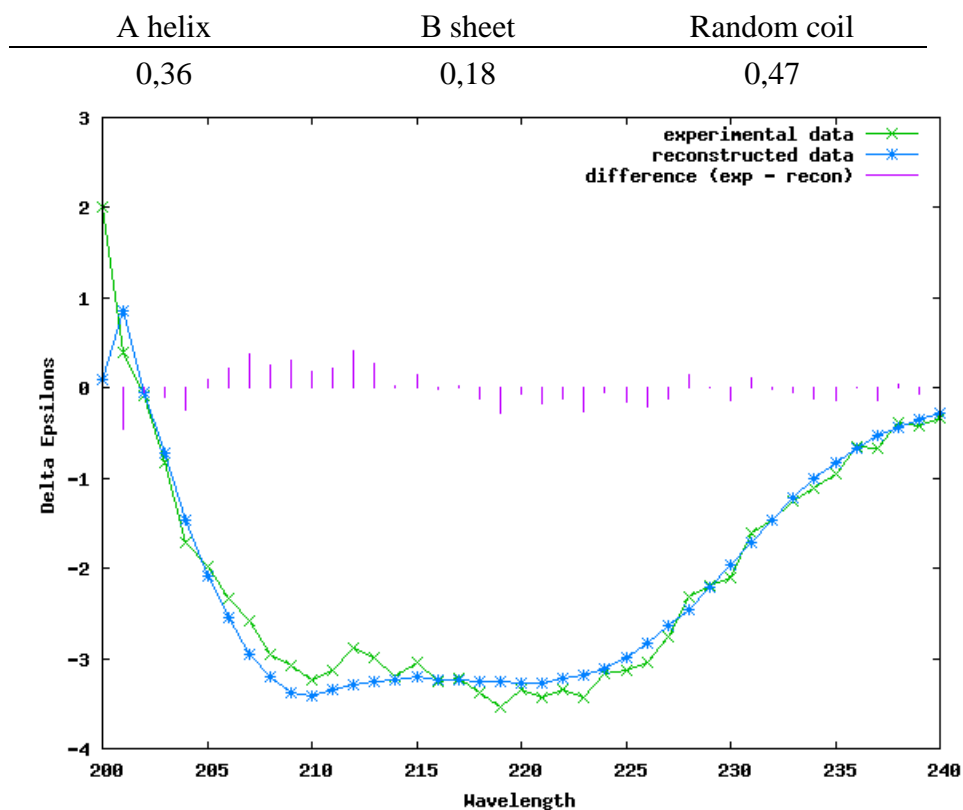
**Figure 24.** Essais de production d'une protéine A1 tronquée de 71 acides aminés en Nter avec une étiquette Histidine-TEV en Nter. Gels SDS-PAGE colorés au bleu de Coomassie (gauche) et Western Blot révélés avec des anticorps anti-A1 (droite) après productions et purifications faites respectivement à 37°C (**A**) et 16°C (**B**). La purification a été faite en solution avec des billes de Ni-NTA (Qiagen).

En conclusion, les expériences de SAXS nous ont permis de confirmer que A1 forme un octamère en présence de  $MnCl_2$  en solution.

#### 1.1.4 Analyse de la structure secondaire de A1 par Dichroïsme Circulaire

La méthode de dichroïsme circulaire (CD) dans l'Ultraviolet est couramment utilisée pour étudier le contenu en structures secondaires des protéines. Nous l'avons utilisée pour connaître le pourcentage d'hélices  $\alpha$  et de feuillets  $\beta$  chez A1 et avoir une courbe de référence pour le repliement de la protéine sauvage en solution. Ce spectre CD de référence est utile pour évaluer la qualité de repliement des mutants de A1 qui seront utilisés dans l'étude fonctionnelle de A1 (voir paragraphe 1.4.1). J'ai analysé les données sur le serveur dichroweb avec le logiciel K2D et sur le serveur Bestsel<sup>120,121</sup>, qui utilisent des algorithmes différents. Les données présentées correspondent à la protéine A1 + MnCl<sub>2</sub> avec un trajet optique de 0,01 mm. La protéine avait une concentration estimée à 4,013 mg/mL au Nanodrop.

Le logiciel K2D indique que A1 possède 36 % d'hélices  $\alpha$ , 18 % de feuillets  $\beta$  et 47 % de structures secondaires non canoniques (**Figure 25**). Le logiciel calcule une erreur maximale grâce à une méthode de "machine learning" basée sur des expériences dont le résultat est connu à l'avance. Cette erreur ne doit pas être supérieure à 0,227. La valeur obtenue pour cette expérience est de 0,080 et l'ajustement des données reconstruites aux données expérimentales est bon, les résultats sont donc considérés comme fiables.

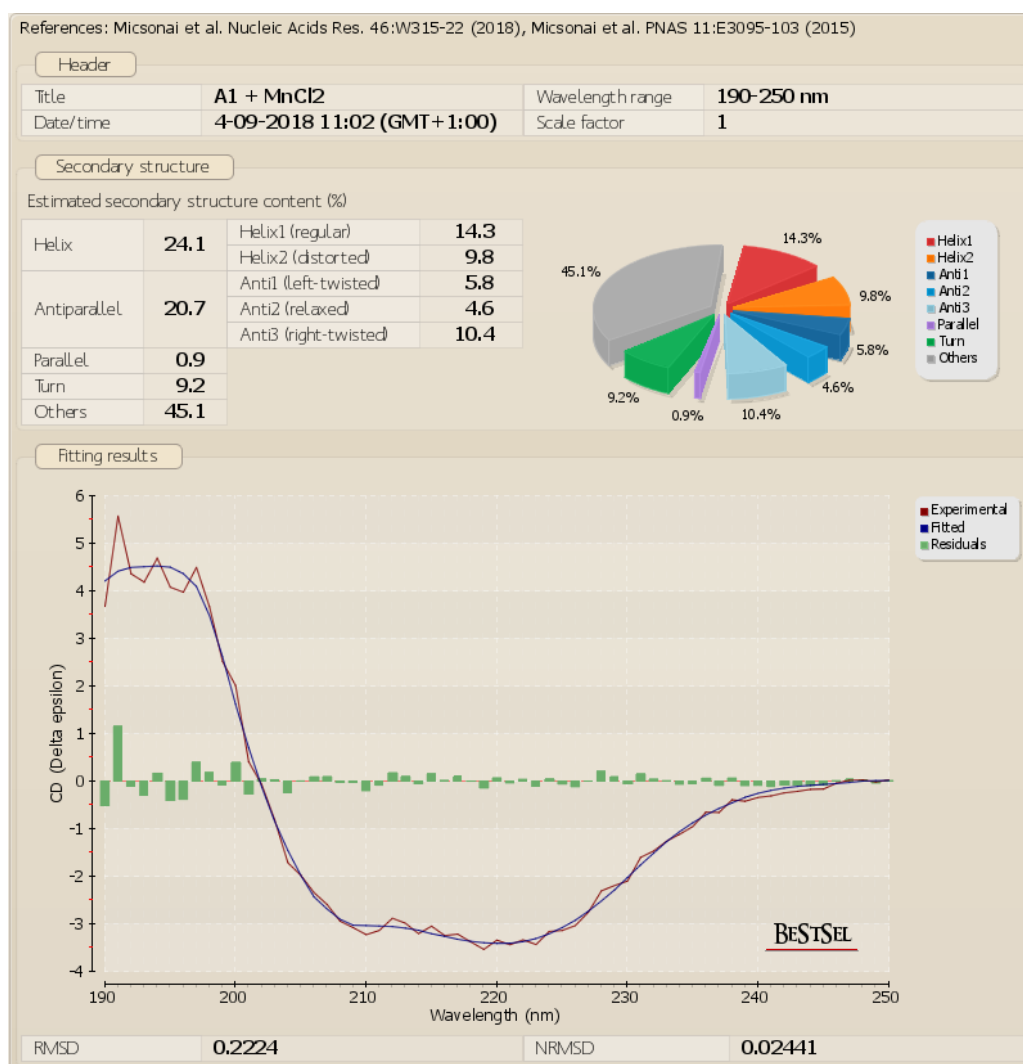


**Figure 25.** Analyse de la courbe de CD de A1 + MnCl<sub>2</sub> par le logiciel K2D.

Le logiciel BestSel indique que A1 possède 24,1 % d'hélices  $\alpha$ , 21,6 % de feuillets  $\beta$ , 9,2 % de coudes et 45,1 % d'autres structures (**Figure 26**). L'ajustement des données reconstruites aux données expérimentales est bon. Le calcul du RMSD (Root-Mean-Square Deviation) indique une distance moyenne entre les atomes du modèle reconstruit et du modèle expérimental. Nous obtenons une valeur de 0,2224 Å, ce qui est satisfaisant.

En comparant les résultats sur les deux serveurs, nous observons quelques différences. K2D indique plus d'hélices  $\alpha$  (36 %) que BestSel (24,1 %) et moins de feuillets  $\beta$  (18 % contre 21,6 %). Les deux serveurs prédisent que A1 possède plus d'hélices  $\alpha$  que de feuillets  $\beta$ .

Nous avons donc obtenu un spectre CD de A1 pour lequel les logiciels d'ajustement aux valeurs expérimentales semblent fiables. Nous devrions donc pouvoir comparer les spectres des protéines A1 mutées par rapport à cette référence pour évaluer le bon repliement de ces protéines modifiées.



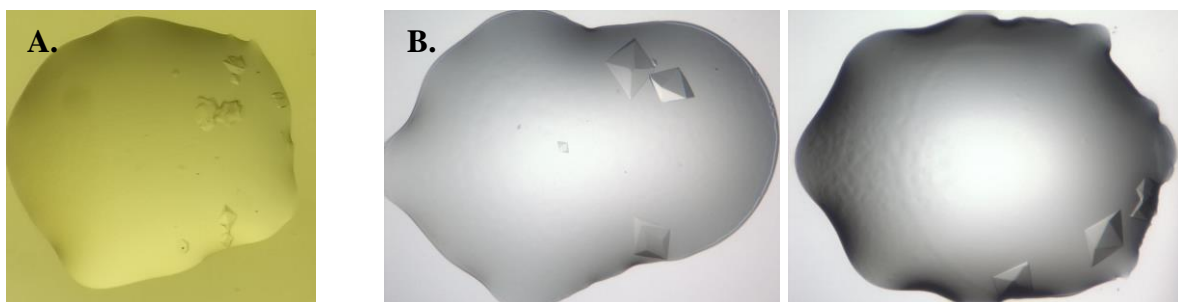
**Figure 26.** Analyse des courbes CD de A1 par BestSel.

### 1.1.5 Cristallographie aux rayons X

Notre but était de résoudre la structure 3D de A1. Dans ce but, nous avons mené en parallèle deux approches : la cristallographie et la cryo-microscopie électronique. Dans les deux cas, la protéine a été purifiée sans  $MnCl_2$ . Elle a ensuite été complexée au manganèse en étant incubée en présence de 1 mM de  $MnCl_2$  sur la nuit.

J'ai réussi à concentrer la protéine A1 jusqu'à 15 mg/mL pour faire des essais de cristallogénèse par la méthode de goutte assise. Pour trouver des conditions favorables, nous avons utilisé des kits de cristallisation qui permettent de tester 96 conditions différentes à la fois. Nous avons utilisé les kits Classics, PEGs, PEGs II et JSCG+ (QIAGEN).

Les premiers cristaux ont été obtenus en 5 jours d'incubation à 18°C, avec 0,1 M  $NaH_2PO_4$  ; 0,1 M MES pH 6,5 ; 2,0 M NaCl ; 0,1 M  $KH_2PO_4$  (**Figure 27 A**). Des optimisations ont été faites pour obtenir des mono cristaux de plus grande taille ayant les meilleures chances de diffracter. Nous avons testé des tampons sans phosphate car c'est un composé qui peut former des cristaux. Nous avons essayé de produire des cristaux à 6°C et nous avons fait varier le pH et la concentration de sel. Enfin, nous avons essayé d'obtenir des cristaux en présence d'un cryoprotectant (Ethylène glycol 25 % ou Glycérol 25 %) pour éviter le stress de changer le tampon au moment de congeler les cristaux, ce qui peut provoquer des dégâts dans les cristaux (fissures, dissolution) de protéine, ce qui peut se traduire par une perte du pouvoir diffractant. Les meilleurs résultats ont été obtenus en 5 jours d'incubation à 18°C avec 0,1 M  $NaH_2PO_4$  ; 0,1 M MES pH 6,5 ; 1,8 M NaCl ; 0,1 M  $KH_2PO_4$ , puis congélation en présence d'Ethylène Glycol 25 % (**Figure 27 B**).



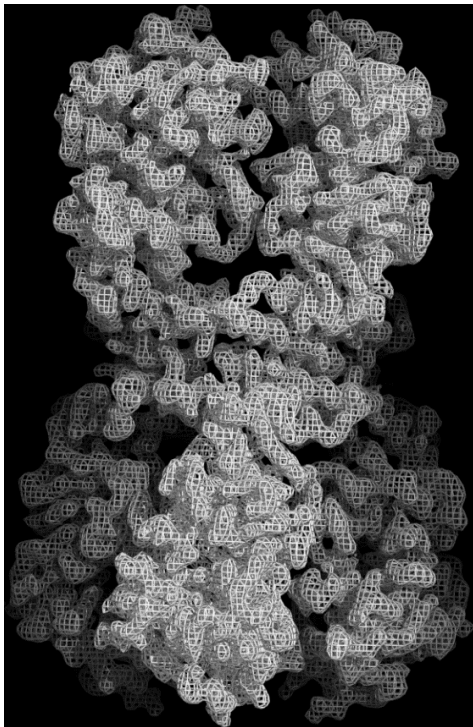
**Figure 27. Crystallisation de la protéine A1. A.** Premiers cristaux obtenus avant optimisation. **B.** Exemples de cristaux obtenus pour lesquels nous avons observé de la diffraction au synchrotron.

Les cristaux ont été analysés sur la ligne PROXIMA-1 au synchrotron SOLEIL et la structure a été résolue par la méthode SAD. Plusieurs jeux ont été collectés sur la ligne PROXIMA-1 et PROXIMA2 (synchrotron SOLEIL) et ID23-1 (synchrotron ESRF). Le meilleur jeu natif a une résolution de 3,3 Å. L'affinement de la structure avec ce dernier jeu est en cours. La protéine

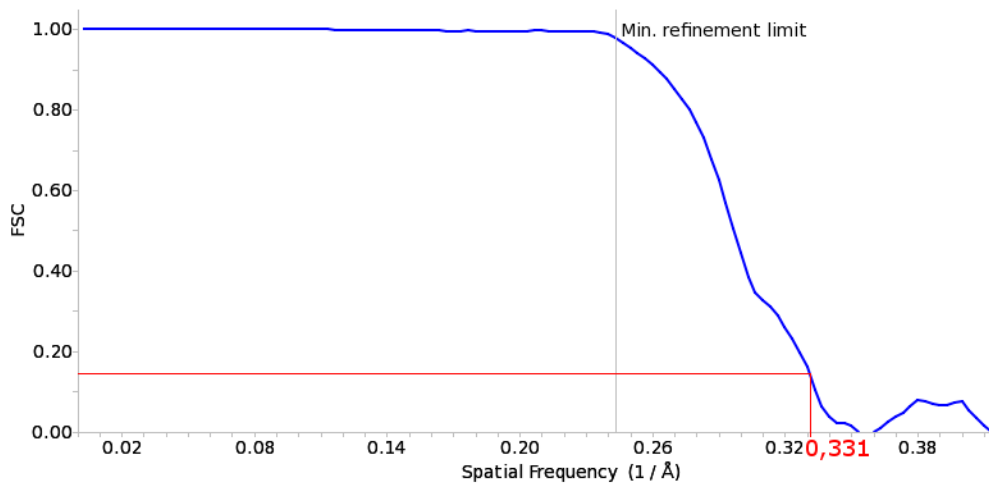
crystallise dans le groupe d'espace P6(1)22, avec 4 molécules de A1 par unité asymétrique. Le pourcentage de solvant des cristaux est de 74,8 %, ce qui est une valeur particulièrement élevée. Les paramètres de maille sont  $a = b = 230,3 \text{ \AA}$  ;  $c = 313,9 \text{ \AA}$  et  $\alpha = \beta = 90^\circ$  ;  $\gamma = 120^\circ$ . Les statistiques du traitement de données sont récapitulées dans la partie Matériel et Méthodes (Page 141, **Tableau 8**). Les valeurs obtenues indiquent que les données sont fiables, en particulier le  $CC_{1/2}$  est proche de 50 % et le rapport signal/bruit  $I/I\sigma$  est supérieur à 1 dans la dernière tranche de résolution pour les deux jeux de données : SAD et natif. Les données anomales ont été collectées à une longueur d'onde égale au seuil de la diffusion anormale des ions  $Mn^{2+}$  car nous avons incubé la protéine en présence de 1 mM  $MnCl_2$  avant les essais de cristallogénèse, ce qui provoque l'oligomérisation de la protéine comme nous l'avons observé par SAXS. Nous pouvions donc prédire que A1 avait interagi avec les ions  $Mn^{2+}$  qui pouvaient être encore présents dans la structure cristalline. Le caractère anomal de cet atome a permis déterminer les phases pour résoudre la structure de la protéine. L'analyse par le programme SHELX a permis de trouver 4 positions par unité asymétrique pour les ions  $Mn^{2+}$ . La reconstruction a permis de placer deux ions  $Mn^{2+}$  par site car la carte de densité électronique n'était pas compatible avec la présence d'un seul ion par site. Le modèle indique que A1 forme un octamère avec une symétrie cristallographique.

### 1.1.6 Cryo-microscopie électronique de la protéine A1

En parallèle, l'analyse des images obtenues par cryo-microscopie électronique a rapidement permis de mettre en évidence la formation d'un octamère par A1. Cet octamère respecte une symétrie locale D4 : c'est un dièdre avec une rotation d'ordre 4, formé d'un axe 4 et d'un axe 2 perpendiculaires. La symétrie 4 engendre un axe 2 perpendiculaire aux deux autres. La carte de densité moyennée permet de voir l'octamère formé par A1 (**Figure 28**). Un graphique de corrélation de Fourier (Fourier shell correlation, FSC) permet d'estimer la limite de résolution de la carte de densité électronique (**Figure 29**). En utilisant un critère de limite de résolution à un score de FSC de 0,143, on trouve une limite de résolution de  $1/0,331 = 3,0 \text{ \AA}$ .



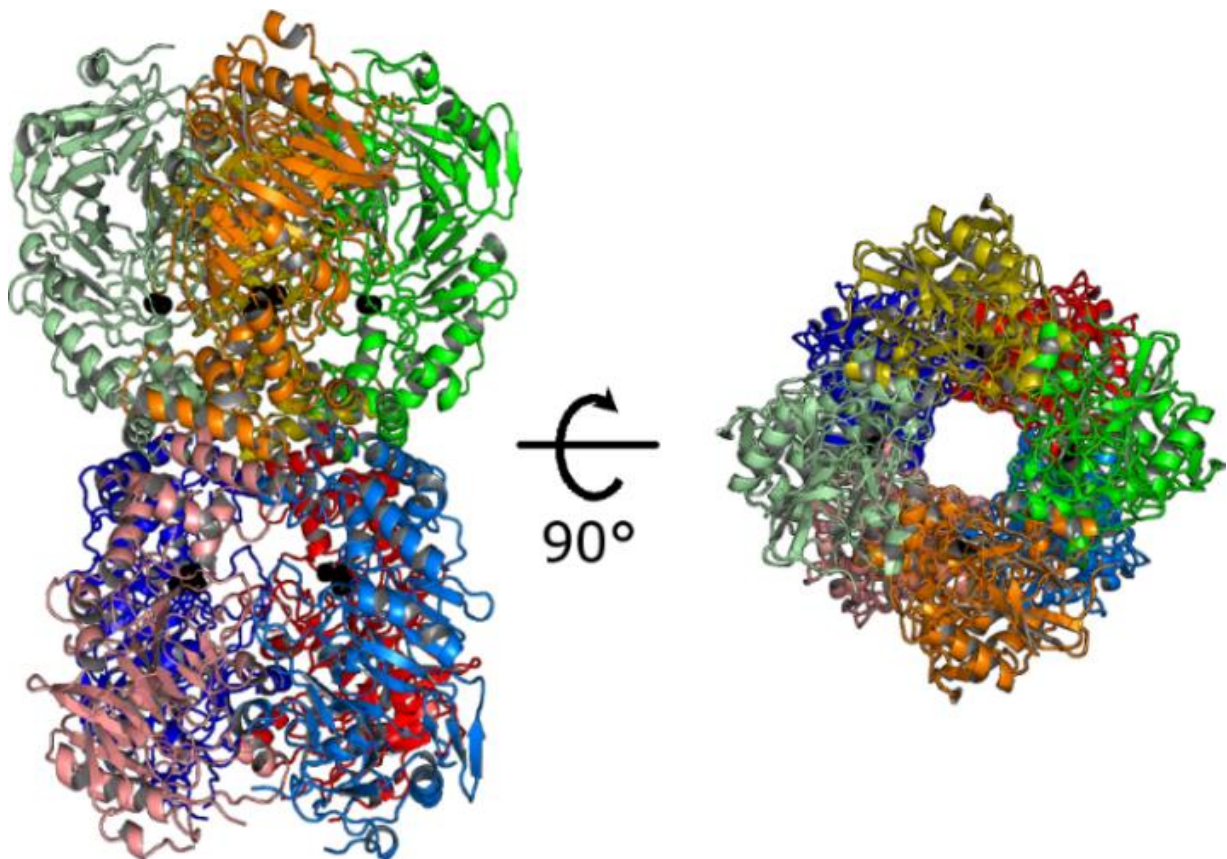
**Figure 28.** Carte de densité électronique moyennée de l'octamère de la protéine A1 obtenue par analyse d'images de cryo-microscopie électronique. Images construites avec le logiciel COOT, à un volume de densité électronique de  $3,0 \text{ \AA}$ . La symétrie D4 est imposée.



**Figure 29.** Fourier shell correlation (FSC) du modèle de A1 obtenu par microscopie électronique. En rouge, limite de résolution (en  $\text{\AA}^{-1}$ ) à un score de FSC de 0,143.

La carte de densité électronique moyennée a été utilisée pour tracer manuellement un modèle initial de A1 à l'aide du logiciel Coot. Les données de ce modèle sont présentées dans le **Tableau 2**. Le logiciel Phenix Real Space Refine a ensuite été utilisé pour optimiser l'accord entre le modèle et la carte de densité tout en tenant compte des contraintes géométriques imposées par les liaisons chimiques. Le score Molprobity et l'écart-type comparant les longueurs et les angles dans le modèle aux valeurs idéales sont suffisamment faibles. Par ailleurs, moins de 1 % des valeurs associées au diagramme de Ramachandran sont aberrantes : notre modèle semble donc fiable et présente une bonne approximation de la réalité. Ce modèle contient un octamère de A1, ainsi que 16 ions  $Mn^{2+}$ . A l'exception des 89 résidus N-terminaux et de la boucle 139-149 qui sont probablement flexibles, tous les résidus de A1 sont présents dans ce modèle.

Cette structure caractéristique formée de 2 tétramères avec une rotation de  $45^\circ$  au niveau de l'axe 2 est visible sur le modèle obtenu (**Figure 30**).



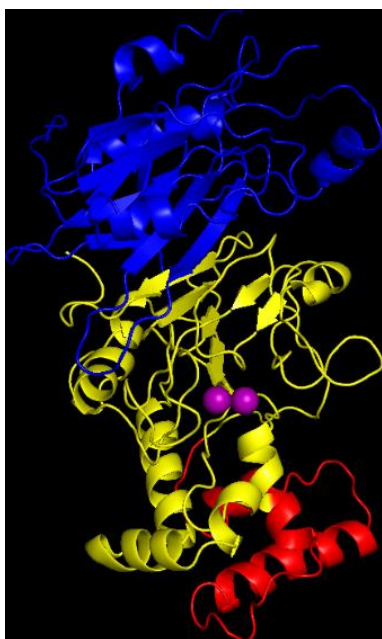
**Figure 30.** Modèle reconstruit de la structure de l'octamère de A1.

<b>Nombre de particules pour la reconstruction cisTEM</b>	59528 (3,0 Å)
<b>Résolution</b>	3,0 Å
<b>Nombre d'atomes du modèle</b>	28608
Protéine	28592 (3656 par chaîne)
Ligands (manganèse)	16
<b>Score Molprobit</b>	2,29
<b>Liaisons chimiques (écart-type)</b>	
Longueur	0,010 Å
Angles	1,287°
<b>Ramachandran</b>	
Permis	92,05%
Toléré	7,73%
Aberrant	0,22%

**Tableau 2. Données du modèle de cryomicroscopie électronique.** Le score Molprobit dépend de la qualité du modèle reconstruit. Ce score augmente si des incohérences géométriques sont présentes dans le modèle (rotamères, superposition des atomes, Ramachandran). Le RMSD des liaisons chimiques indique l'écart moyen observé entre les longueurs et les angles des liaisons chimiques du modèle et les valeurs idéales. Les valeurs associées au diagramme de Ramachandran permettent d'analyser la conformation du squelette polypeptidique de la protéine et d'évaluer si le modèle utilise des valeurs possibles du point de vue énergétique.

#### 1.1.6.1 Structure du monomère

Le monomère de A1 est composé de 3 domaines, indiqués par des couleurs différentes : en bleu le domaine N-terminal (89-295), en jaune le domaine actif (296-401 456-556) et en rouge un domaine impliqué dans la formation de l'octamère (402-455) (**Figure 31**).



**Figure 31. Structure du monomère de A1 obtenue par cryo-microscopie électronique.**

Le domaine N-terminal manquant composé de 89 résidus correspond probablement au petit domaine identifié en SAXS (voir **Figure 23** H) car le modèle obtenu en cryo-microscopie s'intègre convenablement dans le grand domaine identifié en SAXS. Par ailleurs, la dimension de la structure obtenue est de 140 Å pour un Dmax de 200 Å en SAXS : nous pouvons en déduire que les domaines N-terminaux placés aux deux extrémités de l'octamère sont libres de



se mouvoir à l'extérieur de celui-ci, ce qui expliquerait qu'ils ne soient pas visibles en microscopie.

Il y a 20,8 % d'hélices  $\alpha$  et 19,3 % de feuillets  $\beta$  dans cette structure. Ces valeurs sont très proches des résultats obtenus sous BestSel qui prédisait 24,1 % d'hélices  $\alpha$  et 21,6 % de feuillets  $\beta$  (voir **Figure 26**).

#### 1.1.6.2 Domaine 402-455

Le domaine 402-455 (en rouge dans la Figure 31) est inséré dans le domaine DNase (en jaune dans la Figure 31). Il est composé de deux hélices et joue un rôle crucial dans la formation de l'octamère.



**Figure 32.** Structure du domaine 402-455 de A1 qui joue un rôle dans l'oligomérisation.

#### 1.1.6.3 Domaine 89-295 et Domaine actif (296-401 456-556)

Les prédictions bioinformatiques à partir de la séquence de A1 n'ont pas permis d'identifier le rôle du domaine 89-295. Cependant, la détermination de la structure de A1 donne des indices sur son origine et sa fonction.

Le domaine actif coordonne les ions manganèse responsables de l'activité DNase. Il se situe dans la moitié C-terminale de la protéine.



**Figure 33.** Structure du domaine 89-295 de A1.



**Figure 34.** Structure du domaine actif de A1. Les ions manganèse sont représentés en violet.

En utilisant le serveur DaliSearch qui compare des structures de protéines, nous avons pu comparer le domaine 89-295 et le domaine actif avec des protéines connues qui ne peuvent pas être trouvées à l'aide de la séquence primaire seule. Le Z-score est indicatif du nombre de différences entre deux structures. On considère qu'un Z-score supérieur à 4,0 est suffisant pour conclure à une similarité. Chacun de ces domaines est homologue aux DNases d'une même famille comprenant notamment le domaine DNase de la DNA pol II, SbcD ou Mre11 (**Tableau 3** et **Tableau 4**). La faible identité de séquence observée est due à une grande divergence des séquences en acides aminés.

Chain	Z	rmsd	lali	nres	%id	Description
5ihe-A	8.2	3.5	150	440	12	DNA pol II small subunit
3qg5-C	7.0	3.4	130	342	12	Complexe Mre11/Rad50
4lu9-D	7	3.5	133	344	14	Exonuclease subunit SbcD
4yke-A	6.2	3.6	129	404	11	Mre11

**Tableau 3. Alignement du domaine 89-295 de A1 obtenu avec le serveur DaliSearch.** Z : score indicatif du nombre de différences entre les deux structures. Rmsd, écart-type entre les positions des carbones  $\alpha$  entre les deux structures. Lali, nombre de résidus de la cible qui ont pu être alignés avec le modèle. Nres, nombre d'acides aminés du modèle cible. %id, identité de séquence.

Chain	Z	rmsd	lali	nres	%id	Description
5ihe-A	14.2	2.7	178	440	16	DNA pol II small subunit
3qg5-C	12.3	3.1	158	362	13	Complexe Mre11/Rad50
4lu9-B	11.5	3.2	160	342	14	Exonuclease subunit SbcD
4yke-A	9.9	3.3	160	404	12	Mre11

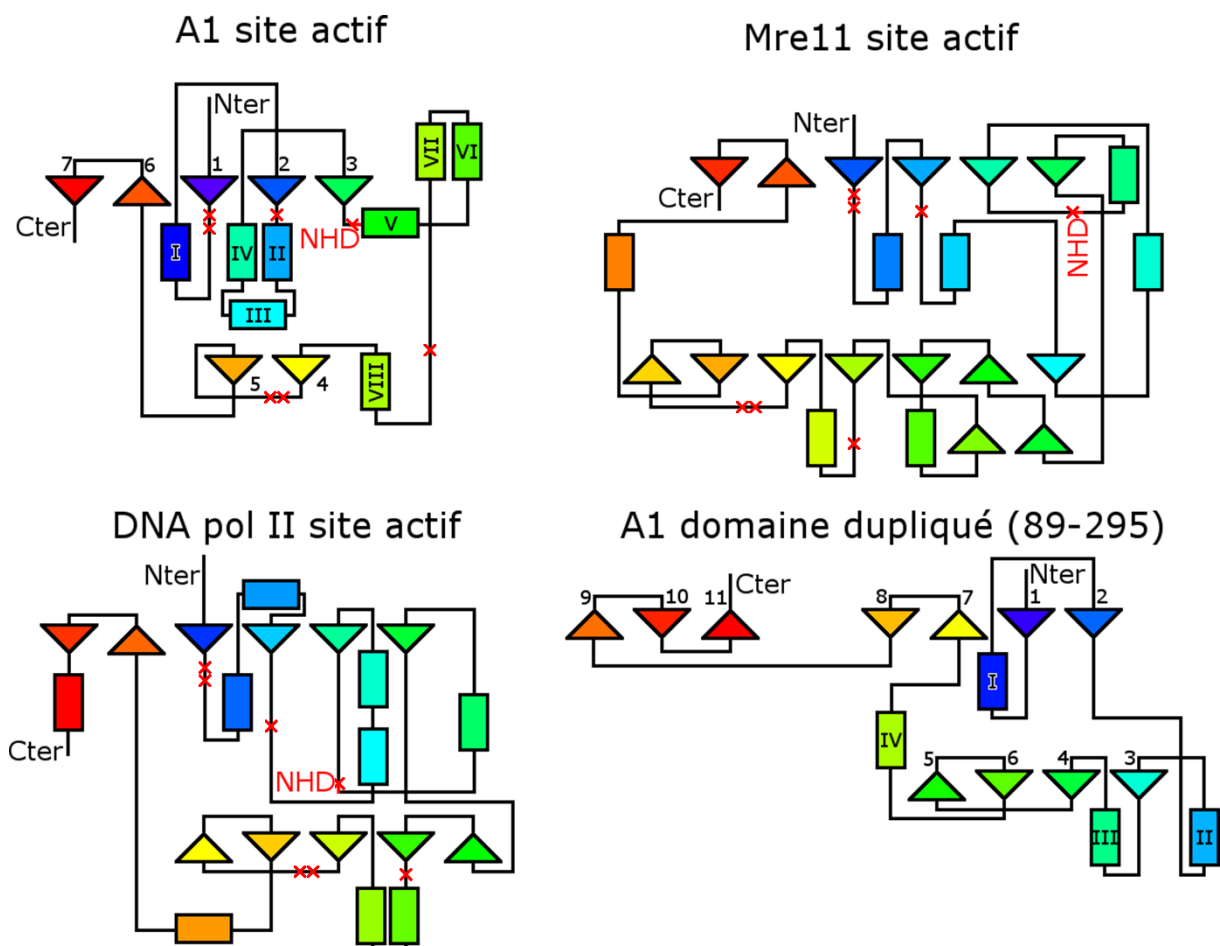
**Tableau 4. Alignement du domaine actif de A1 obtenu avec le serveur DaliSearch.** Voir Tableau 3 pour la signification des différentes colonnes.

Il serait donc possible que le domaine 89-295 soit une duplication du domaine actif. Pour le vérifier, nous avons comparé ces deux domaines avec le serveur DaliSearch (**Tableau 5**). Bien que le pourcentage d'identité de séquence soit faible, le Z-score et le rmsd indiquent qu'ils sont homologues, ce qui confirme que ces domaines proviennent d'une duplication d'un domaine ancestral.

Z	Rmsd	lali	nres	%id
10.3	2.9	136	195	9

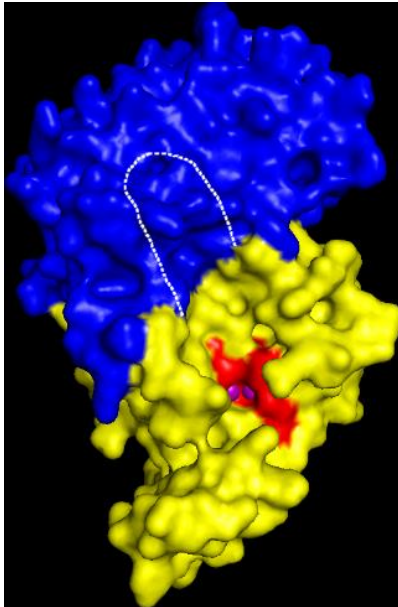
**Tableau 5. Alignement du domaine actif de A1 avec le domaine 89-295, obtenu avec le serveur DaliSearch.** Voir Tableau 3 pour la signification des différentes colonnes.

Afin de pouvoir comparer les structures secondaires de A1 avec les protéines homologues Mre11 et DNA pol II, j'ai construit des schémas représentant l'organisation de leurs structures secondaires (**Figure 35**). Le domaine actif de A1 comme les domaines DNase de Mre11 et de la DNA pol II s'organisent autour de deux feuillets  $\beta$  et de plusieurs hélices. Le domaine actif de A1 apparait comme une version simplifiée de ce domaine. Le premier feuillet  $\beta$  (brin 7, 6, 1, 2, 3) est conservé chez ces trois protéines, cependant Mre11 et la DNA pol II possèdent un brin supplémentaire. De même le second feuillet  $\beta$  ne possède que deux brins chez A1 alors qu'il en a respectivement 7 et 5 chez Mre11 et DNA pol II. Dans ces trois protéines on retrouve à l'extrémité C-terminale du brin 3 le motif fortement conservé dans cette famille NHD. Dans le domaine dupliqué de A1 (89-295), c'est le domaine composé des feuillets  $\beta$  3 à 6 et de l'hélice  $\alpha$  III qui est présent dans les sites actifs de Mre11 et de la DNA pol II. En revanche, le brin 3 du site actif est absent de même que le site NHD.



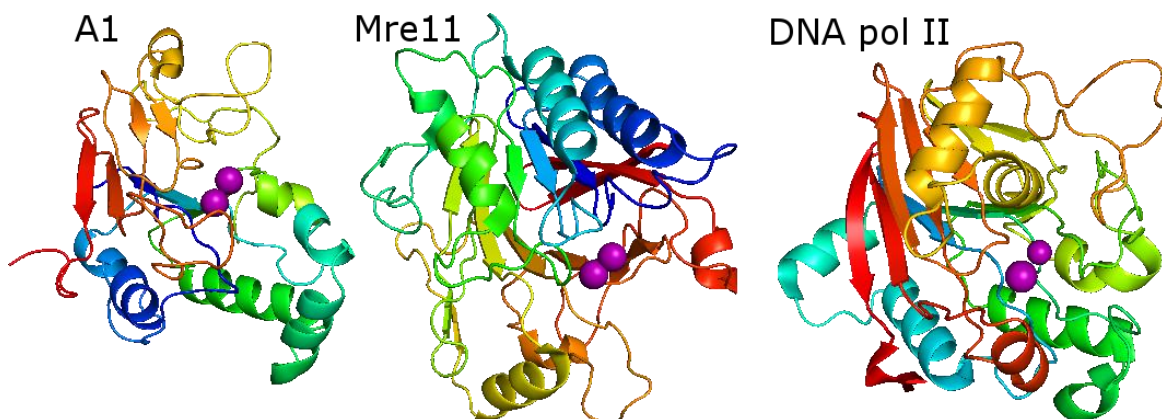
**Figure 35.** Comparaison des structures secondaires du domaine actif de A1 et du domaine N-terminal de A1 avec les domaines DNase de Mre11 (4YKE) et de la DNA pol II (5IHE). NHD : acides aminés très conservés du site actif et impliqués dans l'activité DNase de A1. Les triangles représentent des feuillets  $\beta$  et les rectangles des hélices  $\alpha$ . Les croix rouges représentent les acides aminés responsables de la coordination des cations divalents.

J'ai comparé la structure du domaine dupliqué avec celle du domaine actif de A1 (**Figure 36**). Les éléments caractéristiques de l'activité sont absents dans le domaine dupliqué : il n'y a pas de crevasse permettant d'accueillir l'ADN comme indiqué par les pointillés blancs, ni d'ion manganèse. Par ailleurs les résidus impliqués dans la coordination des ions manganèse fortement conservés dans cette famille sont tous absents.



**Figure 36. Comparaison des structures du domaine dupliqué inactif (en bleu) et du domaine actif de A1 (en jaune).** En rouge, acides aminés responsables de la coordination des ions manganèse. En violet, ions manganèse. Les pointillés blancs représentent la crevasse manquante dans le domaine dupliqué inactif.

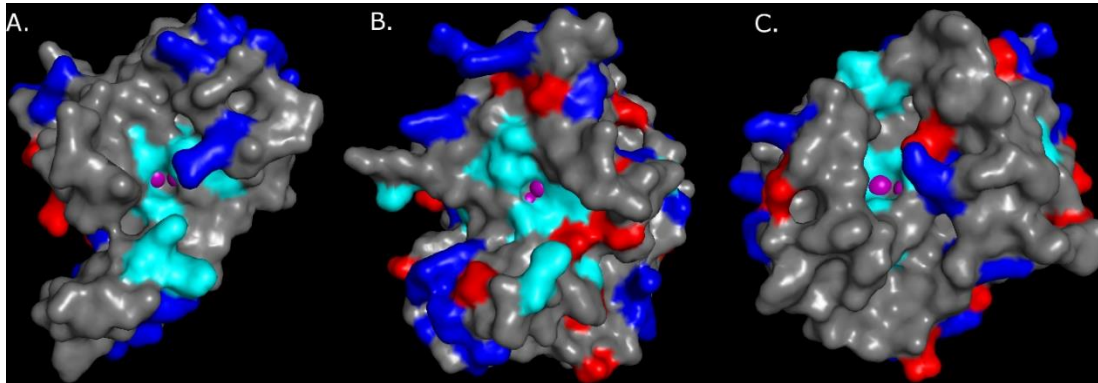
Le logiciel pymol a été utilisé pour faire des alignements entre les différents domaines de A1 et les domaines actifs de Mre11 et de la DNA pol II (**Figure 37**). Le RMS est un score de l'alignement indiquant la distance moyenne entre les atomes des protéines comparées. Il est considéré comme étant très significatif lorsqu'il est inférieur à 5. Les sites actifs de chaque protéine ont une structure similaire, malgré une homologie de séquence inférieure à 10 %.



**Figure 37. Alignement du site actif de A1 avec ceux de Mre11 et de la DNA pol II.**

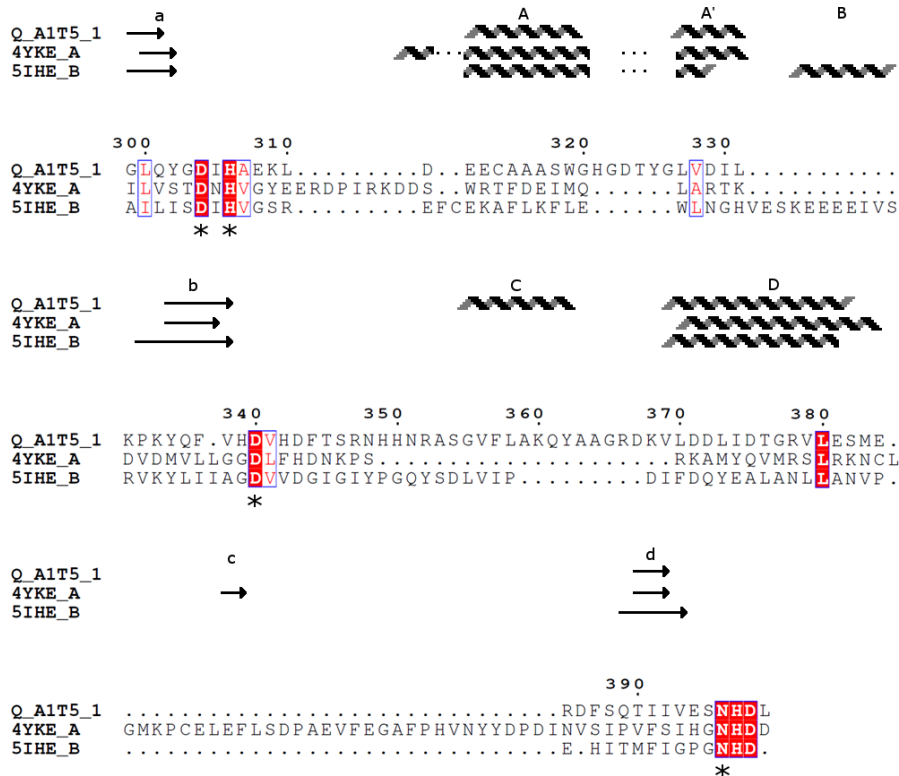
- Site actif A1 / Site actif Mre 11, RMS = 4,063.
- Site actif A1 / Site actif DNA pol II, RMS = 5,623.
- Site actif A1 / domaine dupliqué, RMS = 14,283 (non représenté).

La crevasse permettant d'accueillir l'ADN est organisée de façon similaire chez A1, Mre11 et la DNA pol II (**Figure 38**). Les charges sont distribuées de la même façon autour des sites actifs de ces protéines avec majoritairement des histidines autour des ions manganèse qui doivent participer au positionnement du substrat.



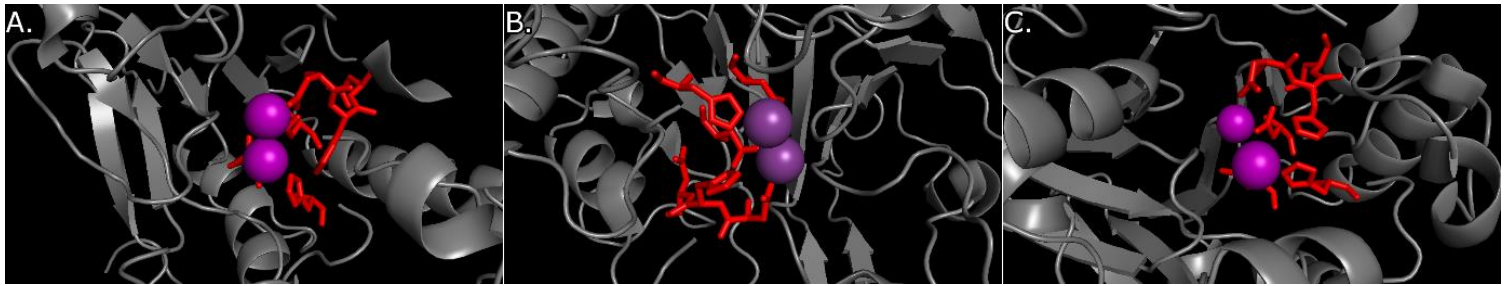
**Figure 38.** Distribution des charges autour du site actif de A1 (A), de Mre11 (B) et de la DNA pol II (C). En rouge, acides aminés acides (Glutamates et Aspartates). En bleu foncé, acides aminés basiques (Lysines et Arginines). En bleu ciel, Histidines.

Un alignement des sites actifs de A1, de Mre11 (4YKE) et de la DNA pol II (5IHE) obtenu par HHPred montre que les structures secondaires sont conservées malgré des séquences très différentes (**Figure 39**).



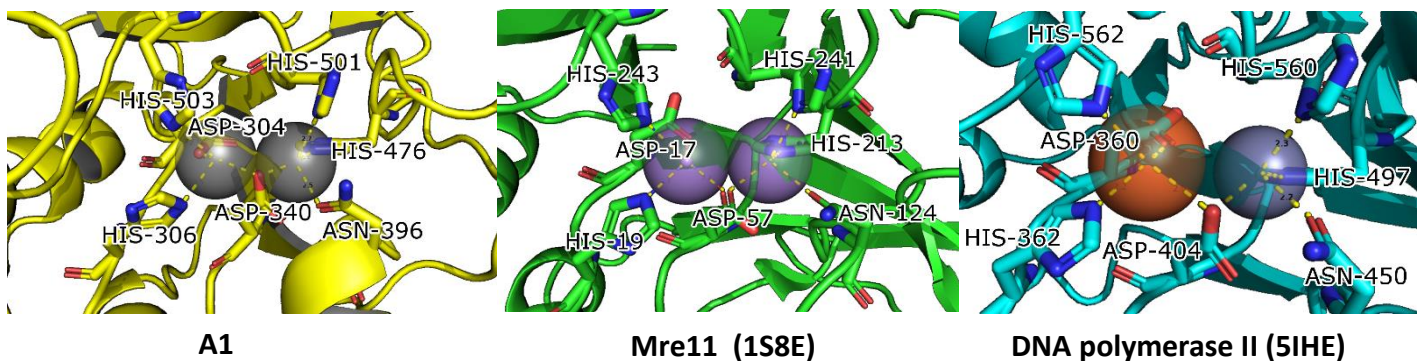
**Figure 39.** Alignement des domaines actifs de A1, de Mre11 (4YKE) et de la DNA pol II (5IHE) obtenu par HHPred. Les résidus coordonnant les ions manganèses sont indiqués par un astérisque. Ils sont identiques chez les trois protéines. Ils ne sont pas présents au-delà du résidu 400 car les séquences sont alors trop divergentes pour être alignées.

Lorsqu'on met en évidence les acides aminés conservés d'après l'alignement, on remarque une organisation similaire pour les trois sites actifs (**Figure 40**).



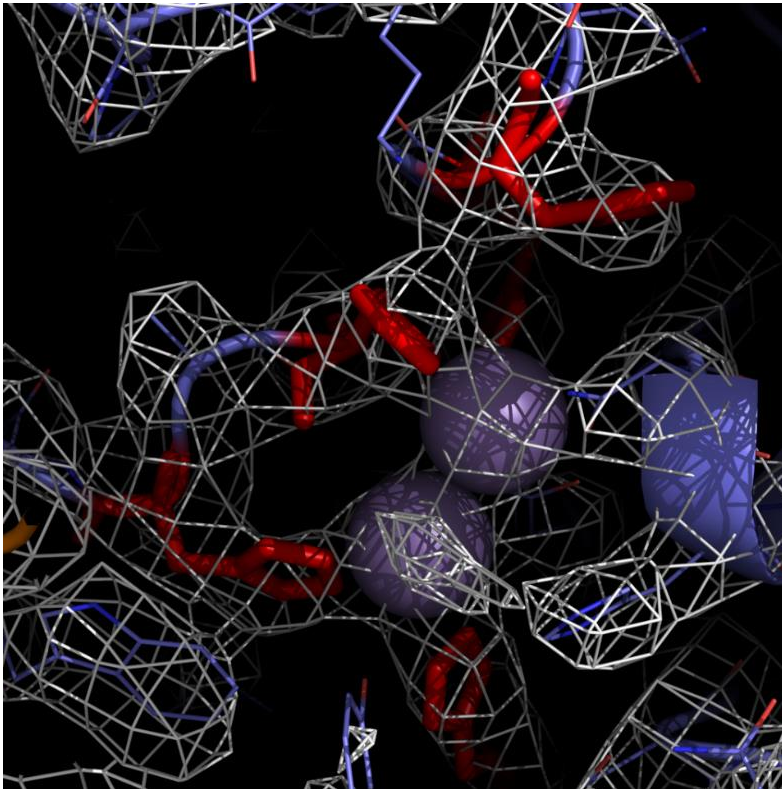
**Figure 40.** Représentation des acides aminés conservés d'après un alignement HHpred dans les sites actifs de A1 (A), de Mre11 (B) et de la DNA pol II (C).

Les sept acides aminés strictement conservés qui coordonnent les ions  $Mn^{2+}$  chez A1 et Mre11 et qui coordonnent les ions  $Fe^{2+}$  et  $Zn^{2+}$  chez la DNA pol II sont organisés de façon identique dans ces trois protéines (**Figure 41**). Cette forte similarité structurale entre le domaine contenant le site actif du monomère de A1 et les domaines exonucléase de la DNA pol II et Mre11 suggèrent que A1 est effectivement une nucléase et qu'elle pourrait être la DNase responsable de la dégradation de l'ADN bactérien observée immédiatement après la première étape de transfert du génome dans la bactérie.



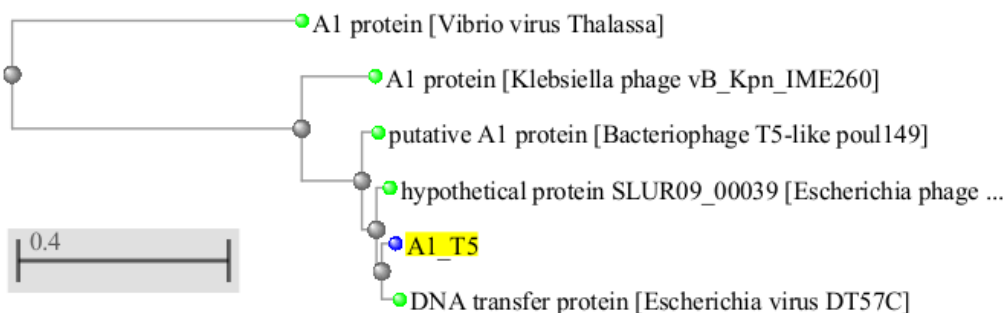
**Figure 41.** Coordination des ions  $Mn^{2+}$  chez A1, Mre11 et des ions  $Fe^{2+}$  et  $Zn^{2+}$  chez la DNA pol II.

Les histidines liant les ions  $Mn^{2+}$  sont visibles dans la carte de densité électronique. Une densité est observable entre ces acides aminés : elle correspond aux ions  $Mn^{2+}$  (**Figure 42**).



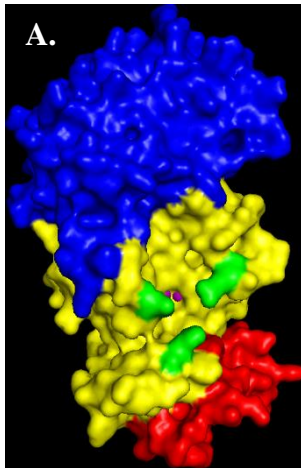
**Figure 42.** Carte de densité électronique montrant les histidines (en rouge) liant les ions  $Mn^{2+}$ .

J'ai construit un arbre phylogénétique avec des protéines A1 de phages de la famille T5 (*T5virus*) sélectionnées avec le serveur Blast (**Figure 43**). Un phage est considéré comme étant un *T5virus* s'il possède les protéines A1 et A2. Aucun phage avec un haut taux de similarité pour T5 sans ces protéines n'a encore été découvert. Le *T5virus* possédant la protéine A1 la plus distante de celle de T5 est le phage Thalassa qui infecte *Vibrio harveyi*. L'identité de séquence de la protéine A1 de ce phage avec celle de T5 n'est que de 31%. La structure et la fonction de A1 sont probablement conservées chez tous les *T5virus*, sans que leurs séquences soient pour autant très conservées.

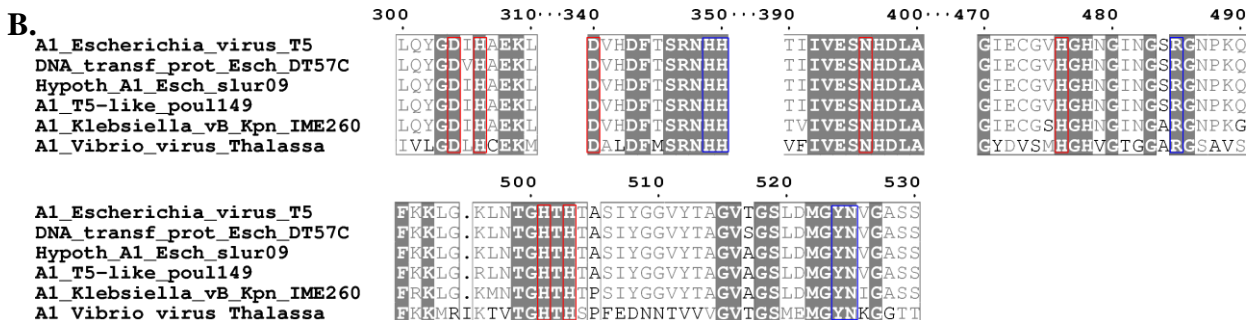


**Figure 43.** Arbre phylogénétique de protéines A1 chez différents phages sélectionnés de la famille T5. Arbre construit sur le serveur BLAST. L'échelle des distances génétiques est arbitraire.

Les acides aminés H350, R485 et N525 sont conservés chez Mre11 et la DNA pol II. Chez A1, ils forment 3 doigts qui positionnent l'ADN (**Figure 44 A**). Ces acides aminés sont conservés chez l'ensemble des phages *T5viruses* testés (encadrés en bleu, **Figure 44 B**). Les acides aminés Y524 et H349 qui sont conservés chez les A1-like et au voisinage du site actif dans la crevasse sont aussi conservés chez les *T5viruses* (encadrés en bleu, **Figure 44 B**). Les acides aminés responsables de la coordination du manganèse sont tous conservés (encadrés en rouge, **Figure 44 B**).



**Figure 44. Conservation des acides aminés prédits pour positionner l'ADN au niveau du site actif.** **A.** Mise en évidence de ces acides aminés, en vert dans la structure. **B.** Alignement de la protéine A1 de T5 et de différents phages *T5viruses* par BLAST.



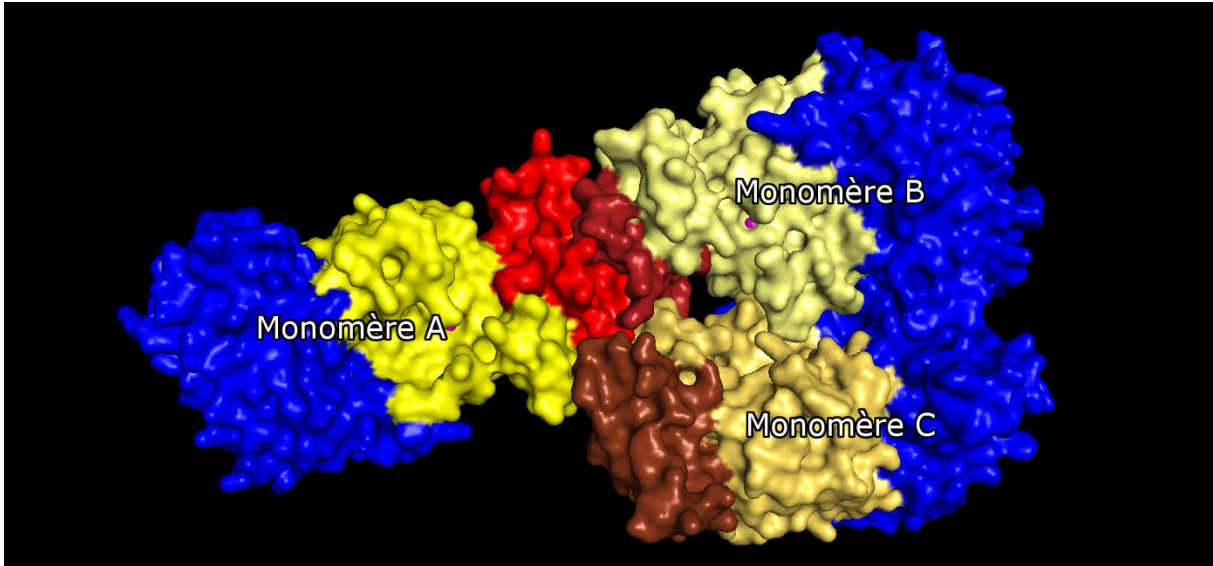
L'ensemble de ces résultats montre que le monomère de A1 a une organisation très conservée malgré une faible homologie de séquence entre les sites actifs des DNases similaires. Les sites actifs de ces DNases présentent les mêmes structures secondaires, une coordination similaire des cations divalents et une conservation des acides aminés probablement responsables de l'activité. Les acides aminés potentiellement responsables du positionnement de l'ADN dans le site actif et de la coordination des ions manganèse sont conservés par rapport à des protéines A1 d'autres phages *T5virus*.



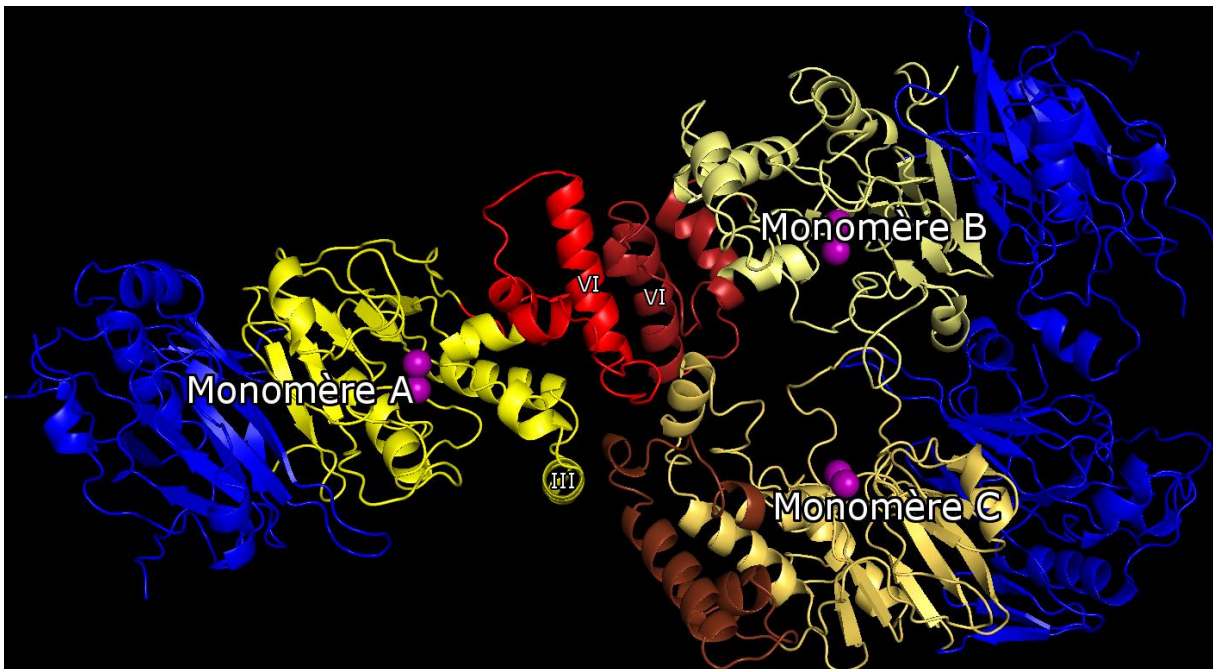
#### 1.1.6.4 Analyse des domaines d'interaction entre les monomères de A1

Au niveau de l'axe de symétrie 2 de l'octamère qui présente une symétrie D<sub>4</sub>, 3 monomères interagissent entre eux (**Figure 45**). Dans le domaine de dimérisation, l'hélice VI du monomère A interagit avec l'hélice VI du monomère B (**Figure 46 A**). L'hélice  $\alpha$  III du site actif du monomère A interagit avec le domaine de dimérisation du monomère C (**Figure 46 B**).

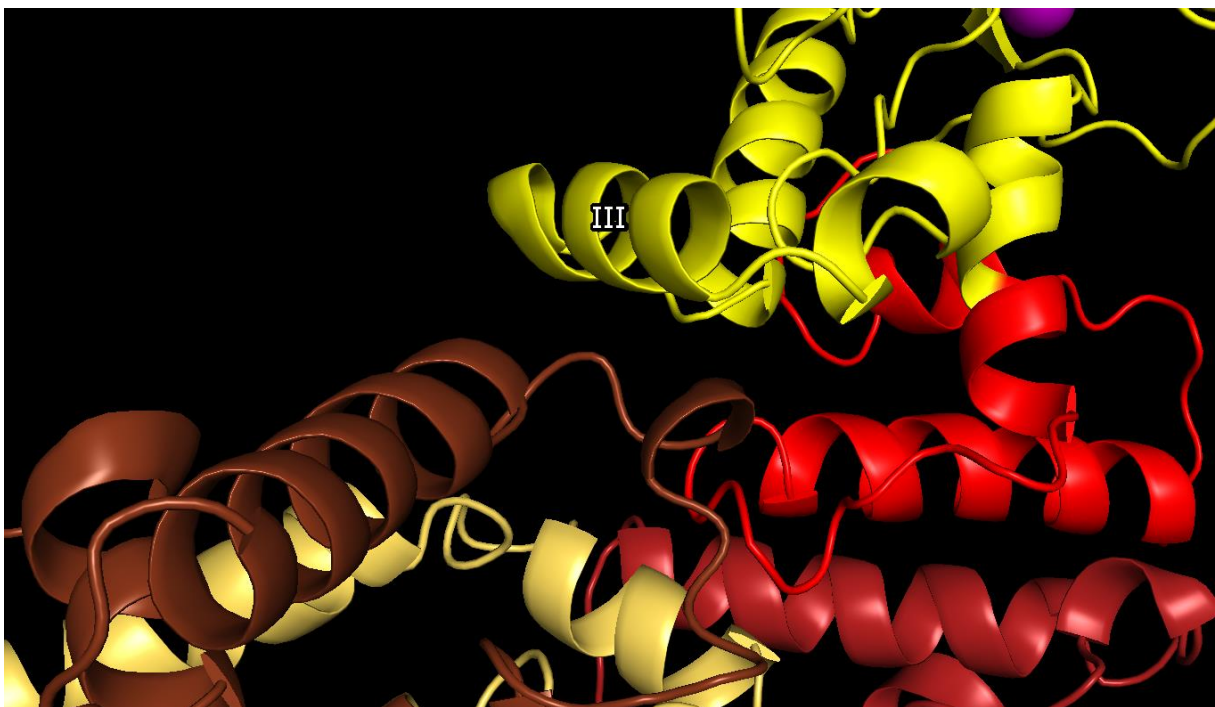
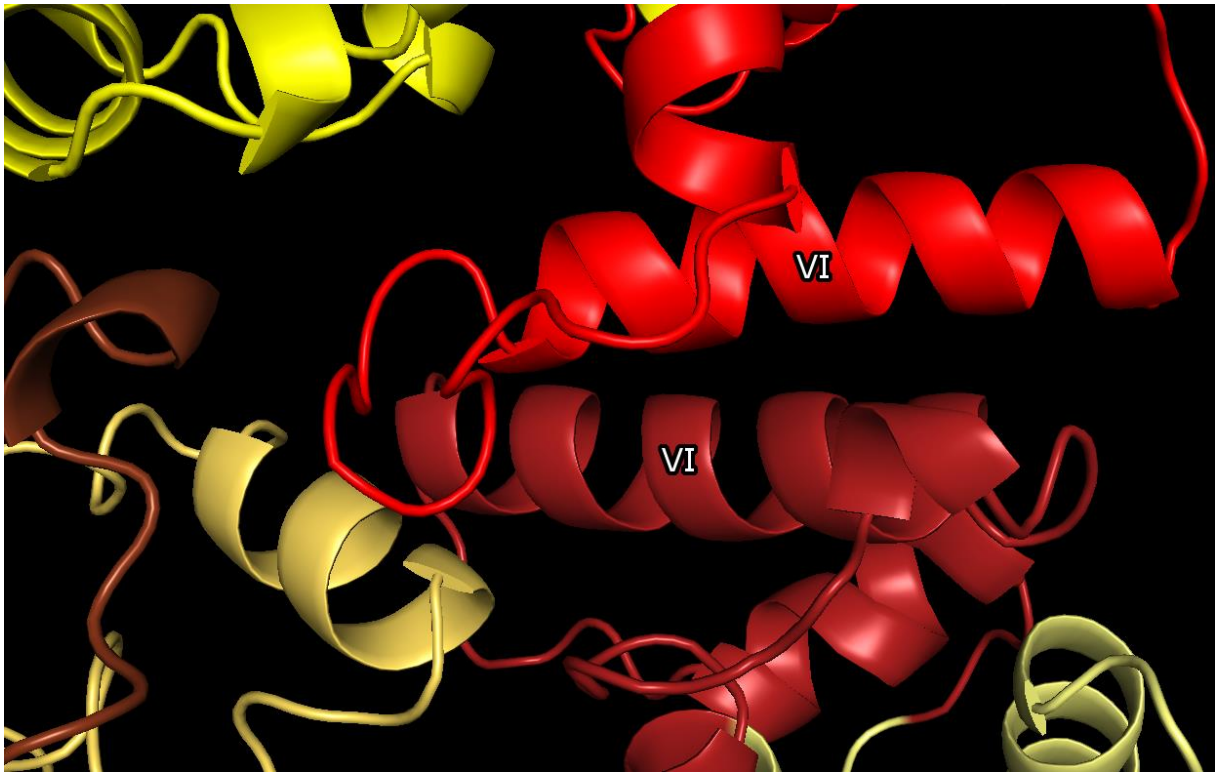
A.



B.

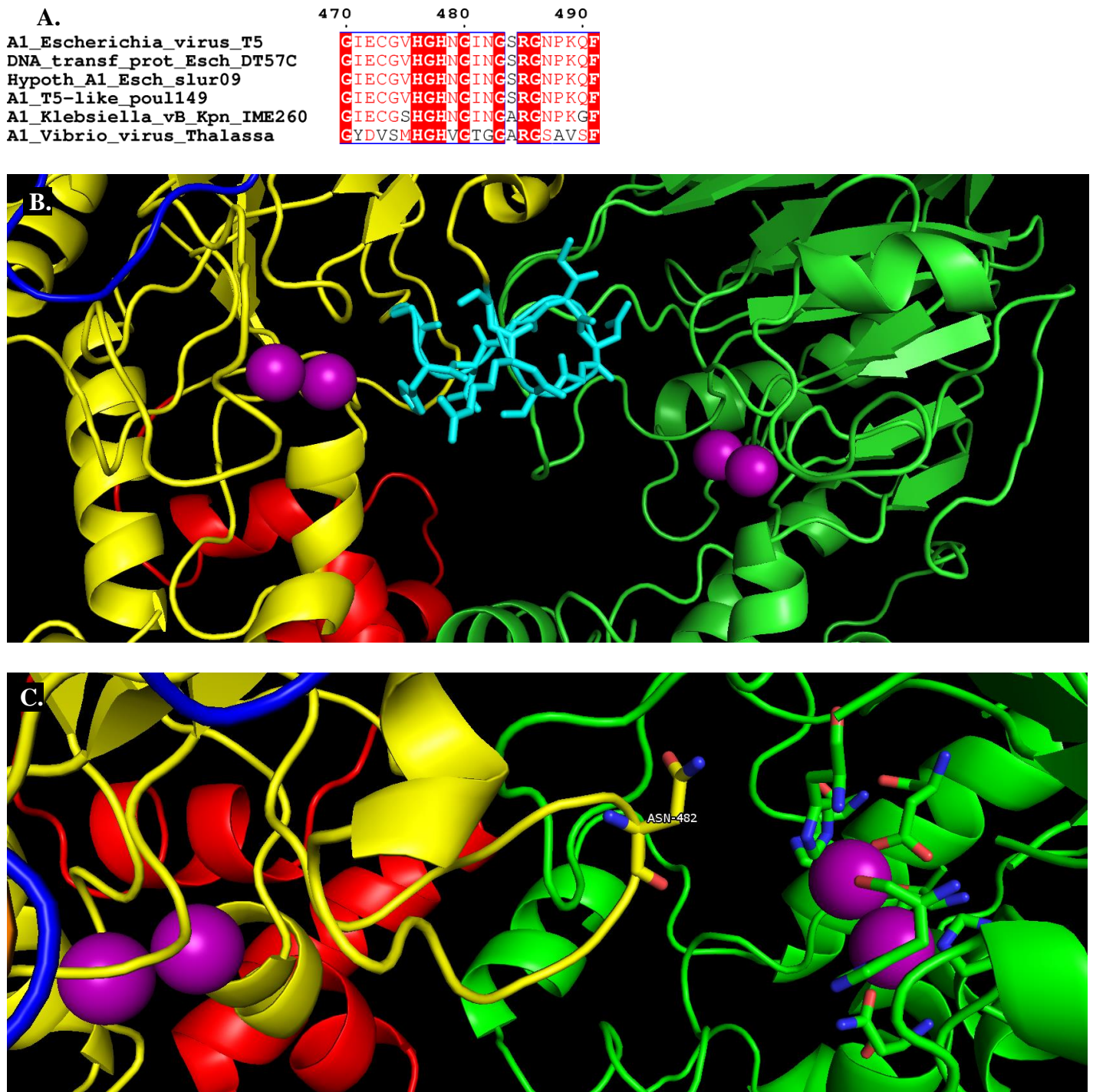


**Figure 45.** Interaction des monomères selon l'axe de symétrie 2 de la protéine, vue globale.  
**A.** Vue en surface. **B.** Vue en ruban.



**Figure 46.** Interaction des monomères selon l'axe 2 de la protéine, vue détaillée. Numérotation basée sur la Figure 35. **A.** Interaction entre les hélices  $\alpha$  du domaine de dimérisation des monomères A et B (VI avec VI). **B.** Interaction entre l'hélice  $\alpha$  (III) du site actif du monomère A et le domaine de dimérisation du monomère C.

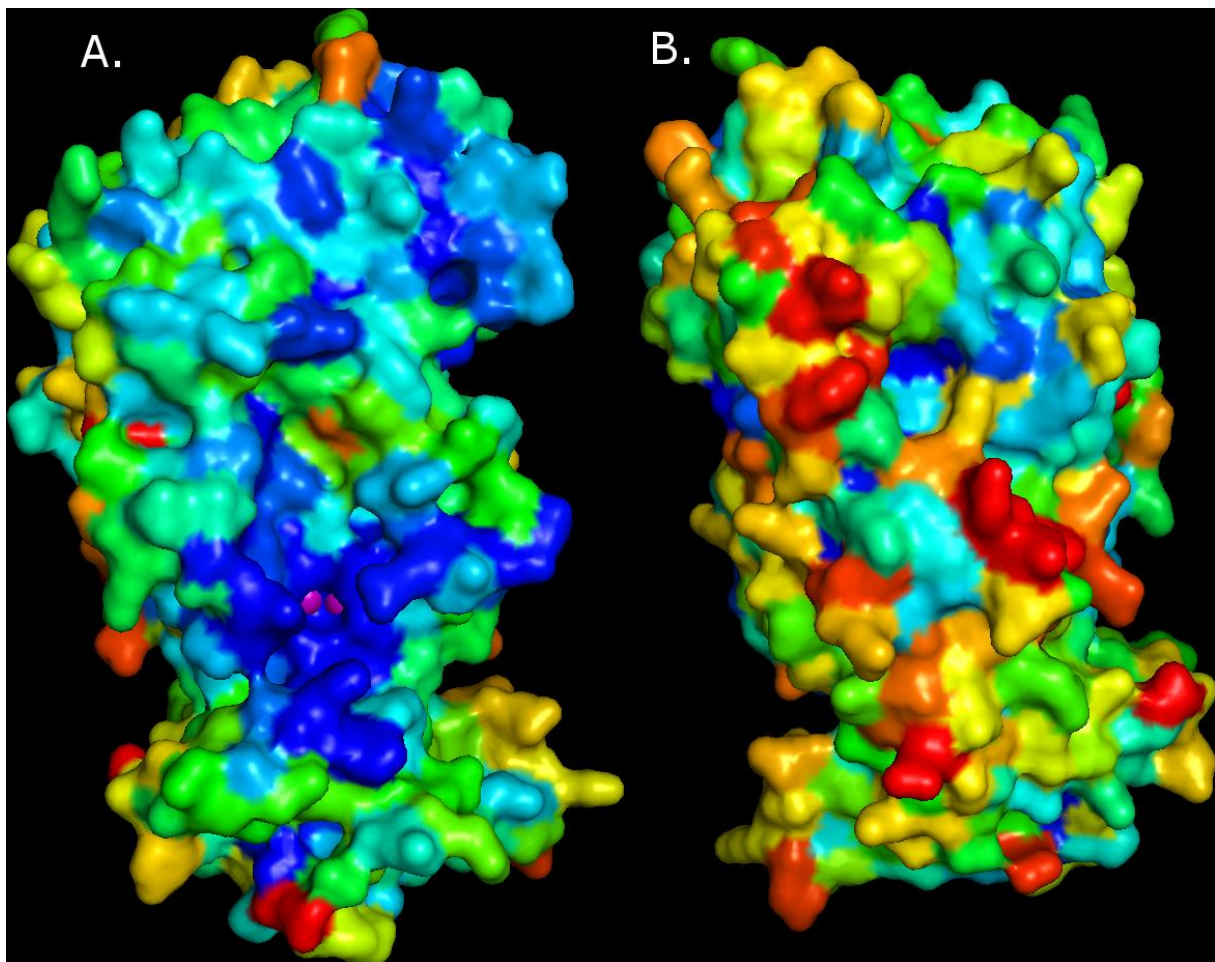
Chaque monomère contient une boucle riche en glycines très conservée qui est proche du site actif du monomère voisin. L'asparagine N482 est à 10 Å des ions  $Mn^{2+}$  du site voisin, ce qui est trop loin pour catalyser la réaction mais à bonne distance pour participer au positionnement du substrat (**Figure 47**). Cette particularité est propre à A1 et n'existe pas chez les DNases dimériques Mre11 et DNA pol II. Cette boucle riche en glycines joue très probablement un rôle majeur dans l'organisation oligomérique de A1.



**Figure 47. Interactions inter-chaînes avec le site actif de A1.** **A.** Conservation de la boucle riche en glycines (en bleu). **B.** Invasion du site actif voisin par la boucle riche en glycine (G477 à G486). **C.** Invasion du site actif voisin par l'asparagine N482.

Nous avons utilisé le serveur Consurf pour obtenir une carte de conservation des acides aminés dans la structure de A1. Consurf effectue un alignement par PSI-BLAST et calcule un score de conservation qui dépend du taux d'évolution de chaque acide aminé. Les positions qui évoluent lentement sont considérées comme étant conservées et se verront attribuer un score bas (en bleu dans la représentation de la structure). Les positions qui évoluent rapidement sont considérées comme étant variables et auront un score élevé (en rouge).

Les acides aminés du site actif sont très conservés. Certains acides aminés du domaine dupliqué sont également assez conservés du côté du site actif. Les acides aminés du côté opposé au site actif sont globalement peu conservés (**Figure 48**).



**Figure 48.** Carte de conservation des acides aminés dans la structure de A1 obtenue avec le serveur Consurf. Acides aminés conservés en bleu, acides aminés peu conservés en rouge. Ions  $Mn^{2+}$  en violet. **A.** Vue du côté du site actif. **B.** Vue du côté opposé au site actif.

### 1.1.6.5 Analyse de la structure et des propriétés de l'octamère

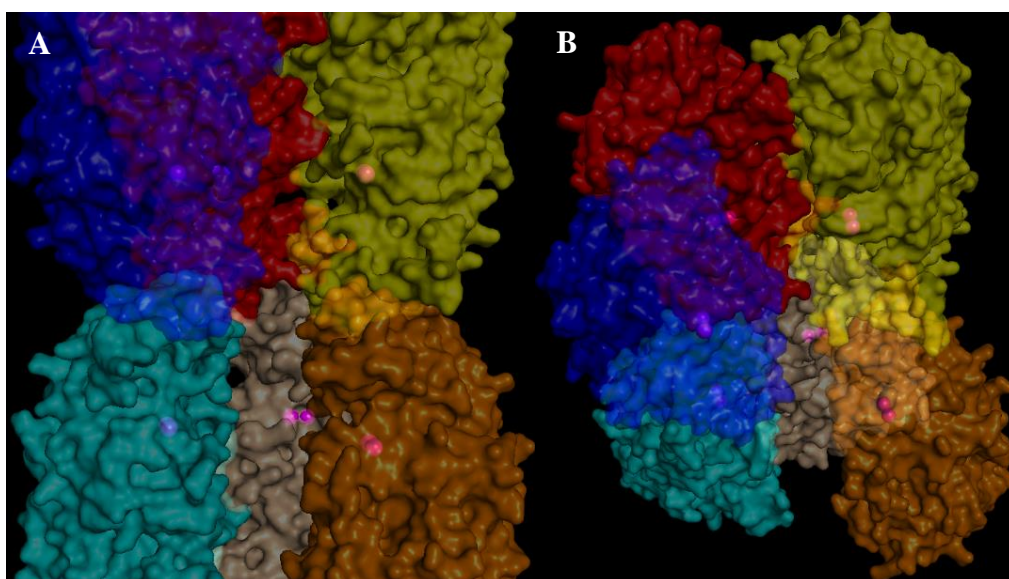
Nous avons calculé l'énergie d'interaction entre chaque résidu pour trouver les acides aminés les plus importants pour former l'octamère grâce à PDBePISA. A partir de la structure, le serveur détermine si chaque position est favorable ou défavorable en calculant la différence d'énergie entre les domaines isolés et associés sous la forme d'un  $\Delta^iG$ . Le  $\Delta^iG$  est négatif si l'interaction est défavorable et il est positif si elle est favorable. Pour A1, les acteurs majeurs de l'interaction ayant un  $\Delta^iG$  supérieur à 1 kJ/mol sont la Leucine 218, la Valine 357 (hélice III), la Lysine 414, la Proline 417 et l'Isoleucine 434 (hélice VI).

Cependant, ces acides aminés sont peu conservés chez les *T5virus* les plus divergents (**Figure 49**). Ils ne sont pas du tout conservés chez Mre11 ou la DNA pol II pour lesquels aucun alignement n'est possible, ce qui est compatible avec le fait que ces protéines ne forment pas d'octamères. Si la protéine A1 du virus *Thalassa* forme aussi un octamère, alors d'autres acides aminés sont responsables des interactions entre les différents monomères.

	210	220	...	350	360	...	410	420	430	440																																												
A1_Escherichia_virus_T5	T	A	E	S	V	P	A	L	K	G	E	H	N	R	A	S	G	V	F	L	A	K	A	N	I	K	D	D	P	A	N	A	E	L	Y	H	R	L	N	A	A	I	Y	G	A	I	A	E	K	D	.	D	T	F
DNA_transf_prot_Esch_DT57C	T	A	E	S	V	P	A	L	K	G	E	H	N	R	A	S	G	V	F	L	A	K	A	N	I	K	D	D	P	A	N	A	E	L	Y	H	R	L	N	A	A	I	Y	G	A	I	A	E	N	D	.	D	T	F
Hypoth_A1_Esch_slur09	T	A	E	S	V	P	A	L	K	G	E	H	N	R	A	S	G	V	F	L	A	K	A	N	I	K	D	D	P	A	N	A	E	L	Y	H	R	L	N	A	A	I	Y	A	I	G	E	H	.	E	T	F		
A1_T5-like_pou1149	T	A	E	S	V	P	A	L	K	G	E	H	N	R	A	S	G	V	F	L	A	K	A	N	I	K	D	D	P	A	N	A	E	L	Y	H	S	L	N	A	A	I	Y	A	I	A	K	Q	.	D	T	F		
A1_Klebsiella_vB_Kpn_IME260	T	A	E	S	V	P	A	L	K	G	D	H	N	R	A	S	G	V	F	L	A	K	A	N	I	Q	D	D	P	A	N	A	H	L	Y	Y	R	L	N	A	A	I	Y	E	A	I	E	N	K	.	D	T	F	
A1_Vibrio_virus_Thalassa	Q	A	K	T	L	S	R	K	N	G	H	N	R	E	D	W	T	F	L	.	F	D	V	E	K	D	P	K	N	A	R	T	Y	H	A	L	M	L	A	I	F	A	M	D	N	N	E	L	G	D	M			

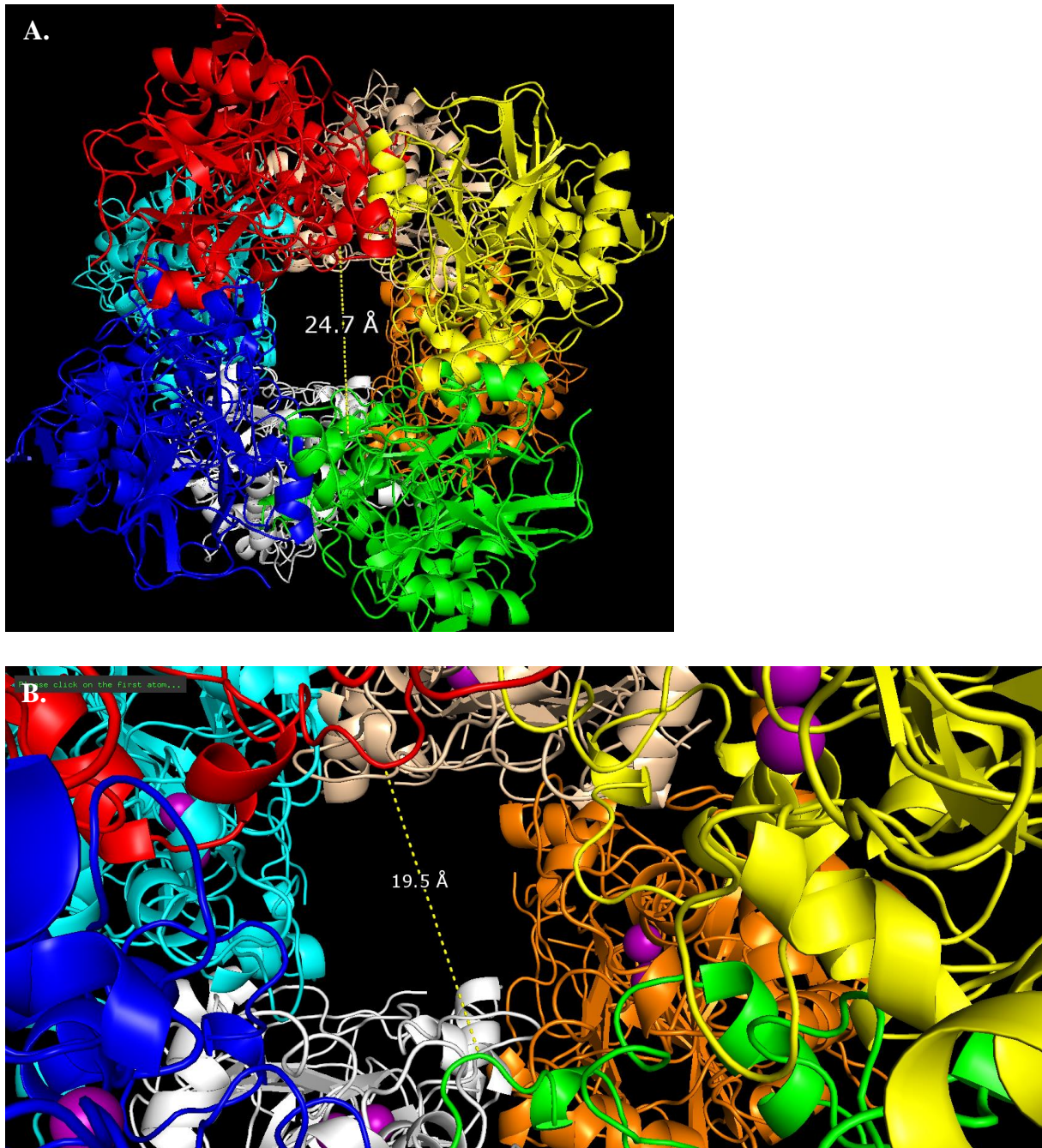
**Figure 49.** Conservation des principaux acides aminés essentiels à l'oligomérisation de A1.

J'ai représenté les manganèses du site actif en violet, ce qui permet de voir qu'ils sont localisés à l'intérieur du canal formé par l'octamère (**Figure 50**). L'ADN devrait donc être positionné à l'intérieur de l'octamère pour pouvoir être dégradé.



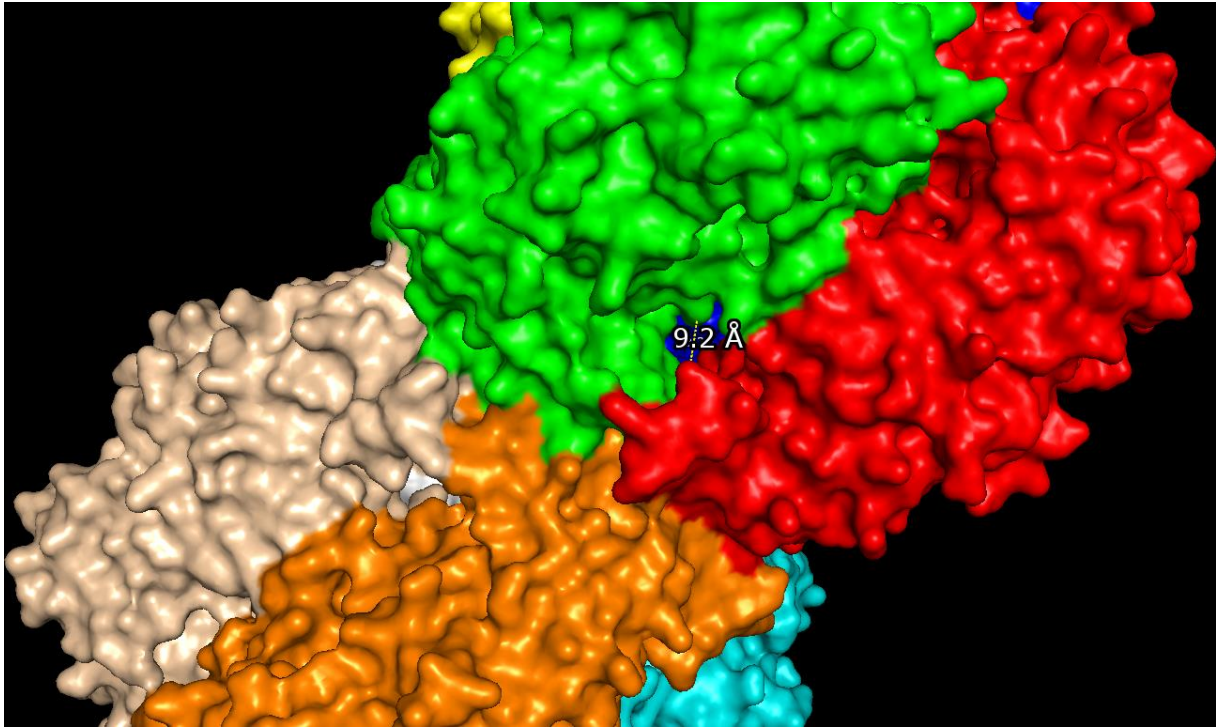
**Figure 50.** Organisation des sites actifs de A1 à l'intérieur du canal formé par l'octamère. Les ions  $Mn^{2+}$  du site actif sont représentés en violet.

Pour mieux comprendre comment A1 peut interagir avec l'ADN, nous avons mesuré le diamètre du canal formé par l'octamère. L'ADN double-brin a un diamètre de 18 à 23 Å (Wolfe 1993). Le diamètre du canal varie entre 19,5 et 24,7 Å, ce qui est tout juste suffisant pour permettre à une molécule d'ADN double-brin de s'y positionner (**Figure 51**).



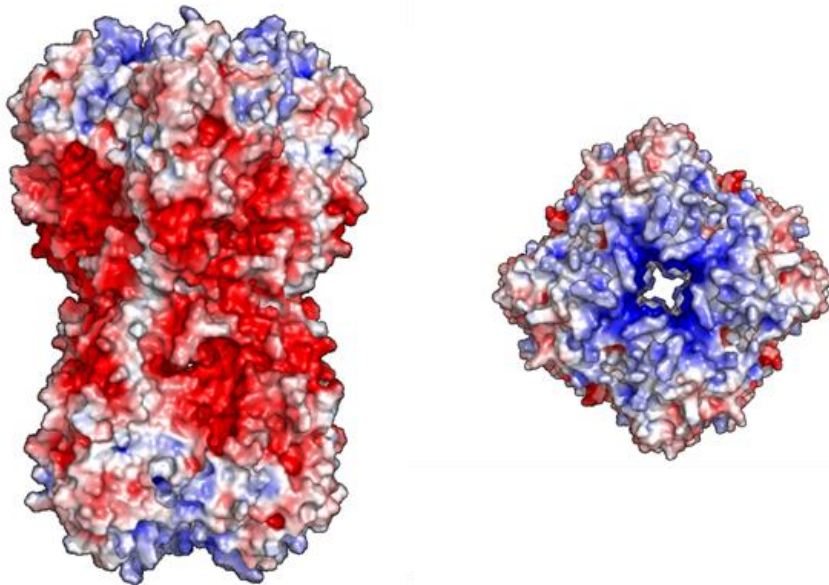
**Figure 51.** Diamètre du canal formé par l'octamère de A1. **A.** Diamètre externe. **B.** Diamètre interne.

Un orifice de 9,2 Å est visible sur le côté de la structure de l'octamère. Ce trou, tel qu'il apparaît dans cette structure est a priori insuffisant pour laisser passer des nucléotides, mais il n'est pas exclu que la réaction de coupure nucléasique induise un changement de conformation qui agrandisse cet orifice. A1 pourrait utiliser ce trou pour libérer les nucléotides résultant de la dégradation de l'ADN (**Figure 52**).



**Figure 52.** Mise en évidence d'un trou de 9,2 Å sur le côté de l'octamère de A1.

Nous avons calculé le potentiel électrostatique de surface de l'octamère pour prédire son comportement en présence d'ADN. L'entrée du canal est très chargée positivement, ce qui n'est pas favorable au déplacement de l'ADN après interaction avec la protéine. L'extérieur de la protéine est très chargé négativement et ne devrait donc pas interagir avec l'ADN. L'intérieur du canal est globalement neutre (**Figure 53**). Il est plutôt rare d'observer une séparation aussi nette des charges dans une protéine.



**Figure 53.** Potentiel électrostatique de surface de l'octamère de A1. En bleu, régions chargées positivement ; En blanc, régions neutres ; En rouge, régions chargées négativement.

En comparant des protéines A1 provenant de phages *T5viruses* divergents, on observe une conservation partielle des acides aminés chargés à l'entrée du canal (positions K132 à K137, E176 à S188, I237 à K244 et I481 à K492) (**Figure 54**).

	130	140	170		180	190	
A1_Escherichia_virus_T5	VSKFI	YNKNGF		R..	FAFMAEINV	LPTADYPLSGF	
DNA_transf_prot_Esch_DT57C	VSKFI	YNKNGF		R..	FAFMAEINV	LPTADYPLSGF	
Hypoth_A1_Esch_slur09	VSKYI	YNKNGF		R..	FAFMAEINV	LPTADYPLSGF	
A1_T5-like_poull49	VSKFI	YNKNGF		R..	FAFMAEINV	LPTADYPLSGF	
A1_Klebsiella_vB_Kpn_IME260	VSKYA	YNKNGF		ETG	FAFMAEINI	LPTADFPLSGF	
A1_Vibrio_virus_Thalassa	IMPIK	YTT..V		A..	.....QT	LPTAKQPINAA	

	230	240	250	480	490	500
A1_Escherichia_virus_T5	TA	TLKNIYIQQKAGQKAEALHN			GINCSRCNPKQFKKLG	KLNTG
DNA_transf_prot_Esch_DT57C	AA	TLKNIYIQQKAGQKAEALHN			GINCSRCNPKQFKKLG	KLNTG
Hypoth_A1_Esch_slur09	TT	TLKNIYIQQKAGQKAEALHN			GINCSRCNPKQFKKLG	KLNTG
A1_T5-like_poull49	TA	TLKNIYIQQKAGQKAEALHN			GINCSRCNPKQFKKLG	RLNTG
A1_Klebsiella_vB_Kpn_IME260	TA	TLKNIYIQQKAGQKAEALHN			GINCARGNPKGFRKLG	KMNTG
A1_Vibrio_virus_Thalassa	VC	IQMHIYTDSTRAGAEAEAHC			CTGCARGSAVSEFKKMR	IKTVTG

**Figure 54.** Conservation des acides aminés chargés positivement dans le canal de l'octamère de A1. Les résidus dont la chaîne latérale pointe vers le cœur du canal sont indiqués par une barre verte.

L'ensemble de ces résultats montre que A1 possède une structure originale par rapport aux DNases connues. L'octamère est une structure unique dont la formation est permise par de nombreuses interactions faibles. Le canal semble très adapté à la fixation d'un d'ADN double-brin car son diamètre est très proche de celui de l'ADN et il possède de nombreuses charges positives à l'entrée pouvant interagir avec l'ADN. En revanche, la circulation de l'ADN dans la protéine semble difficile à cause du grand nombre de charges positives à l'entrée du canal.





## 1.2 Etude de l'activité DNase de la protéine A1 purifiée

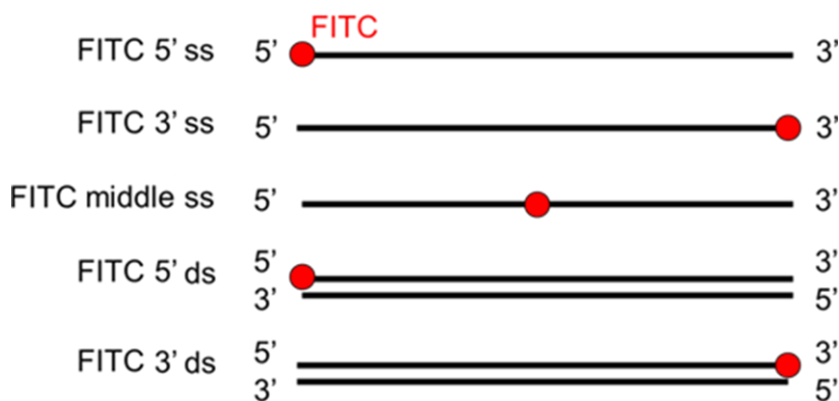
La structure de A1 a révélé l'existence d'un domaine DNase, compatible avec l'activité nucléase observée avec la protéine purifiée avant le début de ma thèse. Au cours de ce travail, j'ai utilisé plusieurs types de substrats afin de caractériser plus précisément le mode d'action de A1.

### 1.2.1 Dégradation d'oligonucléotides de 75 bases

#### 1.2.1.1 ADN fluorescents simples brins et doubles brins

Dans un premier temps, j'ai établi un protocole de test d'activité utilisant des oligonucléotides de 75 bases.

J'ai utilisé des ADN marqués en 3' ou en 5' avec une molécule fluorescente dérivée de la fluorescéine, le FITC (Fluorescein Isothiocyanate), afin de pouvoir observer l'ADN après dégradation car les méthodes classiques de coloration de l'ADN sur gel ne permettent pas d'observer de très petits fragments. L'ADN FITC 5' ss est complémentaire de l'ADN FITC 3' ss. J'ai également hybridé chacun de ces ADN avec leur ADN complémentaire non marqué afin de pouvoir comparer l'activité de A1 sur un ADN simple-brin et double-brin (**Figure 55**). Un des ADN a été marqué au milieu de sa séquence.



**Figure 55.** Fragments d'ADN marqués au FITC utilisés pour le test d'activité DNase (ss, simple brin ; ds, double brin).

Les résultats de bioinformatique ont permis de comparer la protéine A1 avec Mre11, une DNase utilisant les ions  $Mn^{2+}$  pour son activité. De plus, nous avons résolu la structure de A1 en présence de  $Mn^{2+}$ . Nous avons donc effectué les tests d'activité en présence de  $MnCl_2$ .

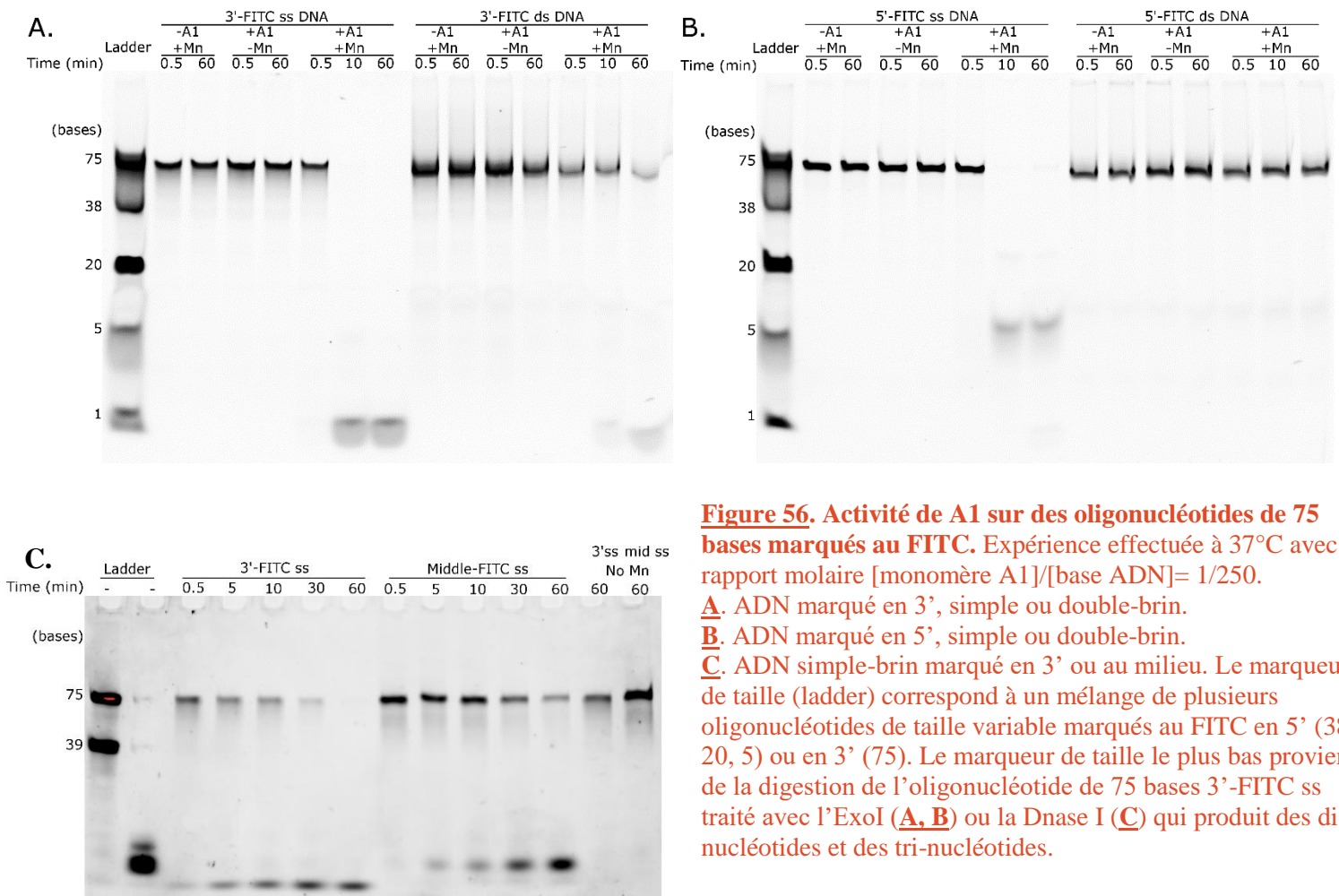
L'ADN simple-brin de 75 bases disparaît rapidement au profit d'une bande de très petite taille qui pourrait correspondre à 1 base pour l'ADN marqué en 3' et d'une bande estimée à 5 bases pour l'ADN marqué en 5' (**Figure 56** A et B). Cette différence entre les deux produits de dégradation ne peut pas être interprétée avec certitude, mais elle pourrait indiquer une digestion incomplète laissant un fragment de 5 bases intact du côté 5'. Une dégradation partielle de l'ADN double-brin est visible pour les deux marquages après 60 min d'incubation (**Figure 56** A et B). L'ADN double-brin semble donc peu dégradé par rapport à l'ADN simple-brin. On ne peut pas exclure que la dégradation observée pour l'ADN double-brin corresponde en partie à la dégradation de quelques ADN fluorescents non-hybridés dans le mélange réactionnel.

Ces résultats confirment donc l'activité DNase de A1. Cette activité apparaît très processive car aucun intermédiaire n'est visible sur le gel, ce qui empêche de connaître le sens de dégradation de l'ADN.

Les résultats de la dégradation de l'oligonucléotide marqué en 3' pourraient être expliqués par la dégradation d'une seule base par A1. Pour vérifier si A1 dégrade tout le fragment d'ADN, j'ai utilisé un ADN simple-brin marqué au FITC au milieu de la séquence (**Figure 56** C). Cet ADN est dégradé sans aucun intermédiaire visible. L'ADN est donc entièrement dégradé par A1. La dégradation d'un ADN marqué au milieu est plus lente que pour un ADN marqué en 3'. La liaison de A1 à l'ADN pourrait être affectée par le FITC : la structure nous montre que le canal formé par l'octamère de A1 n'est pas suffisamment grand pour accommoder un ADN d'un diamètre supérieur à 23 Å (voir **Figure 51**). Lorsque le marquage au FITC est au milieu de la séquence, il pourrait être responsable d'une contrainte stérique. De nouveau, dans cette expérience, la position des produits finaux de dégradation par A1 obtenus avec les oligonucléotides marqués en 3' ou au milieu n'apparaît pas comme étant identique sur le gel. La taille apparente est légèrement supérieure pour l'ADN marqué au milieu par rapport à l'ADN marqué en 3' et pourrait correspondre à un di-nucléotide, au même niveau que le produit de dégradation de la DNase I qui produit majoritairement des di-nucléotides et minoritairement des tri-nucléotides<sup>122</sup>. On ne peut pas exclure que ce résultat soit dû à la façon dont l'ADN est marqué au FITC selon la position du marquage.

La structure de A1 suggère que le manganèse pourrait être le cation divalent nécessaire à son activité, comme pour Mre11. Nous confirmons ici que A1 n'a aucune activité sur les ADN testés en absence de manganèse dans le milieu réactionnel.

Dans ces expériences, nous avons utilisé un rapport molaire [monomère A1] / [ADN] de 0,3. On ne sait pas, à priori, si c'est le monomère ou l'octamère qui dégrade l'ADN. Dans tous les cas, ces expériences prouvent qu'une seule molécule de A1 peut dégrader plusieurs molécules d'ADN.



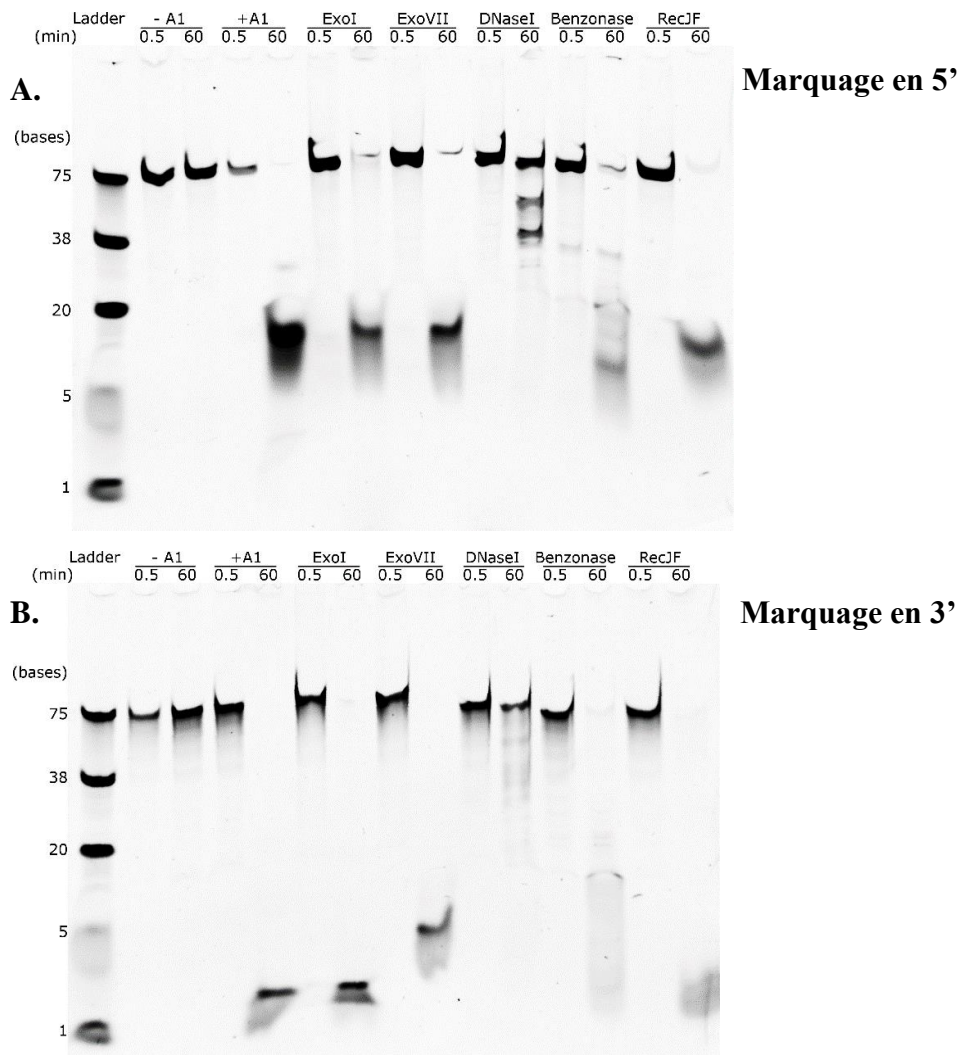
**Figure 56. Activité de A1 sur des oligonucléotides de 75 bases marqués au FITC.** Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN]= 1/250.  
**A.** ADN marqué en 3', simple ou double-brin.  
**B.** ADN marqué en 5', simple ou double-brin.  
**C.** ADN simple-brin marqué en 3' ou au milieu. Le marqueur de taille (ladder) correspond à un mélange de plusieurs oligonucléotides de taille variable marqués au FITC en 5' (38, 20, 5) ou en 3' (75). Le marqueur de taille le plus bas provient de la digestion de l'oligonucléotide de 75 bases 3'-FITC ss traité avec l'ExoI (**A, B**) ou la Dnase I (**C**) qui produit des dinucléotides et des tri-nucléotides.

### 1.2.1.2 Comparaison de l'activité de A1 avec différentes nucléases

Afin de comparer l'activité DNase de A1 à celle de nucléases déjà caractérisées, les oligonucléotides simple brin marqués en 5' ou 3' ont été incubés avec A1 ou, en parallèle, avec (i) des exonucléases dégradant exclusivement l'ADN simple brin : ExoI (polarité 3'-5'), RecJF (polarité 5'-3') et ExoVII (sans polarité préférentielle) ou (ii) les endonucléases DNase I et benzonase (qui produit des fragments de moins de 8 nucléotides).

Avec l'ADN marqué en 5', les produits de dégradation obtenus sont similaires pour toutes les enzymes testées à l'exception de la benzonase (NB : la DNase I n'a pas été utilisée à une concentration suffisante pour dégrader totalement l'ADN) (**Figure 57**). En revanche, pour l'ADN marqué en 3', A1 produit un fragment de taille comparable à ceux produits par ExoI et RecJF et de taille inférieure au fragment produit par ExoVII ou la benzonase. Ces tests

préliminaires ne permettent pas de différencier le mode d'action de A1 de ceux de ExoI et RecJF, suggérant que A1 pourrait être une exonucléase agissant dans les deux directions 3'→5' et 5'→3'. Ces réactions ont été analysées après une heure d'incubation à 37°C. Une comparaison plus poussée est nécessaire avec des études de cinétique et en testant la dégradation sur d'autres substrats, par exemple en modifiant la composition en bases de l'ADN ou en bloquant une partie de l'ADN avec des liaisons non-hydrolysables.

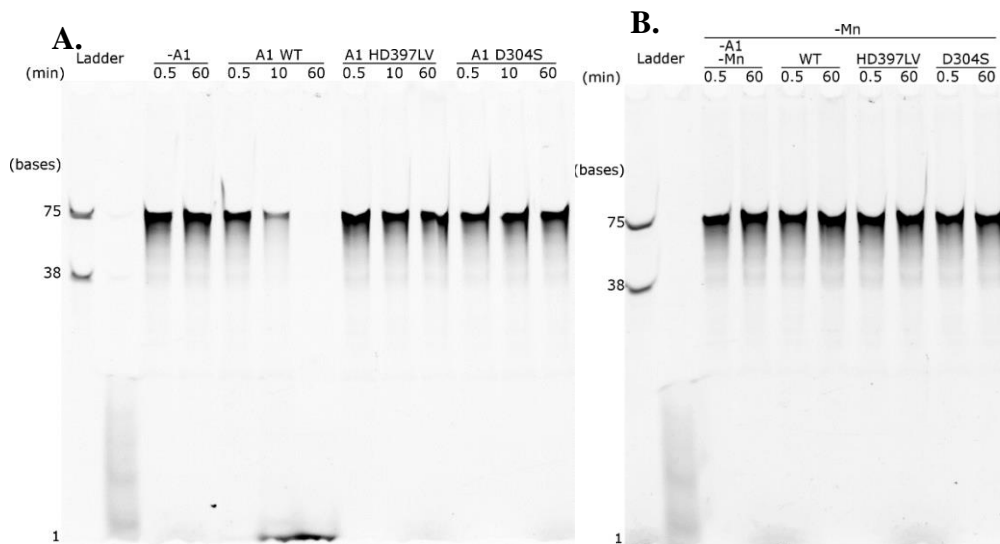


**Figure 57. Activité de A1 et de différentes DNases sur des oligonucléotides de 75 bases marqués au FITC en 5' (A) ou en 3' (B).** Expérience faite à 37°C avec un rapport molaire [monomère A1]/[base ADN] = 1/250. ExoI, ExoVII et RecJF : 20 Unités par réaction. DNase I et Benzonase : 0,001 U par réaction. **A.** ADN simple-brin marqué en 5'. **B.** ADN simple-brin marqué en 3'. NB : Dans cette expérience, la DNase I a été accidentellement trop diluée et n'a pas digéré totalement l'ADN. Il a déjà été observé que les deux oligonucléotides testés sont dégradés par la DNase I et donnent un nucléotide marqué de la même taille que celui obtenu avec A1. Bande à 1 base du marqueur : digestion par ExoI.

### 1.2.1.3 Absence d'activité DNase chez les mutants du site actif

La structure de A1 a permis d'identifier plusieurs acides aminés potentiellement essentiels pour l'activité de A1. L'aspartate 304 est responsable de la coordination d'un des deux ions manganèse. L'histidine 397 et l'aspartate 398 ne participent pas directement à la coordination des ions manganèse, mais ce sont deux acides aminés très conservés chez les DNases similaires à A1. Nous avons muté ces acides aminés pour tester leur importance pour l'activité DNase de A1. Les mutants testés sont HD397LV et D304S : les substitutions par ces acides aminés ne montrent que les positions les plus sensibles à la mutation des protéines car ces substitutions n'empêchent généralement pas la fonction des protéines <sup>123</sup>.

Les deux protéines mutées n'ont aucune activité de dégradation sur un ADN de 75 bases par rapport à la protéine sauvage (**Figure 58 A**). Aucune activité n'est observée sans MnCl<sub>2</sub> dans tous les tests (**Figure 58 B**). Les deux mutants testés ont donc perdu leur activité DNase. Cette perte d'activité peut être due à un rôle crucial joué par les acides aminés mutés dans l'activité ou à une déstructuration de la protéine. La conservation de la structure des mutants a été démontrée par microscopie électronique (voir paragraphe 1.3.3) et par dichroïsme circulaire (voir paragraphe 1.4.1). Nous avons donc identifié des acides aminés essentiels pour l'activité DNase de A1.



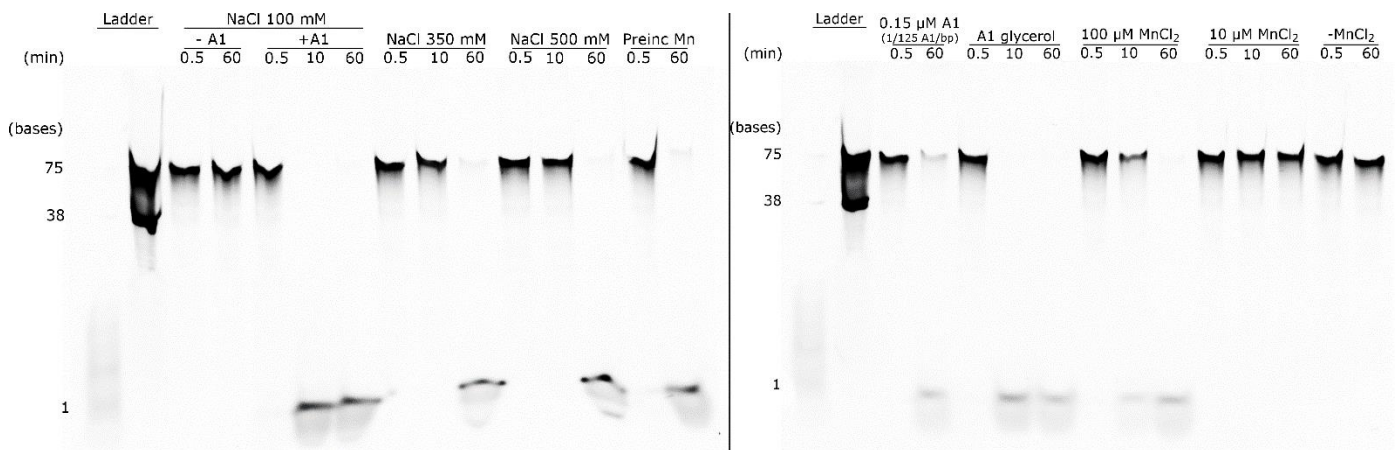
**Figure 58. Activité des protéines A1 mutées dans le site actif par rapport à la protéine A1 sauvage (WT) sur l'ADN 3'-FITC ss.** Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN] = 1/250 en présence (**A**) ou en absence (**B**) de 1 mM MnCl<sub>2</sub>. Bande à 1 base du marqueur : digestion par la DNase I.

#### 1.2.1.4 Caractérisation des conditions optimales pour observer l'activité de A1

Nos tests d'activité *in vitro*, même s'ils peuvent ne pas être totalement représentatifs de l'activité de A1 *in vivo*, nous renseignent sur les propriétés de A1. J'ai optimisé les concentrations de NaCl et de MnCl<sub>2</sub> utilisées pour maximiser cette activité. L'activité DNase de A1 est plus faible lorsque la concentration en NaCl augmente (**Figure 59**) : il pourrait y avoir une compétition avec la fixation de l'ADN. A1 a moins d'activité à 100 µM de MnCl<sub>2</sub> et perd toute son activité à 10 µM ou sans MnCl<sub>2</sub> (**Figure 59**) : il est donc nécessaire de travailler avec 1 mM de MnCl<sub>2</sub> pour avoir la meilleure activité.

J'avais fait l'hypothèse que la pré-incubation de A1 en présence de MnCl<sub>2</sub> 1 mM augmenterait l'activité de A1 en permettant à l'enzyme de démarrer plus rapidement la réaction. Cependant, A1 dégrade l'ADN moins efficacement à 60 minutes que lorsque la protéine n'a pas été pré-incubée en présence de MnCl<sub>2</sub>. Ce résultat suggère que lorsque la protéine est déjà majoritairement organisée sous forme octamérique en présence de Mn<sup>2+</sup> (voir SEC-MALLS paragraphe 1.1.2 et SAXS paragraphe 1.1.3), son activité est diminuée.

A1 est moins active lorsqu'elle est utilisée à 0,15 µM plutôt qu'à 0,3 µM. Les expériences doivent donc être effectuées à 0,3 µM de protéine au minimum. Un autre lot de protéine a été congelé en glycérol 40 % à -80°C. L'activité de ce lot n'est pas diminuée par rapport au lot habituellement utilisé malgré la présence de 2 % de glycérol dans la réaction finale. La congélation en glycérol n'empêche donc pas l'activité de A1 et pourrait présenter l'avantage de protéger la protéine A1 au cours de sa conservation à -20°C.



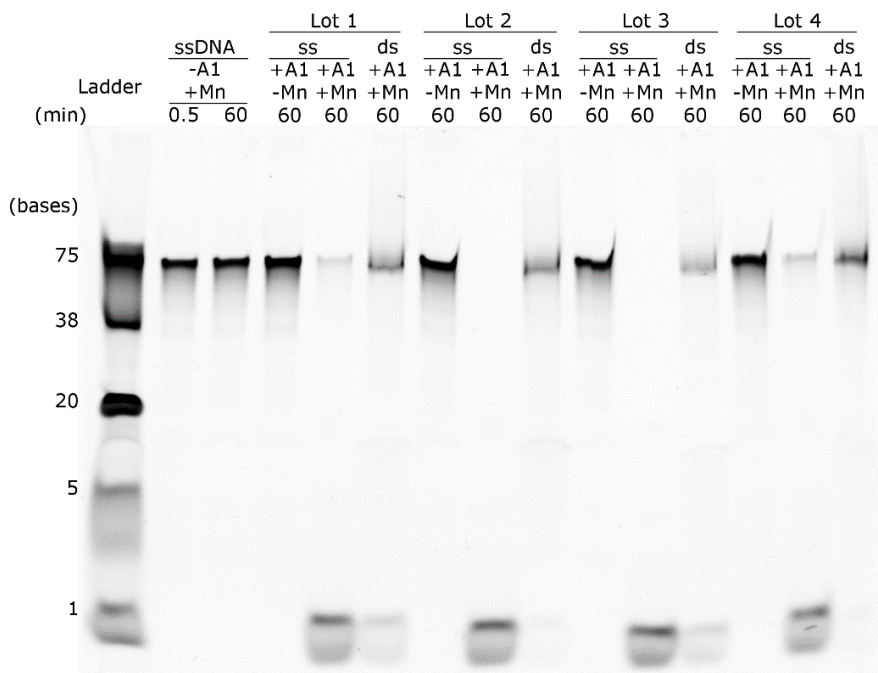
**Figure 59.** Optimisation du test d'activité de A1 sur l'ADN FITC 3' ss. Sauf indication contraire, [A1] = 0,3 µM, [MnCl<sub>2</sub>] = 1 mM, [NaCl] = 100 mM. Expérience faite à 37°C avec un rapport molaire [monomère A1]/[base ADN] = 1/250. A1 glycérol : lot différent de A1 congelé avec 40% de glycérol à -80°C, soit 2 % de glycérol dans la réaction finale. Bande à 1 base du marqueur : digestion par la DNase I.

Je n'ai donc pas découvert de meilleures conditions pour étudier l'activité de A1 *in vitro*. J'ai démontré que la congélation en glycérol n'empêche pas son activité. Il faudra comparer l'activité d'un lot congelé en glycérol et non congelé pour vérifier qu'il n'y a pas de diminution de l'activité. Le stockage à long terme de la protéine pourrait alors être simplifié.

#### 1.2.1.5 Comparaison de l'activité de différents lots de A1 sur des oligonucléotides marqués

J'ai testé 4 lots de A1 provenant de purifications différentes sur l'ADN simple brin FITC 3' ss et l'ADN double brin FITC 3' ds (**Figure 60**). Tous les lots ont dégradé l'ADN FITC 3' ss, les lots 2 et 3 étant les plus efficaces alors que nous avons utilisé des molarités identiques. Des variations d'activité en fonction du lot sont aussi observables sur les ADN double brin. Il existe donc une variabilité de l'activité spécifique de la protéine A1 en fonction du lot utilisé.

Au moins 2 hypothèses sont possibles pour expliquer ces variations d'activité spécifique en fonction des lots de A1 : i) Selon les lots, on purifie une protéine contaminante non détectable sur gel SDS-PAGE ou un cofacteur qui favorise l'activité. ii) Si A1 est instable en solution elle peut perdre plus ou moins son activité au cours des étapes successives de purification. Comme les temps de purification peuvent varier selon les lots, on peut observer des activités spécifiques différentes au moment de la congélation des aliquots dans l'azote liquide. Comprendre l'origine de la différence d'activité entre ces lots de protéines est essentiel pour caractériser l'activité de A1 sur les différentes formes d'ADN.



**Figure 60. Activité de 4 lots de A1 sur deux types d'ADN (FITC 3' ss simple-brin ; FITC 3' ds double-brin).** Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[pb ADN] = 1/250. Bande à 1 base du marqueur : digestion par Exo I.



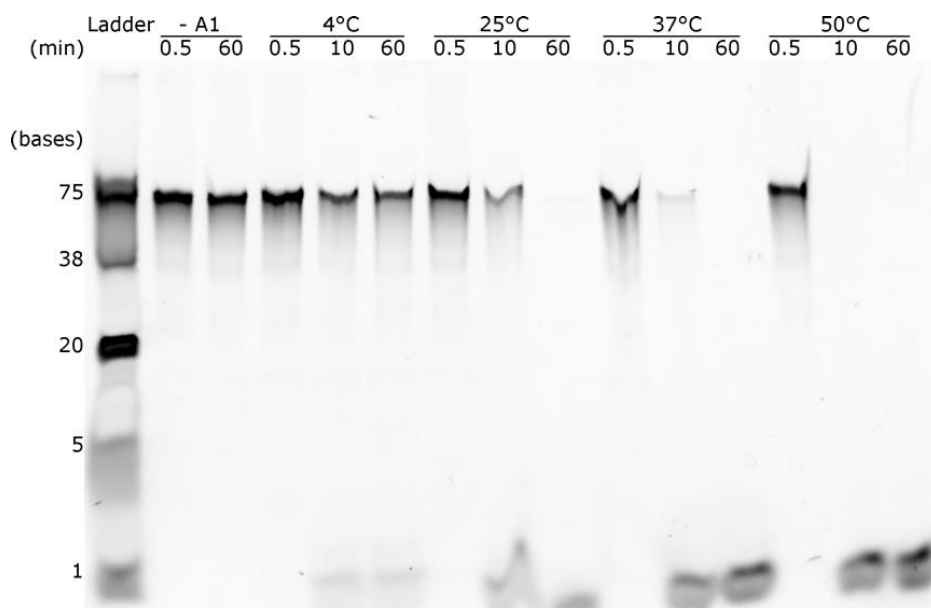
### 1.2.2 Tests à différentes températures

Aucune des expériences décrites jusqu'ici n'ayant permis d'identifier d'intermédiaires de réaction, nous ne pouvons pas déterminer le sens de digestion de A1. Le principal obstacle pour caractériser cette direction est la très grande processivité de cette DNase. J'ai donc testé l'activité de A1 à différentes températures pour ralentir cette processivité sur l'ADN simple-brin et tenter d'observer des intermédiaires de réaction.

A1 a dégradé l'ADN FITC 3' ss à toutes les températures testées, sans aucun intermédiaire de réaction visible (**Figure 61**). La dégradation est très lente à 4°C et elle est ralentie à 25°C par rapport à la dégradation à 37°C. De manière surprenante, A1 semble être plus active à 50°C qu'à 37°C. Si cette observation est confirmée par de nouveaux essais à 50°C et à plus haute température, cela pourrait ouvrir des perspectives pour une utilisation en biotechnologie.

L'absence d'intermédiaires de réaction indique que A1 est très processive. Ces intermédiaires n'apparaissent pas lorsque la réaction est ralentie : c'est donc probablement la liaison de A1 à l'ADN qui est l'étape limitante dans la réaction.

Comme la diminution de la température de réaction n'a pas permis l'observation d'intermédiaires de réaction pour identifier le sens de digestion de l'ADN par A1, il est nécessaire de trouver un autre moyen pour étudier cette activité.



**Figure 61.** Test d'activité de A1 à différentes températures sur l'ADN FITC 3' ss avec un rapport molaire [monomère A1]/[base ADN] = 1/250. Bande à 1 base du marqueur : digestion par ExoI.

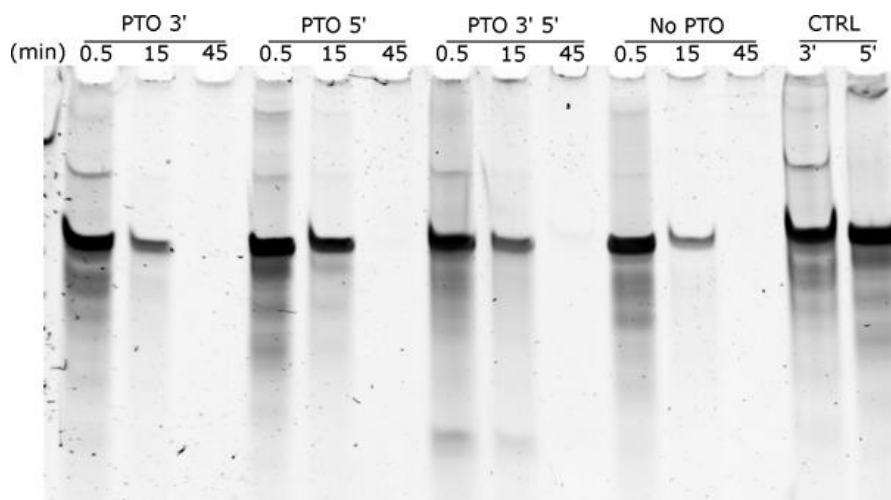
### 1.2.3 Tests d'activité endonucléase avec des oligonucléotides portant des liaisons phosphorothioate aux deux extrémités

#### 1.2.3.1 Blocage des extrémités par 3 liaisons phosphorothioate

Pour déterminer le sens de l'activité de A1, j'ai utilisé des oligonucléotides possédant des liaisons phosphorothioate (PTO). Ces liaisons, dans lesquelles un groupement sulfure remplace l'oxygène libre de la liaison phosphodiester, ne peuvent pas être dégradées par la plupart des nucléases<sup>124</sup>, il est donc possible de bloquer l'activité DNase sur des positions spécifiques de l'ADN. Le premier essai a été fait avec des ADN bloqués en 3' ou en 5' par 3 liaisons PTO.

On observe une dégradation totale de tous les ADN testés en 45 minutes (**Figure 62**). A1 a donc été capable de couper l'ADN malgré des extrémités non-hydrolysables, ce qui suggère qu'elle possède une activité endonucléase. Comme les extrémités non hydrolysables de 3 bases ne sont pas détectées en SYBR Gold, nous ne pouvons pas vérifier si A1 a pu ou non dégrader l'ADN malgré les liaisons PTO.

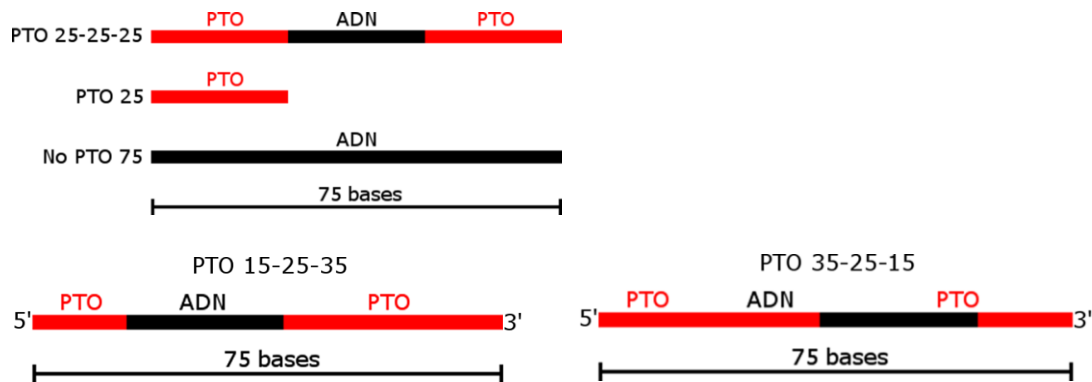
Afin d'être en mesure de détecter les fragments non digérés par A1 et conclure sur son activité endonucléase, j'ai reproduit l'expérience sur un ADN avec des extrémités non-hydrolysables plus longues.



**Figure 62.** Activité de A1 sur un oligonucléotide de 75 bases bloqué pour la digestion en 3', en 5' ou des deux côtés par une extrémité PTO non-hydrolysable de 3 bases. Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN] = 1/250. L'ADN est coloré avec du SYBR Gold. PTO 3', PTO 5' et PTO 3' 5' portent 3 liaisons PTO en respectivement en 3', 5', 3'et 5'. Le témoin No PTO correspond à un oligonucléotide de 75 bases sans liaison PTO. Le témoin CTRL correspond aux ADN mélangés dans le tampon de réaction en absence de A1 et de MnCl<sub>2</sub>, sans incubation.

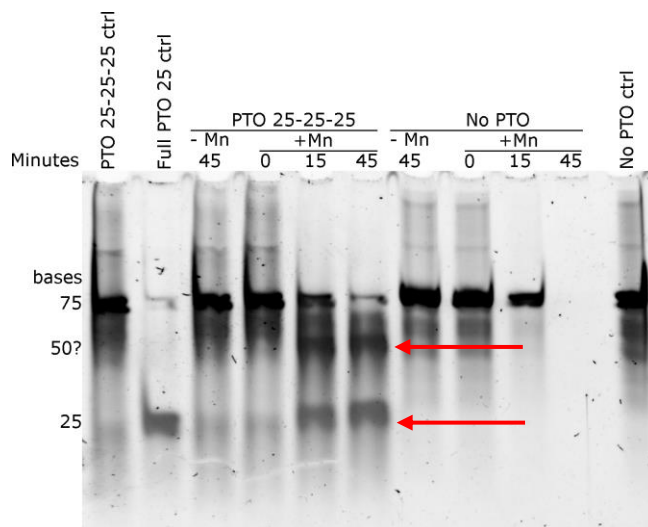
### 1.2.3.2 PTO 25-25-25

J'ai utilisé des ADN portant 25 liaisons PTO non-hydrolysables à chaque extrémité pour confirmer l'activité endonucléase de A1. Pour vérifier que cette longueur d'ADN modifié est visible sur gel, j'ai utilisé un fragment de 25 bases possédant uniquement des liaisons PTO.



**Figure 63.** Oligonucléotides partiellement non hydrolysables utilisés dans les tests d'activité de A1 : ADN de 75 bases bloqués pour la digestion par des extrémités phosphorothioate (PTO) non-hydrolysables (en rouge).

A1 dégrade l'ADN bloqué en 45 minutes, avec deux bandes intermédiaires : une proche de la position 25 bases et une d'environ 50 bases (**Figure 64**). La bande à 25 bases correspond probablement aux fragments PTO non-hydrolysables, mais la bande d'environ 50 bases est nécessairement un oligonucléotide composé d'une partie hydrolysable et d'une partie non-hydrolysable.



**Figure 64.** Activité de A1 sur l'ADN PTO 25-25-25, bloqué pour la digestion en 3' et en 5'. La digestion est comparée avec un ADN de 75 bases sans liaisons PTO. Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN]= 1/250. Les ADN témoins n'ont été incubés ni en présence de A1 ni en présence de MnCl<sub>2</sub>. Flèches rouges : intermédiaires de réaction observés après digestion par la protéine A1.

Ces expériences confirment l'activité endonucléase de A1 sur l'ADN simple-brin car les extrémités PTO qui bloquaient les deux côtés de l'ADN n'ont pas été dégradées. En revanche, A1 ne fait pas ou seulement très rarement plusieurs coupures endonucléase, sinon le fragment de 50 bases aurait été dégradé. A1 possède donc une double activité endonucléase et une activité exonucléase sur l'ADN simple brin.

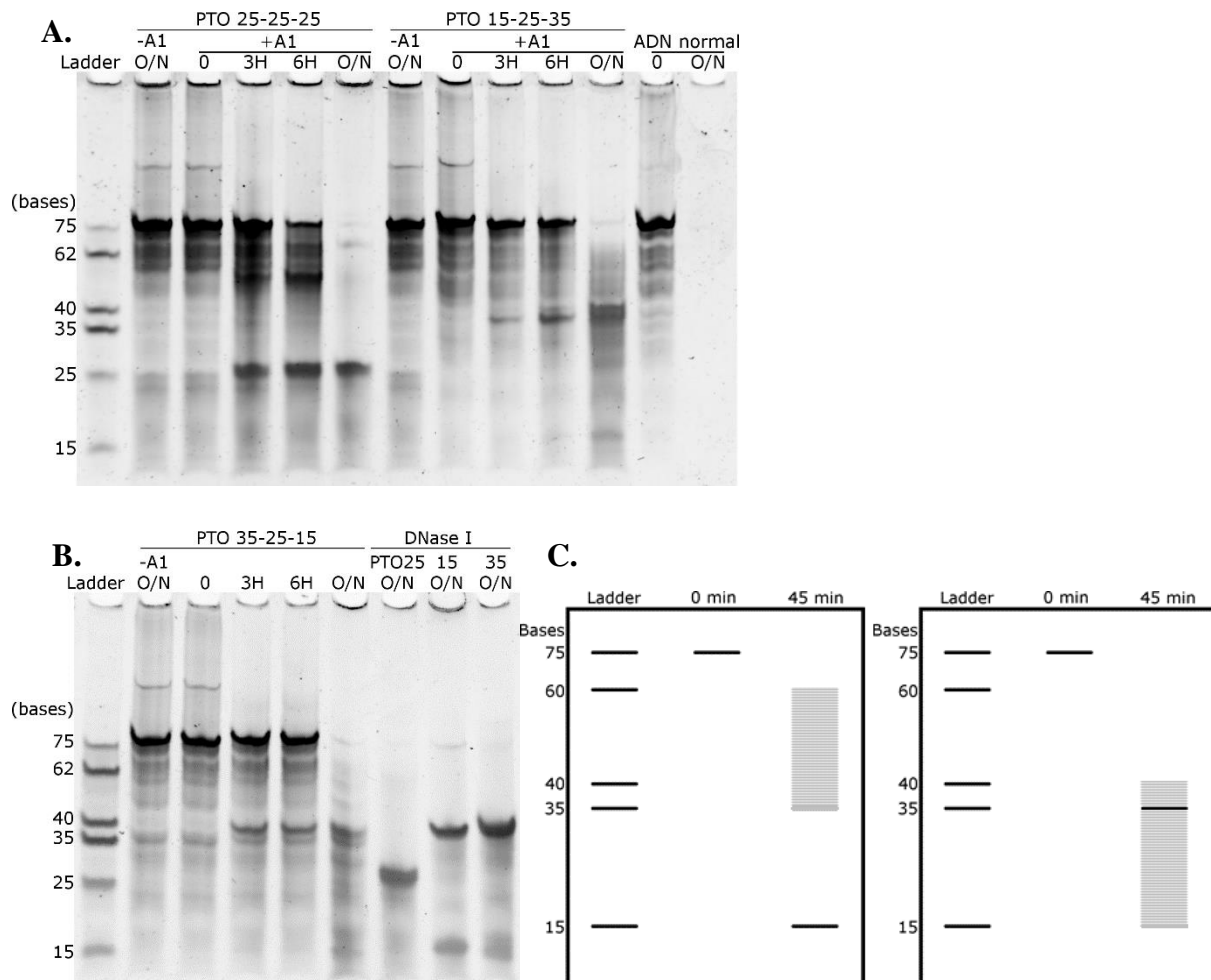
Il est possible que A1 se "bloque" sur l'ADN PTO après avoir effectué une coupure. Elle empêcherait alors la fixation d'une autre protéine A1 qui pourrait dégrader la partie non PTO (25 bases centrales) de l'ADN de 50 bases restant.

Ces résultats ne permettent pas de conclure sur le sens de digestion de A1. Puisque nous pouvons voir les ADN PTO non dégradés par A1, nous pouvons utiliser des ADN avec des extrémités PTO de longueurs différentes. Si A1 a un sens de digestion privilégié, elle ne les dégradera pas de la même manière.

#### *1.2.4 Détermination du sens de la digestion avec les oligonucléotides PTO 15-25-35 et PTO 35-25-15*

Pour déterminer le sens de digestion de l'ADN par A1, nous avons utilisé des ADN dont les extrémités PTO ont des longueurs différentes, soit 15 et 35 bases, positionnées en 5' ou 3' (**Figure 63**).

Les intermédiaires de réaction à 25 et 50 bases apparaissent entre 0 et 3 heures pour l'ADN PTO 25-25-25. En fin de réaction, seul l'intermédiaire à 25 bases est encore présent, comme pour le contrôle positif DNase I (**Figure 65** A et B). Le temps nécessaire à la digestion maximale est donc nettement plus long (plusieurs heures contre environ 30 minutes) pour les ADN PTO par rapport aux ADN non modifiés. A1 pourrait donc "stationner" sur la région PTO, en raison de son affinité importante pour l'ADN et de son incapacité à couper, et ainsi fortement ralentir la digestion des liaisons phosphodiester non modifiées. Par ailleurs, nous utilisons un rapport molaire [monomère de A1]/[molécules d'ADN] de 0,3, il y a donc environ 1 octamère de A1 pour 25 molécules d'ADN. Si la protéine A1 est séquestrée sur l'ADN PTO, la concentration de A1 libre pouvant dégrader l'ADN diminue rapidement. Cette hypothèse expliquerait qu'une nuit d'incubation soit nécessaire pour la digestion totale jusqu'aux fragments non hydrolysables.



**Figure 65.** **Activité de A1 sur des ADN PTO non hydrolysables en 3' et 5' (voir Figure 63).** La digestion est comparée avec celle d'un ADN de 75 bases sans liaison PTO ou avec la digestion par la DNase I sur les ADN PTO. L'expérience est effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN] = 1/250. DNase I : 1 unité dans un tampon Tris 50 mM, MgCl<sub>2</sub> 1 mM, CaCl<sub>2</sub> 1 mM. **A.** ADN PTO 25-25-25, PTO 15-25-35 et sans PTO. **B.** ADN PTO 35-25-15 et dégradation par la DNase I. **C.** Représentation des résultats attendus sur les ADN PTO 15-25-35 et 35-25-15 si A1 a un sens de digestion privilégié, dans cet exemple 3' → 5'. La digestion aurait lieu vers une extrémité contenant 35 ou 15 liaisons PTO. La digestion devrait donc produire une bande majoritaire à 15 bases pour un des ADN et une bande majoritaire à 35 bases pour l'autre. De plus, si A1 peut effectuer une coupure endonucléase de façon aléatoire sur les ADN testés, des oligonucléotides ADN/PTO de différentes tailles pourront être obtenus. En effet, si A1 a un sens de digestion privilégié, elle ne peut dégrader l'ADN que vers l'une des extrémités en laissant l'autre bout intact. Une "trainée" d'ADN devrait alors être observable entre 60 et 35 bases ou entre 40 et 15 bases selon la taille maximale de l'oligonucléotide ADN/PTO possible.

Nous observons une apparition dès 3 heures de l'intermédiaire à 35 bases pour les ADN PTO 15-25-35 et PTO 35-25-15. Une bande à 15 bases devient visible après une nuit de digestion (**Figure 65** A et B). Les quantités d'ADN déposées pour l'échelle de taille sont inversement proportionnelles à la longueur de l'ADN utilisé. Pourtant, le fragment de 15 bases est peu visible en raison de sa faible masse. Il n'est donc pas surprenant que l'intermédiaire à 15 bases soit difficile à voir lors de la réaction. Cet intermédiaire est peut-être déjà présent dès 3 heures

de réaction sans être encore visible sur gel et il semble commencer à apparaître à partir de 6 heures de réaction pour les deux ADN testés.

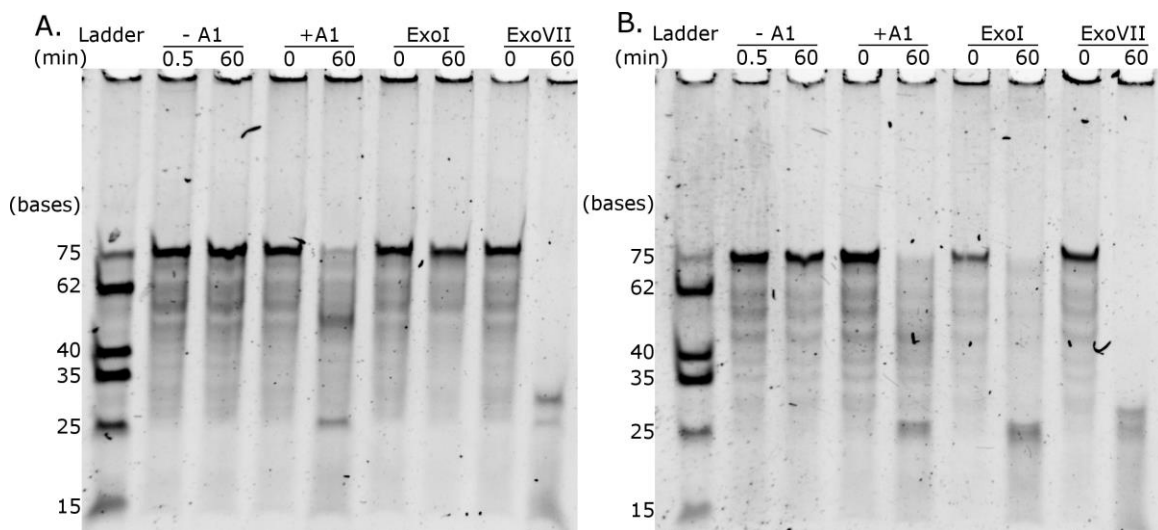
Nous n'observons pas les différences attendues si A1 avait un sens de digestion privilégié dans les profils de digestion des ADN PTO 15-25-35 et PTO 35-25-15 (**Figure 65 C**). Ces résultats sont donc une forte indication d'une activité bidirectionnelle.

Ces résultats suggèrent que A1 a une activité bidirectionnelle, mais ne constituent pas une preuve suffisante.

#### 1.2.5 Comparaison avec d'autres enzymes connues sur PTO 25-50 et 50-25

Pour confirmer que A1 dégrade l'ADN dans les deux directions, j'ai comparé son activité à celles d'enzymes connues : l'exonucléase bidirectionnelle ExoVII<sup>125</sup> et l'exonucléase 3'→5' ExoI<sup>126</sup>.

Les deux ADN testés sont des ADN 25-50 PTO 5' et 50-25 PTO 3'. Ils sont dégradés jusqu'à 25 bases par A1 et ExoVII, alors que l'ADN n'est dégradé que s'il est bloqué du côté 5' pour ExoI (**Figure 66**). A1 semble donc dégrader l'ADN dans les deux directions de la même manière que ExoVII, mais son activité endonucléase pourrait avoir joué un rôle dans l'obtention de ce résultat.



**Figure 66.** Activité de A1 sur des ADN 50-25 PTO 3' (A) et 25-50 PTO 5' (B) bloqués pour la digestion en 5' ou en 3' par 25 bases non hydrolysables. Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN] = 1/250. ExoI et Exo VII : 20 U par réaction.

Bloquer l'extrémité de l'ADN avec une partie non-hydrolysable n'est pas suffisant pour conclure avec certitude sur l'activité bidirectionnelle de A1. Pour mieux comprendre le mode

de digestion de A1, j'ai utilisé des oligonucléotides partiellement hybridés pour imiter des substrats d'hélicase.

#### 1.2.6 *Activité sur des ADN partiellement hybridés*

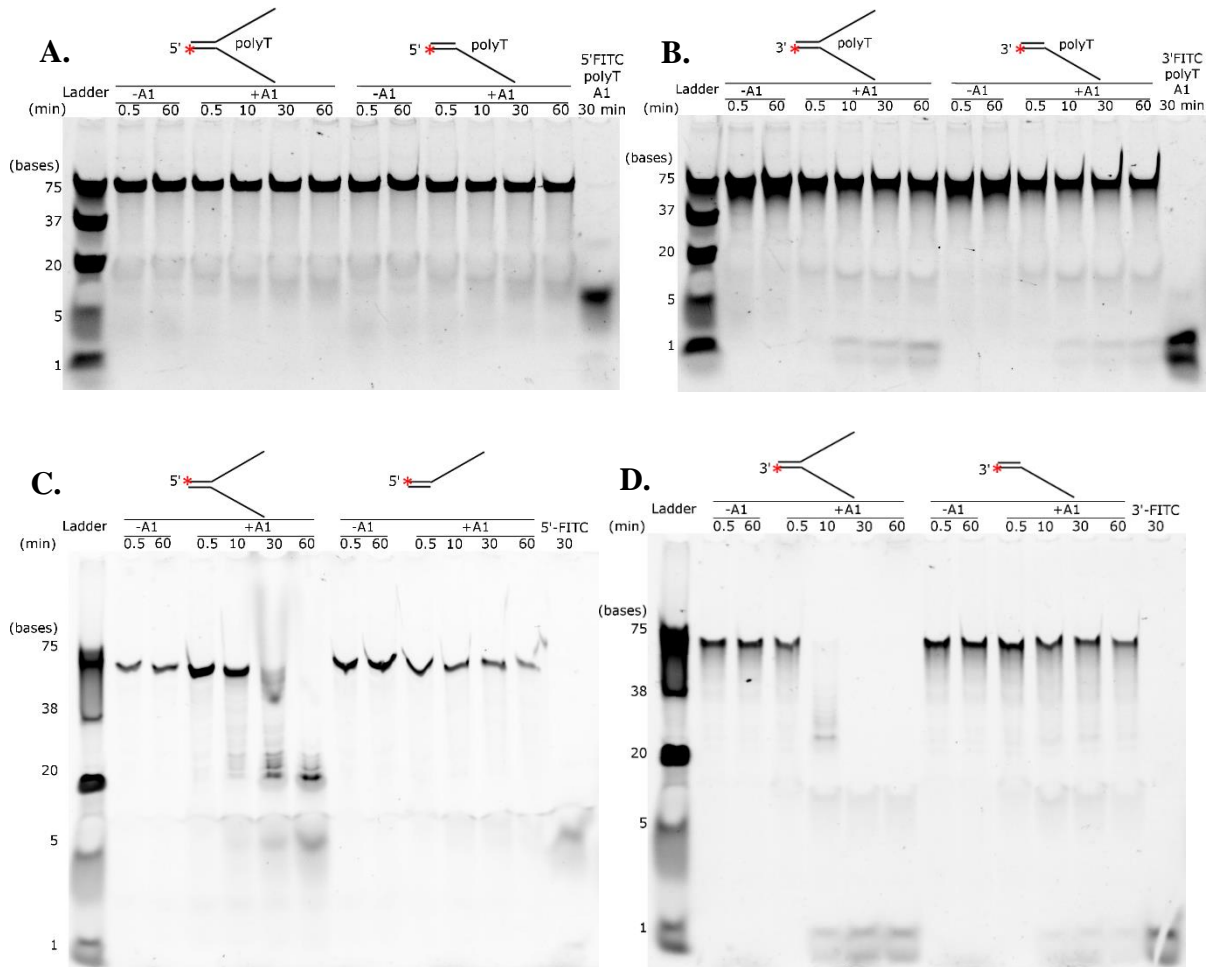
Il est surprenant que *in vitro*, A1 dégrade l'ADN double-brin avec très peu d'efficacité par rapport à l'ADN simple-brin. En effet, si on admet que A1 est directement responsable de la dégradation massive de l'ADN bactérien au début de l'infection par le bactériophage T5, on s'attend à observer la dégradation de l'ADN double brin *in vitro*. L'organisation de l'ADN chromosomique dans la bactérie est plus complexe que celle des ADN double brin de 75 pb utilisés dans mes tests d'activité précédents : cet ADN contient par exemple des fourches de réplication ou des bulles de transcription. Pour savoir si A1 peut envahir les fourches de réplication pour dégrader le chromosome, j'ai utilisé des ADN imitant des substrats d'hélicase.

Dans un premier temps, j'ai utilisé des oligonucléotides partiellement hybridés sur 20 bases et portant une région poly-T de 55 bases. Une activité quasi-nulle est observée sur les oligonucléotides hybridés possédant des fourches ayant une séquence poly-T alors que le contrôle avec de l'ADN simple-brin avec une partie poly-T montre que A1 est active (**Figure 67** A et B). Deux hypothèses sont possibles : A1 ne peut pas dégrader l'ADN simple brin partiellement hybridé ou A1 ne peut pas dégrader les séquences composées de poly-T. Les séquences poly-T sont du côté non marqué dans le contrôle positif : si A1 ne peut pas dégrader des séquences polyT, elle peut quand même dégrader l'ADN composé d'une séquence différente du côté marqué du contrôle positif. La méthode de détection utilisée ici (présence du fluorophore FITC) ne permet pas de détecter si les fragments poly-T sont dégradés complètement, mais seulement de détecter la taille de fragments portant le FITC. Il conviendrait de vérifier par coloration d'ADN total avec du SYBR Gold par exemple, si A1 est capable de dégrader les ADN poly-T.

Dans un deuxième temps, j'ai utilisé des oligonucléotides avec une séquence copiée sur le génome de T5 qui ne sont capables de s'hybrider que partiellement (20 bases sur 75). On observe une dégradation des oligonucléotides imitant des substrats d'hélicase avec un intermédiaire de 20 pb qui correspond probablement à l'ADN double-brin plus difficile à digérer (**Figure 67** C et D). Cette dégradation est totale en 30 min dans le sens 5'→3' alors qu'elle n'est pas terminée au bout de 60 min dans le sens 3'→5'. En effet, le fragment de 20 pb est toujours visible mais a été partiellement dégradé puisqu'il y a apparition d'une bande à 5

bases. En revanche, l'activité est quasi-nulle sur un ADN double brin avec une extension simple-brin de 55 bases. Des intermédiaires de réaction sont visibles avec les substrats d'hélicase pour les deux marquages, ce qui est caractéristique d'une activité exonucléase dans les deux directions.

Puisque A1 peut dégrader entièrement un ADN imitant un substrat d'hélicase, il serait possible qu'elle dégrade le chromosome en envahissant les fourches de réplication.



**Figure 67. Activité de A1 sur des ADN de 75 bases partiellement hybridés (20 bases sur 75).** Un des brins est marqué avec le fluorophore FITC (astérisque rouge). Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN] = 1/125 (0,6 μM A1). Un contrôle de l'activité de A1 avec un ADN simple-brin non-hybridé est effectué pour chaque réaction. Bande à 1 base du marqueur : digestion par ExoI.

Il est surprenant que l'ADN avec une extension simple-brin de 55 bases soit dégradé de manière très peu efficace. On peut se demander si A1 est capable d'interagir avec la région double brin, bloquant ainsi la progression d'autres molécules de A1 sur l'ADN simple-brin. Il est plus difficile pour A1 de se fixer à l'extrémité double-brin d'un ADN imitant un substrat d'hélicase



car il possède 55 paires de bases non complémentaires dont l'appariement est défavorable, ce qui pourrait expliquer que cet ADN soit dégradé.

Pris dans leur ensemble, mes résultats indiquent que A1 a une activité exo- et endonucléase sur l'ADN simple brin ainsi que sur certains ADN partiellement hybridés.

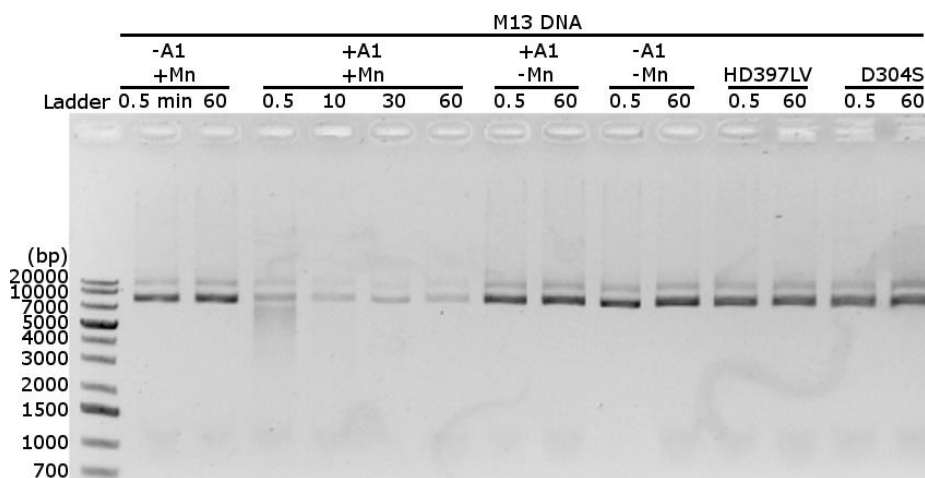
### 1.2.7 Tests sur des ADN double-brin circulaires et sur un ADN chromosomique

Les expériences réalisées jusqu'à présent ont utilisé des fragments d'ADN très courts et linéaires. Afin de poursuivre la caractérisation de l'activité de A1, j'ai également travaillé sur des ADN plus longs ainsi que des formes circulaires.

#### 1.2.7.1 Test d'activité sur des ADN double-brin circulaires

Les tests d'activité faits sur des petits ADN de 75 bases à 37°C ont montré que A1 à 0,3 µM était active sur l'ADN simple-brin et très peu active sur l'ADN double-brin à 1 µM. Cependant, on s'attend à ce qu'elle exerce son activité sur l'ADN double-brin pendant l'infection. À la suite des résultats obtenus avec les substrats d'hélicase, j'ai fait l'hypothèse que A1 pouvait envahir les fourches de réplication et les bulles de transcription du chromosome, mais est-elle capable de dégrader de l'ADN double-brin n'étant pas en cours de réplication ?

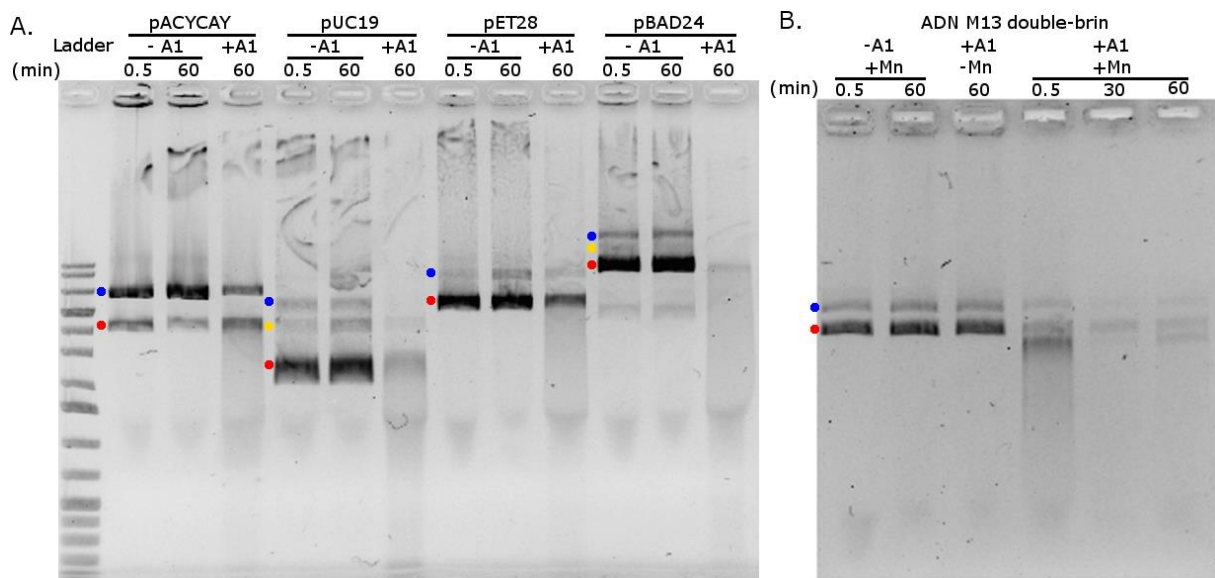
Pour le savoir, j'ai observé l'activité de A1 sur de l'ADN double-brin circulaire du bactériophage M13 (**Figure 68**). A1 commence à dégrader l'ADN de M13 en seulement 30 secondes, puis elle poursuit la dégradation jusqu'à 10 minutes. A1 semble donc pouvoir dégrader l'ADN double-brin circulaire de manière efficace. En revanche, l'ADN n'est pas dégradé davantage au-delà de 10 minutes. A1 pourrait avoir distingué entre différents niveaux de surenroulement de l'ADN.



**Figure 68. Activité de A1 sur l'ADN double-brin circulaire M13mp18 RF I commercial.** Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[pb ADN] = 1/250.

J'ai effectué cette expérience avec les mutants de A1 testés sur les ADN courts (Paragraphe 1.2.1.2) et en absence de  $MnCl_2$ . Dans les deux cas, aucune activité n'a été observée, ce qui confirme les résultats précédents : nous avons identifié deux mutants sans activité DNase et la protéine A1 sauvage n'a pas d'activité sans  $MnCl_2$ .

J'ai reproduit ces expériences avec un autre lot de A1 sur différents ADN double-brin circulaires. On observe une activité dans tous les cas, qui est plus ou moins efficace selon le plasmide utilisé (**Figure 69**). Nous confirmons donc que A1 peut avoir une activité sur des ADN double-brin circulaires. Cette activité est globalement plus lente que celle observée sur des ADN simple-brin de 75 bases et les ADN ne sont jamais totalement dégradés dans nos conditions expérimentales.

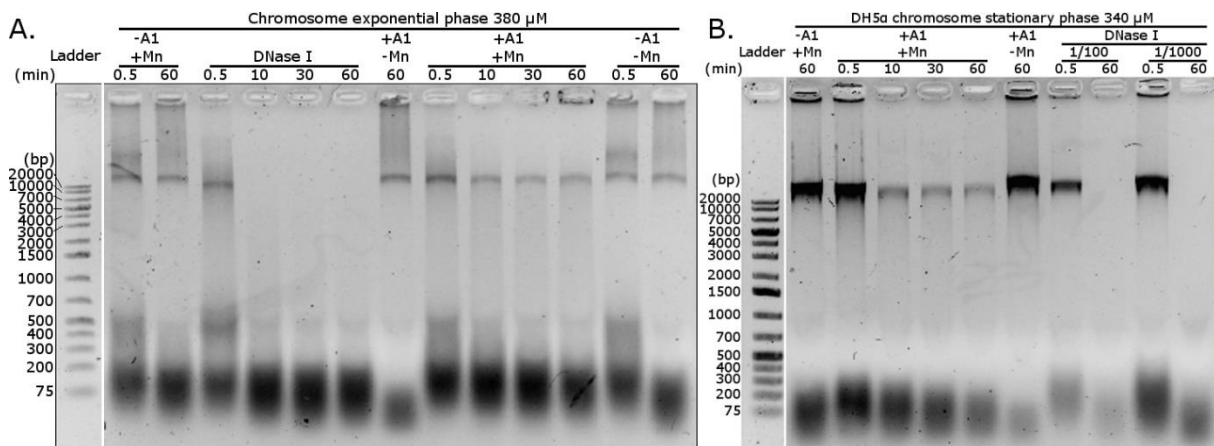


**Figure 69. Activité de A1 sur des ADN plasmidiques double-brin circulaires ou sur l'ADN M13mp18 RF I commercial.** Expérience effectuée à 37°C avec un rapport molaire  $[monomère\ A1]/[pb\ ADN] = 1/250$ . Le lot de A1 utilisé dans cette expérience est différent de celui de la Figure 68. En rouge, forme surenroulée ; en bleu, forme relâchée probable ; en jaune, forme linéaire probable.

### 1.2.8 Test d'activité sur un chromosome extrait de *E. coli*

Nous nous attendons à observer une activité de A1 sur un ADN chromosomique bactérien car c'est cet ADN qui est dégradé pendant l'infection par le bactériophage T5. J'ai testé l'activité de A1 sur un chromosome de *E. coli* extrait à l'aide d'un kit, sans traitement par une protéase et sans traitement par phénol-chloroforme.

L'ADN est partiellement dégradé en 10 minutes par A1, puis la dégradation semble être stoppée (**Figure 70**). Ce résultat ressemble à ce qui avait été obtenu avec l'ADN double-brin de M13 (voir **Figure 68**). Aucune activité n'est obtenue en présence de  $MnCl_2$  et en absence de A1 : il est donc peu probable qu'une DNase contaminant l'échantillon d'ADN chromosomique soit activée dans ces expériences. A1 ne semble pas dégrader les ARN de transfert présents dans l'échantillon, visibles en bas du gel. A1 est donc active sur un ADN chromosomique : cette observation renforce l'hypothèse selon laquelle elle pourrait être la DNase responsable de la dégradation massive du chromosome de l'hôte pendant l'infection par le bactériophage T5. Pour que ces hypothèses soient confirmées il faut répéter l'expérience sur de l'ADN chromosomique extrait en présence de phénol-chloroforme et traité avec la protéinase K pour être certain d'éliminer tous les contaminants bactériens de nature protéique.



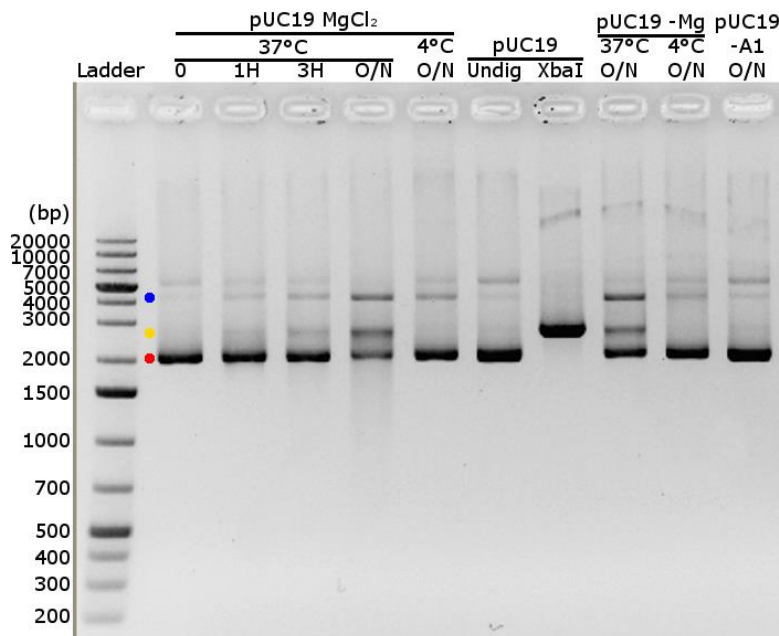
**Figure 70. Activité DNase de A1 sur un chromosome extrait de *E. coli* en phase exponentielle (A) ou en phase stationnaire (B).** La concentration des chromosomes extraits est imprécise à cause de l'ARN contaminant. Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[pb ADN] = 1/250.

### 1.2.9 Mise en évidence d'activités nickase et endonucléase de A1 sans cation divalent

Jusqu'ici, j'ai fait les expériences avec du  $MnCl_2$  car c'est l'ion que nous avons utilisé pour former l'octamère de A1 et résoudre sa structure. J'ai fait un test d'activité avec du  $MgCl_2$ , un autre cation divalent couramment utilisé par les nucléases.

Après une nuit d'incubation du plasmide pUC19 en présence de A1 avec sans magnésium, on observe que la forme surenroulée (bande du bas) est moins intense et on voit clairement une augmentation de l'intensité des bandes correspondant aux formes relâchée et linéaire, respectivement la bande la plus haute et la bande intermédiaire (**Figure 71**). A1 provoque donc des coupures simple et double-brin sur le plasmide. Cette observation est faite en présence ou en absence de magnésium.

A1, sans magnésium ni manganèse, a une activité nickase et endonucléase sur de l'ADN double-brin. Cette activité est lente *in vitro* dans nos conditions expérimentales, mais *in vivo* une forte production de A1 pourrait conduire à une activité beaucoup plus importante. Nous avons donc mis en évidence une activité lente de A1 sur des plasmides, y compris en absence de cation divalent.



**Figure 71. Effet du magnésium sur l'activité de A1 sur un plasmide pUC19 double-brin circulaire.**

Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[pb ADN] = 1/250 (0,6 μM A1). Le plasmide testé montre principalement une bande correspondant à la forme surenroulée lorsqu'il est incubé sans A1. En rouge, forme surenroulée ; en bleu, forme relâchée ; en jaune, forme linéaire.

#### 1.2.10 Discussion sur l'activité de A1 *in vivo*

Des expériences de microscopie de fluorescence ont été effectuées par Ombeline Rossier dans l'équipe en parallèle de mon travail. Des bactéries ont été transformées par un plasmide pBAD24 exprimant A1 sous le contrôle d'un promoteur arabinose. L'ADN est visualisé par marquage au DAPI ou par marquage fluorescent de la protéine HU avec le fluorophore mCherry<sup>127</sup>. HU est une protéine apparentée aux histones qui s'associe au nucléoïde bactérien. En absence de A1, l'ADN forme un nucléoïde au contour nettement défini dans les cellules. Lorsque A1 est exprimée, ce nucléoïde disparaît presque totalement, tout comme le marquage fluorescent de HU. L'ADN des bactéries a été extrait et déposé sur gel et on retrouve significativement moins d'ADN chez des bactéries ayant exprimé A1. Le plasmide pBAD24

semble peu dégradé en comparaison. Enfin, le mutant d'activité HD397LV a été testé et ne montre aucune dégradation significative, même si le phénotype des bactéries est légèrement différent avec un nucléoïde moins bien délimité.

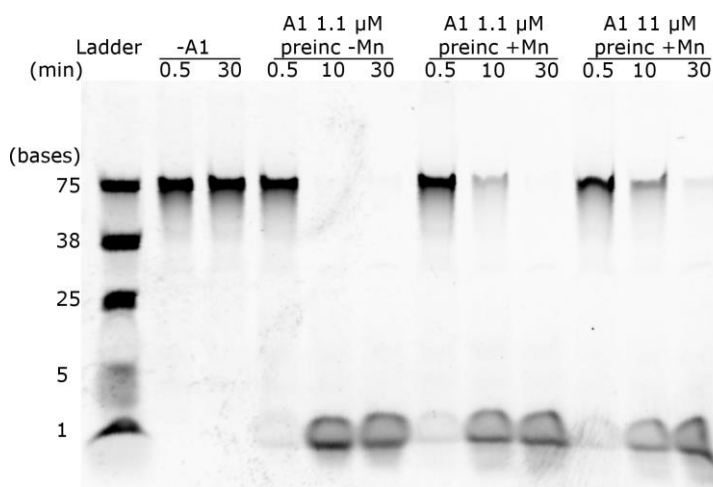
Ces résultats confirment *in vivo* l'activité DNase de A1 que j'ai observée *in vitro*. L'expression du gène viral A1 dans la bactérie ne nécessite donc pas d'autres gènes viraux pour déclencher la dégradation du chromosome. Des partenaires bactériens pourraient participer à cette activité, conduisant à la dégradation rapide du chromosome.

### 1.2.11 Altération de l'activité de A1 en préincubation avec du MnCl<sub>2</sub>

Nous avons vu précédemment que la préincubation de A1 avec du MnCl<sub>2</sub> diminuait l'activité de la protéine au lieu de l'augmenter comme nous nous y attendions (**Figure 59**). J'ai reproduit l'expérience en testant l'effet de la concentration de A1 lors de la préincubation.

On observe une plus forte activité quand A1 est préincubée sans MnCl<sub>2</sub> (ajout de MnCl<sub>2</sub> au dernier moment pour démarrer la réaction de digestion de l'ADN) que lorsque A1 est préincubée avec du MnCl<sub>2</sub> (la réaction de digestion est démarrée par l'addition de l'ADN) (**Figure 72**). Plus la concentration de A1 est élevée lorsqu'elle est préincubée avec du MnCl<sub>2</sub>, plus l'activité diminue. Ce résultat est inattendu : si A1 possède déjà le manganèse nécessaire à son activité, elle devrait pouvoir être immédiatement active.

La microscopie électronique à coloration négative permet d'observer directement A1. Nous allons voir comment cette technique nous a apporté les éléments manquants pour comprendre cette perte d'activité en pré-incubation avec du MnCl<sub>2</sub>.



**Figure 72.** Activité de A1 sur l'ADN FITC 3' ss après préincubation sur la nuit avec ou sans MnCl<sub>2</sub>. A1 est incubée à 1,1 μM ou 11 μM en absence ou en présence de 1 mM de MnCl<sub>2</sub>, puis utilisée à la concentration de 0,3 μM dans la réaction. Expérience effectuée à 37°C avec un rapport molaire [monomère A1]/[base ADN]= 1/250. Bande à 1 base du marqueur : digestion par ExoI.

### 1.3 Microscopie électronique à coloration négative

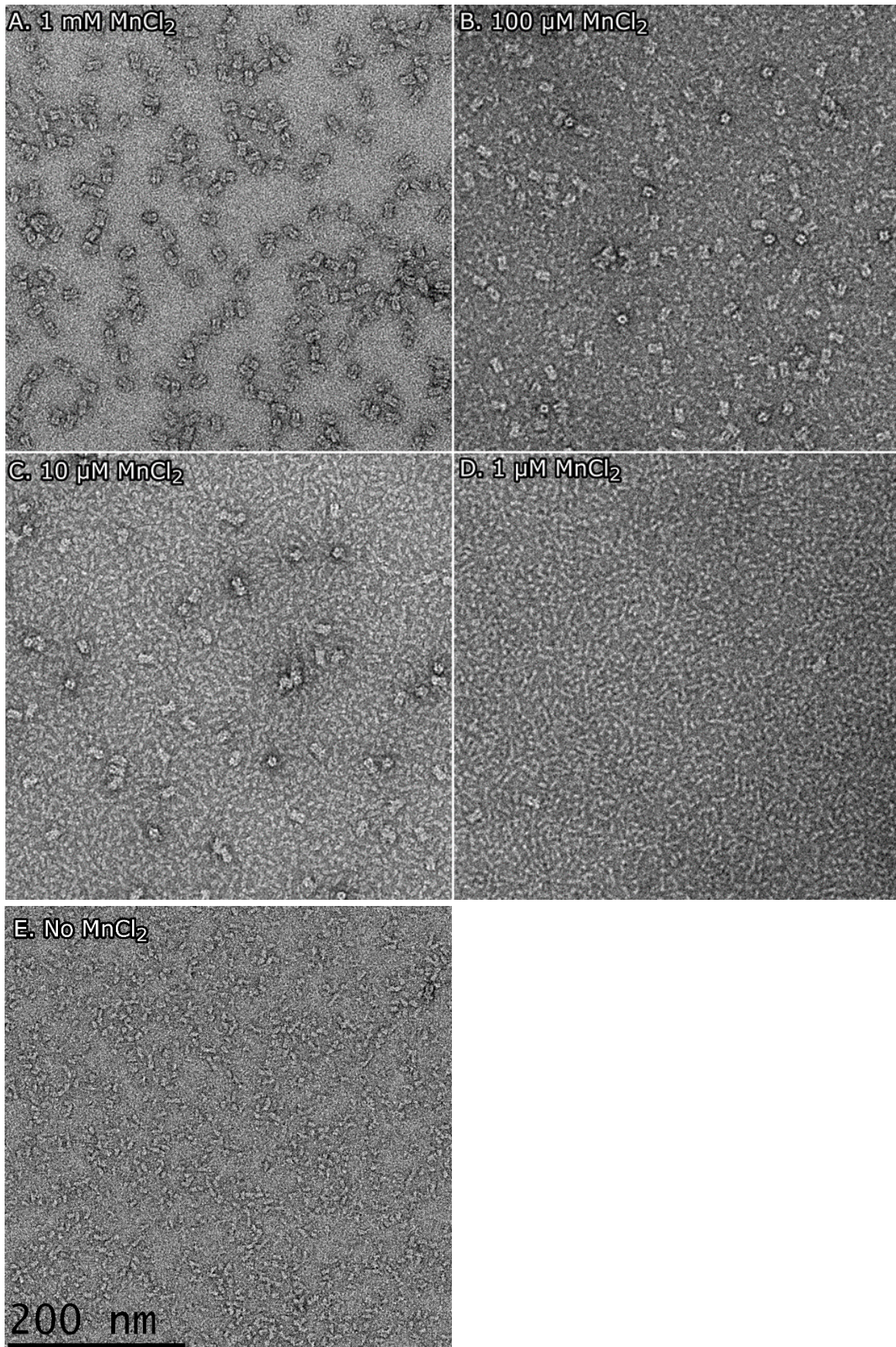
#### 1.3.1 Oligomérisation de A1 en présence de $MnCl_2$

##### 1.3.1.1 Variation de la concentration en A1 et en $MnCl_2$

Les résultats de SEC-MALLS et de SAXS montrent qu'il existe un équilibre entre oligomère et monomère qui varie selon la présence de  $MnCl_2$ . La microscopie électronique à coloration négative est une technique particulièrement adaptée pour observer l'évolution de cet équilibre, car elle permet d'observer directement les protéines. Le monomère de A1 a une masse de 61,7 kDa, ce qui est observable en microscopie.

L'équilibre est complètement déplacé vers la formation de l'octamère lorsque A1 est incubée pendant la nuit en présence de  $MnCl_2$  dans la glace (**Figure 73** A) alors qu'elle reste monomérique en absence de  $MnCl_2$  (**Figure 73** E). Lorsqu'on applique des concentrations variables de  $MnCl_2$ , on observe une variation de l'équilibre entre le monomère et l'octamère. A 100  $\mu M$  de  $MnCl_2$ , nous observons autant de monomères que d'oligomères de A1. Quand la concentration de  $MnCl_2$  diminue à 10  $\mu M$  puis 1  $\mu M$  l'équilibre est déplacé progressivement vers le monomère (**Figure 73** B-D).

Bien que la concentration en  $Mn^{2+}$  dans la cellule soit de l'ordre de 10  $\mu M$ <sup>128</sup>, il est préférable d'effectuer les manipulations de microscopie à 1 mM de  $MnCl_2$  car c'est la concentration nécessaire à l'oligomérisation totale de A1 *in vitro*. Puisque le manganèse est important pour la formation de l'octamère, est-il possible de déplacer à nouveau l'équilibre vers la formation de monomère en présence d'EDTA ?



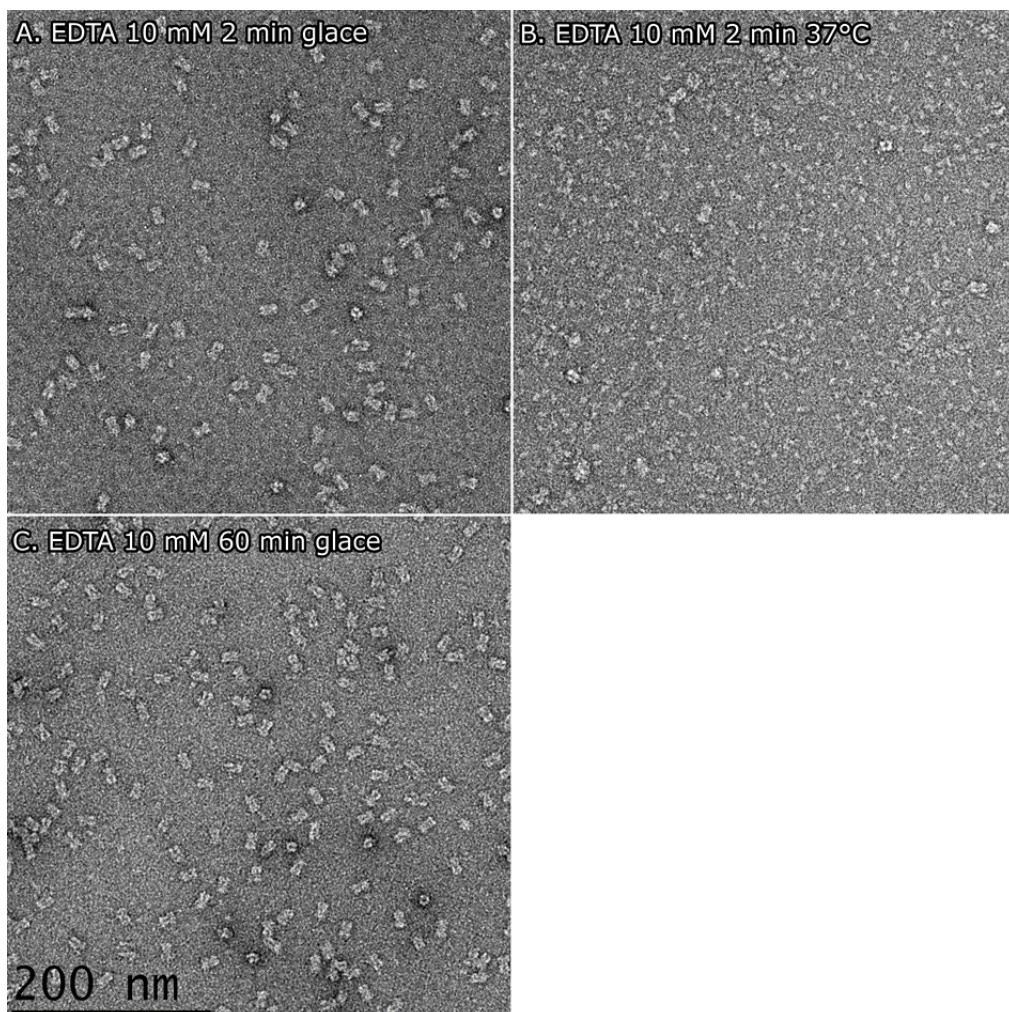
**Figure 73.** Observation par microscopie électronique en coloration négative de la protéine A1 (0,6 μM) pré-incubée pendant une nuit à 4°C, en présence de différentes concentrations de MnCl<sub>2</sub>.

### 1.3.1.2 Dissociation de l'octamère en EDTA 10 mM

A1 a été préincubée avec 1 mM de  $MnCl_2$  pour provoquer son oligomérisation, puis nous avons utilisé 10 mM d'EDTA pour voir s'il était possible de dissocier l'octamère en séquestrant le manganèse. Après 2 minutes d'incubation dans la glace, l'oligomère était intact alors que l'équilibre a fortement été déplacé vers le monomère après 2 minutes d'incubation à 37°C (**Figure 74** A et B). Après 60 minutes d'incubation dans la glace, l'oligomère était toujours intact (**Figure 74** C).

Il est donc possible de dissocier l'octamère de A1 avec de l'EDTA, mais seulement à 37°C. La structure de l'octamère pourrait être stable à 4°C en absence de manganèse, ou l'affinité de A1 pour le manganèse pourrait être plus forte que celle de l'EDTA à 4°C.

A1 étant une DNase, nous avons ensuite observé en microscopie électronique son comportement en présence d'ADN.



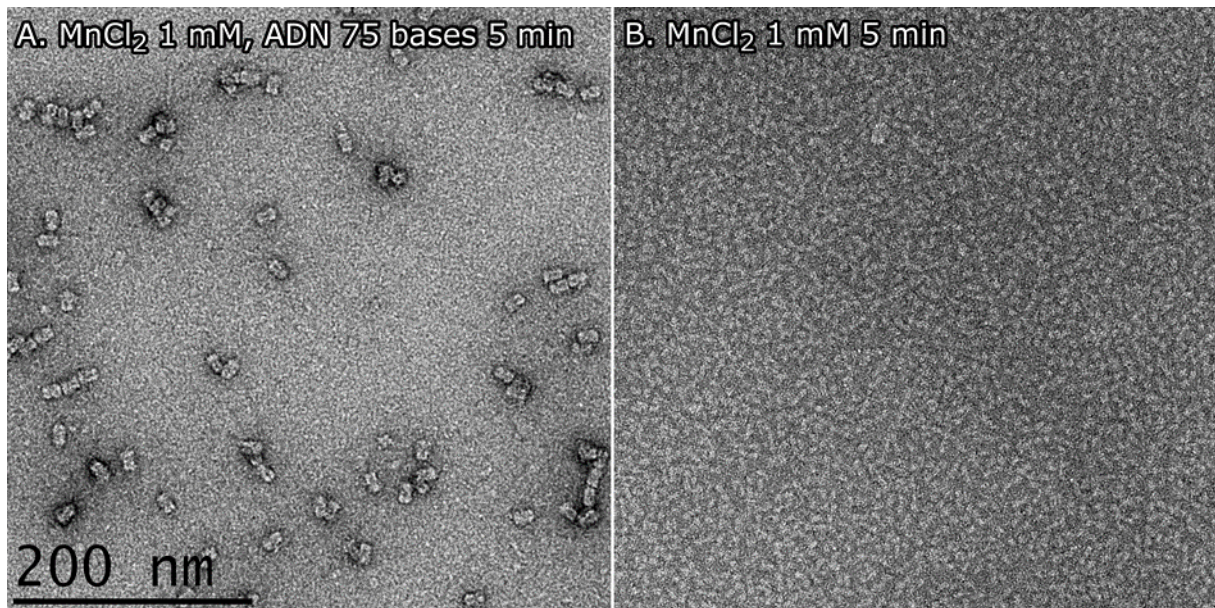
**Figure 74.** Observation de A1 (0,6  $\mu M$ ) pré-incubée en présence de 1 mM  $MnCl_2$  pendant une nuit à 4°C, puis en présence de 10 mM EDTA dans la glace ou à 37°C.



### 1.3.2 Incubations avec l'ADN

#### 1.3.2.1 Comparaison de la vitesse d'oligomérisation avec ou sans l'ADN

Pour étudier le comportement de A1 avec l'ADN par microscopie, nous avons fait une incubation dans les mêmes conditions que pour les tests d'activité, dans la glace. A1 est observée sous forme oligomérique après seulement 5 minutes d'incubation en présence d'un ADN simple-brin de 75 bases et de 1 mM de  $\text{MnCl}_2$  (**Figure 75 A**). En comparaison, A1 est encore sous forme de monomère si elle est incubée dans les mêmes conditions sans ADN (**Figure 75 B**).



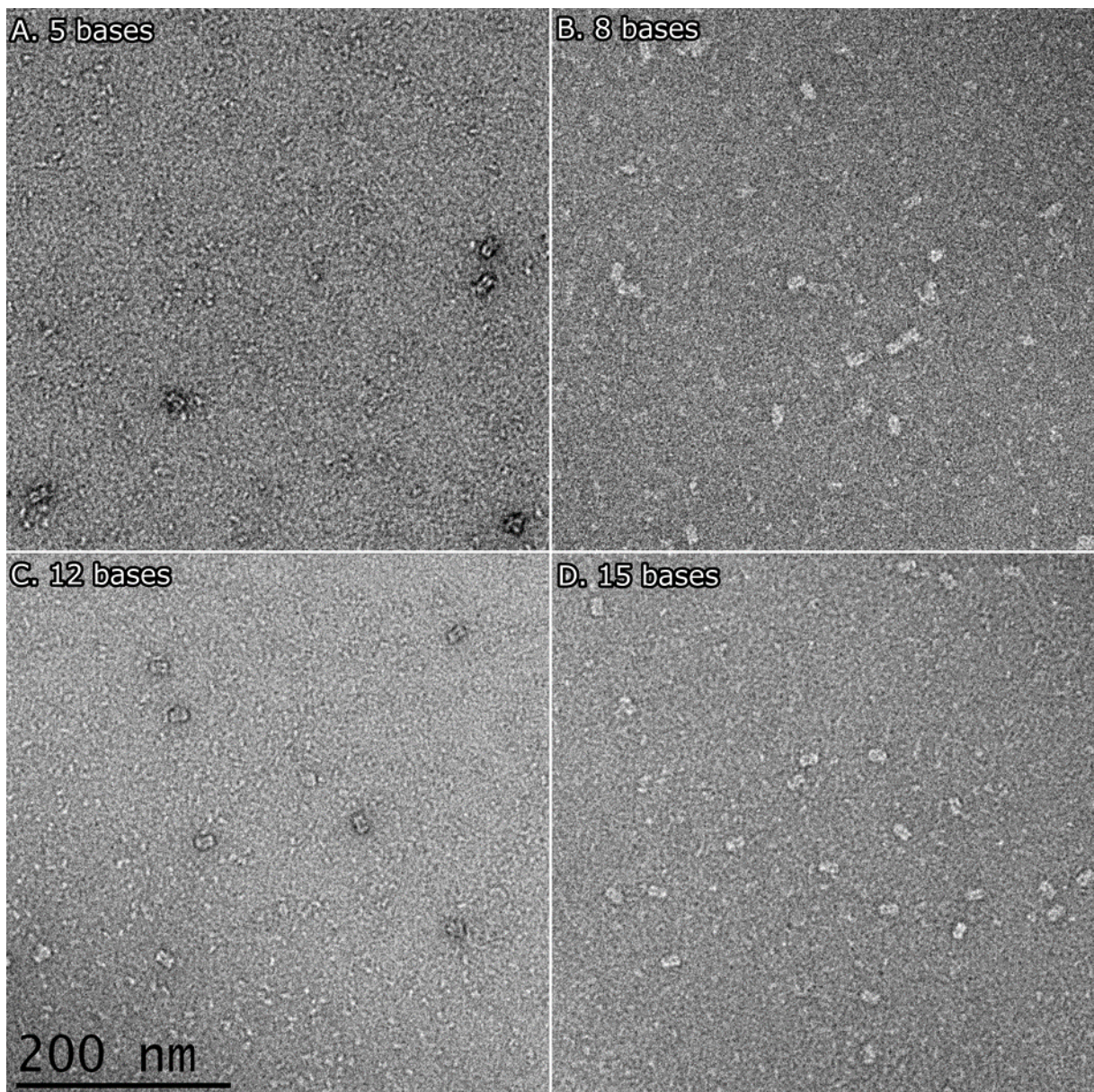
**Figure 75.** Observation de A1 (0,6  $\mu\text{M}$ ) incubée pendant 5 minutes dans la glace, en présence de 1 mM de  $\text{MnCl}_2$  avec ADN (**A**) ou sans ADN (**B**).

L'équilibre est donc rapidement déplacé vers la formation de l'oligomère lorsque A1 est en présence d'ADN et de  $\text{MnCl}_2$ , de manière bien plus efficace qu'en absence d'ADN. Cette observation suggère que la forme active de A1 serait la forme oligomérique. Cependant, nos tests d'activité en solution ont montré que A1 préincubée en  $\text{MnCl}_2$  est majoritairement oligomérique et peu active si elle est ajoutée à l'ADN dans ces conditions. C'est donc la capacité de A1 à former un octamère qui semble indispensable à l'activité DNase de la protéine.

### 1.3.2.2 Taille minimale d'ADN permettant l'oligomérisation

Pour mieux comprendre le mécanisme d'action de A1, nous avons cherché à partir de combien de bases d'ADN nous pouvions rapidement induire la transition vers l'état oligomérique.

Nous avons incubé A1 en présence de 1 mM de  $MnCl_2$  avec différentes molécules d'ADN simple-brin de 5 à 15 bases pendant 30 minutes dans la glace, puis nous l'avons observée en microscopie (**Figure 76**). Les oligomères commencent à se former avec un ADN d'au moins 8 bases. On observe davantage d'oligomères avec un ADN de 15 bases, mais l'équilibre n'est pas totalement déplacé vers la formation de l'octamère.



**Figure 76.** Tests sur la longueur minimale d'ADN permettant l'oligomérisation de A1. A1 0,6  $\mu M$ , incubation 5 minutes dans la glace avec 1  $\mu M$  d'ADN.

Il faut donc un ADN de plus de 5 bases pour permettre l'oligomérisation de la protéine. À 8 bases, la longueur n'est cependant pas suffisante pour déclencher l'oligomérisation de toutes les molécules de A1.

Jusqu'à présent, tous nos tests d'incubation A1/ADN pour les observations en microscopie électronique ont été faits dans la glace. Mais les tests d'activité en solution sont faits à 37°C, il est donc important d'observer le comportement de A1 dans les mêmes conditions que lors d'une cinétique de dégradation de l'ADN à 37°C.

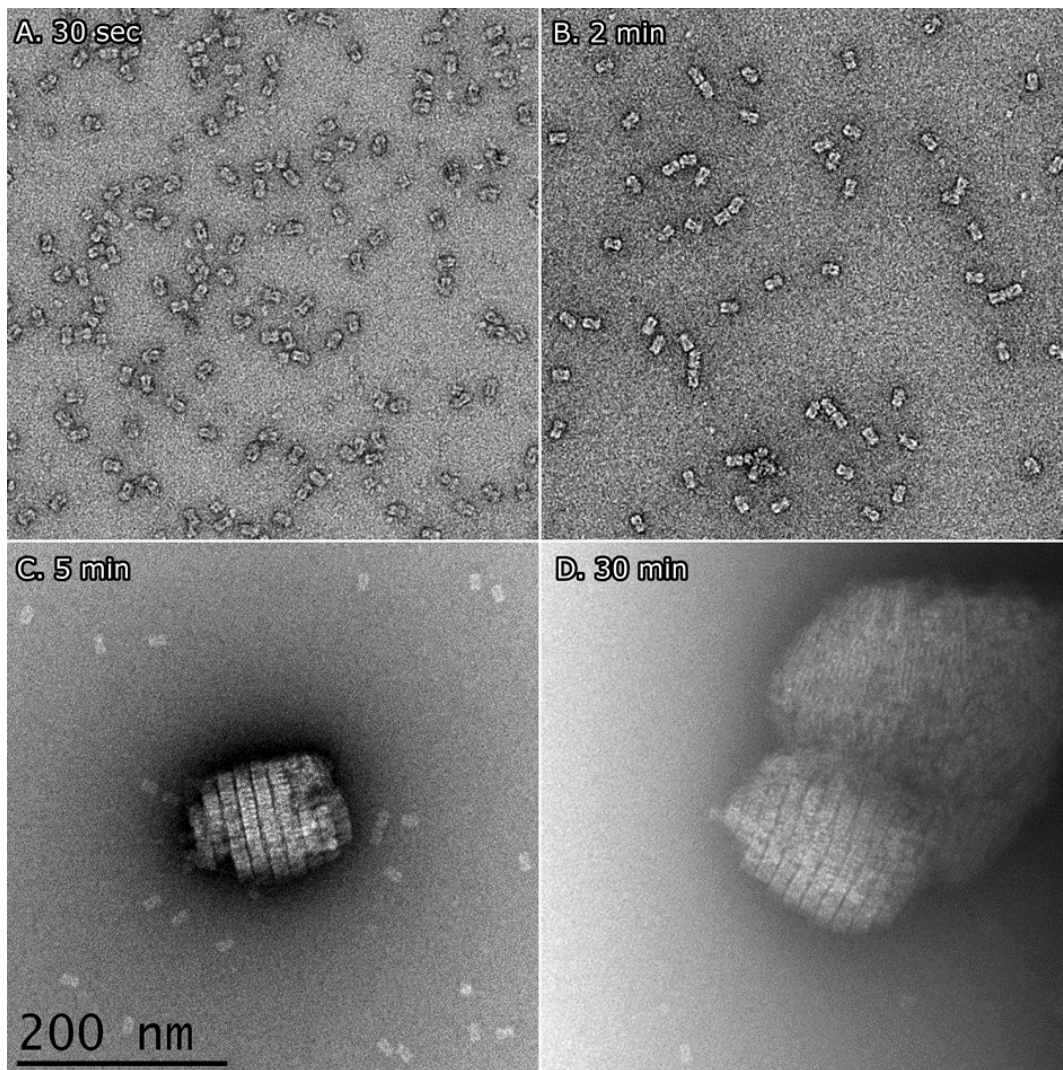
#### 1.3.2.3 Observation de A1 pendant la dégradation de l'ADN

Nous avons incubé A1 en présence d'ADN simple-brin dans les mêmes conditions que pour les tests d'activité et nous avons déposé les échantillons sur gel pour pouvoir comparer les résultats d'activité avec les images obtenues.

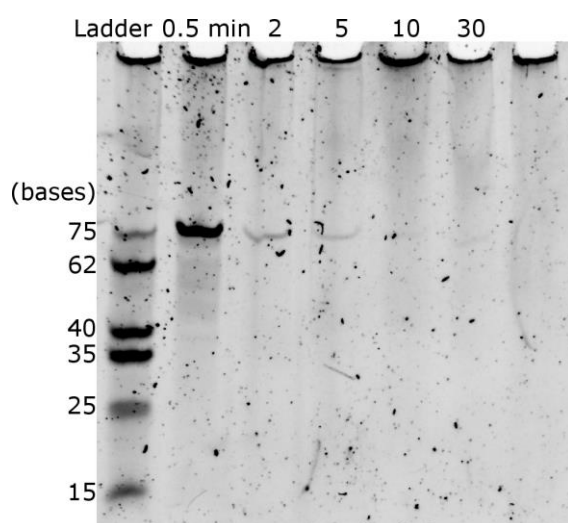
En 30 secondes, de nombreux oligomères sont déjà observables (**Figure 77 A**). À partir de 2 minutes, ces oligomères commencent à former des structures composées de 2 ou 3 octamères alignés dans le sens de la longueur (**Figure 77 B**). Ces structures pourraient être causées par la liaison de plusieurs octamères sur une même molécule d'ADN. La concentration de A1 sur la grille est plus faible qu'à 30 secondes. A 5 minutes, les oligomères sont rares sur la grille. Un examen approfondi révèle que A1 s'ordonne pour former des cristaux de protéine qui pourraient être désordonnés (**Figure 77 C**). A 30 minutes, il n'y a presque plus d'oligomères isolés sur la grille et les cristaux sont plus gros qu'à 5 minutes (**Figure 77 D**).

Nous avons cherché ces cristaux en absence d'ADN, mais l'observation de la grille n'a révélé que la présence d'agrégats de protéine désorganisés. Il est possible que ces structures existent en plus faible nombre en absence d'ADN et que nous ne les ayons pas trouvées. L'ADN semble donc favoriser l'organisation de A1 en cristaux à 37°C et il pourrait être indispensable à leur formation.

On observe la dégradation de l'ADN sur gel dès 2 minutes d'incubation (**Figure 78**). L'ADN disparaît totalement du gel à 10 minutes d'incubation. Les puits sont marqués mais il ne s'agit pas d'ADN car le dernier puits n'en contient pas, il s'agit donc bien d'une dégradation mais pas d'un retard de l'ADN par A1. Nous confirmons donc que malgré la formation de ces structures cristallines, A1 a bien été active pendant les observations de microscopie.



**Figure 77.** Observation de A1 pendant la dégradation de l'ADN. Incubation de A1 à  $0,6 \mu\text{M}$  avec  $1 \text{ mM}$  de  $\text{MnCl}_2$  et  $1 \mu\text{M}$  d'un ADN simple-brin de 75 bases à  $37^\circ\text{C}$ .



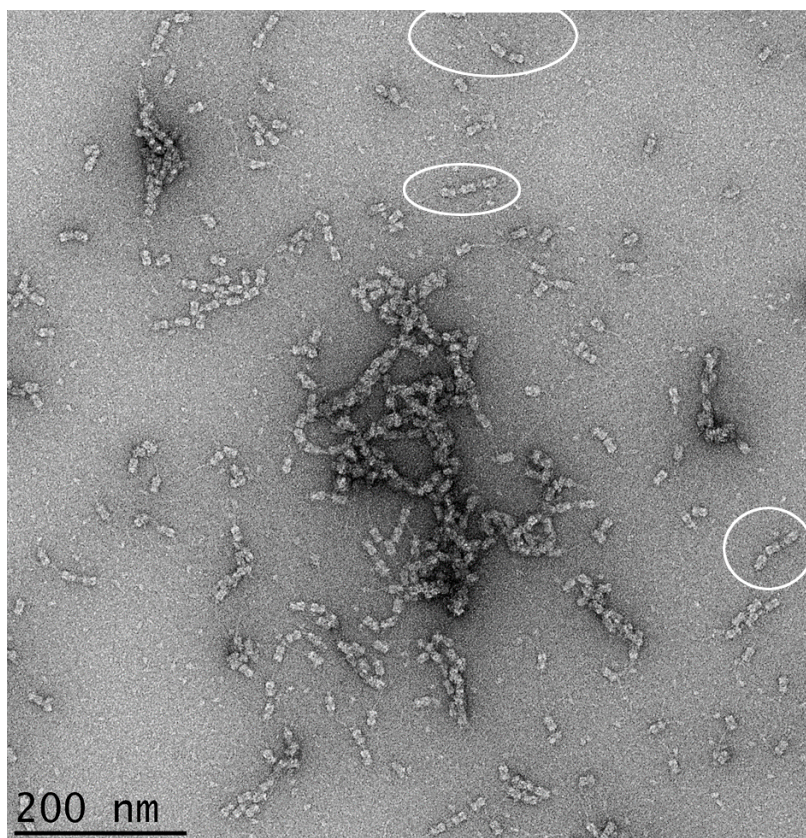
**Figure 78.** Migration des ADN observés en microscopie en [Figure 77](#). Expérience effectuée à  $37^\circ\text{C}$  avec un rapport molaire  $[\text{monomère A1}]/[\text{pb ADN}] = 1/125$ .

La formation de ces cristaux explique la disparition des oligomères de A1 sur la grille.

D'après sa structure, la surface extérieure de l'octamère formé par A1 est fortement chargée négativement dans une large partie centrale (paragraphe 1.1.6.5, **Figure 53**), alors que la charge globale est moins marquée aux extrémités et l'entrée du canal est fortement chargée positivement. Bien que ce potentiel électrostatique de surface de l'octamère ne semble pas favorable à la formation d'un réseau cristallin (en raison d'une répulsion probable entre octamères), on ne peut pas exclure que la partie N-terminale manquante dans la structure soit impliquée dans l'organisation des cristaux. L'analyse de sa séquence révèle que cette partie est fortement chargée positivement. Par sa flexibilité, le domaine N-terminal manquant pourrait être exposé à l'extérieur de l'octamère et ainsi permettre une interaction électrostatique favorable avec un octamère voisin.

#### 1.3.2.4 Incubation avec un ADN long

Pour étudier le comportement de A1 avec un ADN plus long, nous avons incubé A1 avec un ADN double-brin de 500 bases pendant 15 minutes à 37°C (**Figure 79**). A1 forme des oligomères qui semblent se suivre sur l'ADN, avec parfois une formation de plusieurs rangées de protéines collées les unes contre les autres. L'utilisation d'un ADN plus long a empêché la formation de cristaux, mais on retrouve une organisation similaire à ce qui avait été observé avec un ADN de 75 bases. Les brins d'ADN sont visibles : ils semblent parfois sortir du canal formé par A1. Bien que l'on ne puisse pas exclure qu'il s'agisse d'une simple co-localisation de l'ADN avec la protéine, l'observation de ce type de structures en "collier de perles" est fréquente dans l'ensemble des images que nous avons enregistrées.



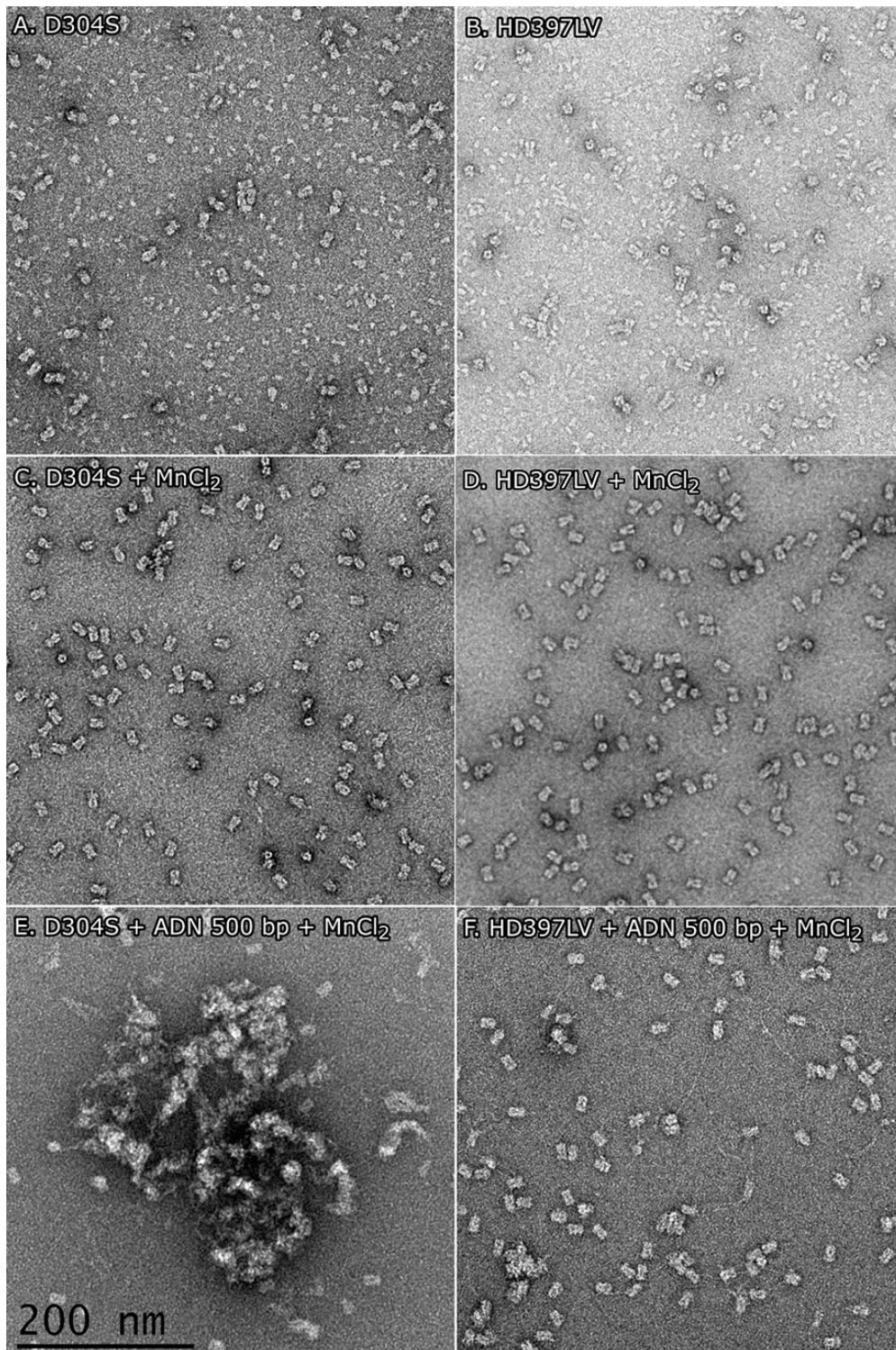
**Figure 79.** Observation de A1 (0,6  $\mu\text{M}$ ) incubée pendant 15 minutes à 37°C avec un ADN double-brin de 500 paires de bases en présence de 1 mM de  $\text{MnCl}_2$ . Certaines structures en collier de perles sont indiquées par des cercles blancs.

### 1.3.3 Comportement des mutants pour l'activité DNase

Nous avons construit plusieurs mutants de A1 inactifs pour la digestion de l'ADN simple-brin. Afin de savoir si ces mutants adoptent une structure similaire à celle de la protéine sauvage, nous avons observé deux d'entre eux : D304S et HD397LV.

On observe que ces deux mutants forment spontanément des oligomères en absence de manganèse et d'ADN (**Figure 80** A et B) à une concentration de A1 (0,6  $\mu\text{M}$ ) pour laquelle on n'observe que des monomères dans le cas de A1 sauvage. Cette oligomérisation est plus importante pour le mutant HD397LV que pour le mutant D304S. Lorsqu'ils sont pré-incubés avec 1 mM de  $\text{MnCl}_2$  pendant une nuit, les deux mutants montrent un déplacement total de l'équilibre vers la formation d'octamères (**Figure 80** C et D). La fixation du manganèse par la protéine est donc probablement encore possible, y compris pour le mutant D304S qui a perdu l'Aspartate 304 impliquée dans la coordination d'un ion  $\text{Mn}^{2+}$ . Après 5 minutes d'incubation avec 1 mM de  $\text{MnCl}_2$  et un ADN de 500 bases dans la glace, le mutant D304S adopte un comportement similaire à celui de la protéine sauvage (**Figure 80** E) : on observe une oligomérisation totale et une agrégation des protéines avec l'ADN. Les oligomères formés par

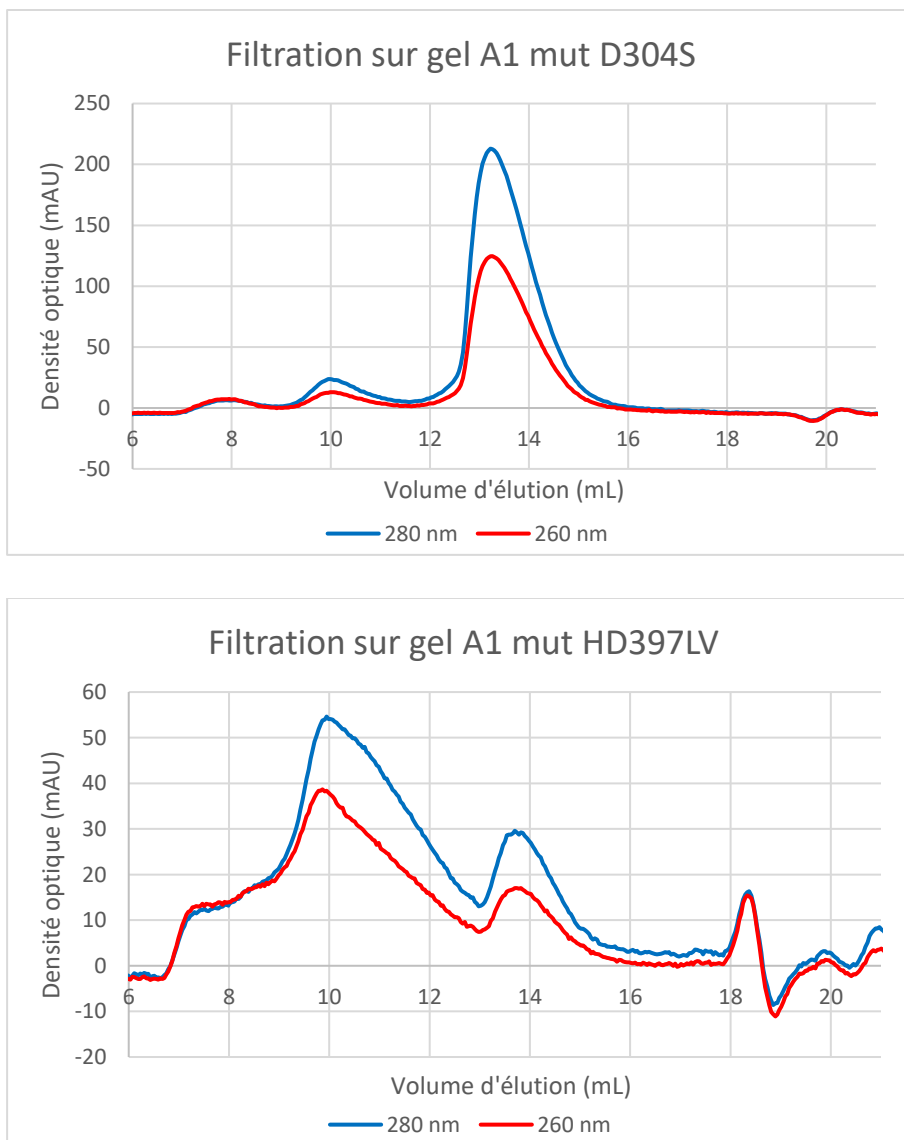
le mutant HD397LV sont colocalisés avec l'ADN, mais nous n'observons pas d'agrégation de l'ADN (**Figure 80 F**).



**Figure 80.** Observation du comportement des mutants A1D304S et A1HD397LV (0,6  $\mu$ M), inactifs pour la digestion de l'ADN, dans différentes conditions. **A**, **B** : sans MnCl<sub>2</sub>. **C**, **D** : avec 1 mM MnCl<sub>2</sub>. **E**, **F** : avec ADN db 500bp 0,15  $\mu$ M + 1 mM MnCl<sub>2</sub>

Pour confirmer les résultats observés en microscopie électronique, j'ai analysé l'état oligomérique des versions mutantes de la protéine A1 sans manganèse sur une colonne de filtration sur gel.

A1 D304S sort majoritairement à un volume d'élution de 13,5 mL, ce qui correspond au monomère. Ce résultat ne correspond pas à ce qui est observé sur la grille : l'octamère se fixe-t-il mieux sur la grille que le monomère ? L'état oligomérique sans  $MnCl_2$  est-il dissocié sur la colonne ? A1 HD397LV est élué dans un pic majoritaire à 10 mL qui correspond à l'octamère, mais ce pic est fortement dissymétrique ce qui est en faveur d'une polydispersité (**Figure 81**). Ce comportement indique probablement que l'octamère se dissocie au cours de l'élution en oligomères de plus petites tailles jusqu'au monomère.



**Figure 81.** Analyse de l'état d'oligomérisation de mutants d'activité DNase de A1 par filtration sur gel. Colonne Superdex 200 10/300 GL. Tampon Tris 50 mM pH 7,4, NaCl 500 mM.

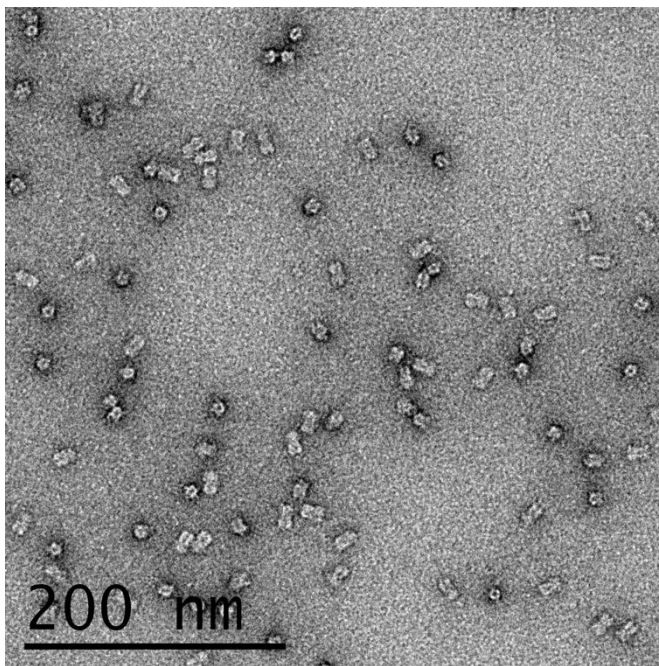


Ces résultats confirment que les mutants d'activité testés ont une structure et une organisation oligomérique similaires à celles de la protéine sauvage. En plus de la perte d'activité DNase, le mutant HD397LV ne provoque pas l'agrégation de l'ADN.

#### 1.3.4 L'oligomérisation *in vivo*

Pour tester si l'oligomère que nous observons en microscopie existe *in vivo*, nous avons extrait A1 en absence d'EDTA et fait une purification classique sur une colonne His Trap FF en présence d'une forte concentration de NaCl pour éliminer l'ADN puis sur une colonne d'héparine. L'objectif était de ne pas dissocier les octamères s'ils se forment *in vivo* en présence du  $Mn^{2+}$  intracellulaire et de les observer sans ajout d'ADN ni de  $MnCl_2$ .

Des oligomères sont observables sur la grille, sans aucun monomère visible (**Figure 82**). Il n'y a pas d'ADN ni de  $MnCl_2$  dans la solution : ces oligomères sont donc probablement des structures formées *in vivo* qui ont résisté aux étapes de purification. L'octamère de A1 semble donc bien se former *in vivo*.



**Figure 82.** Observation de A1 (environ 0,6  $\mu$ M) extraite et purifiée sans traitement à l'EDTA.

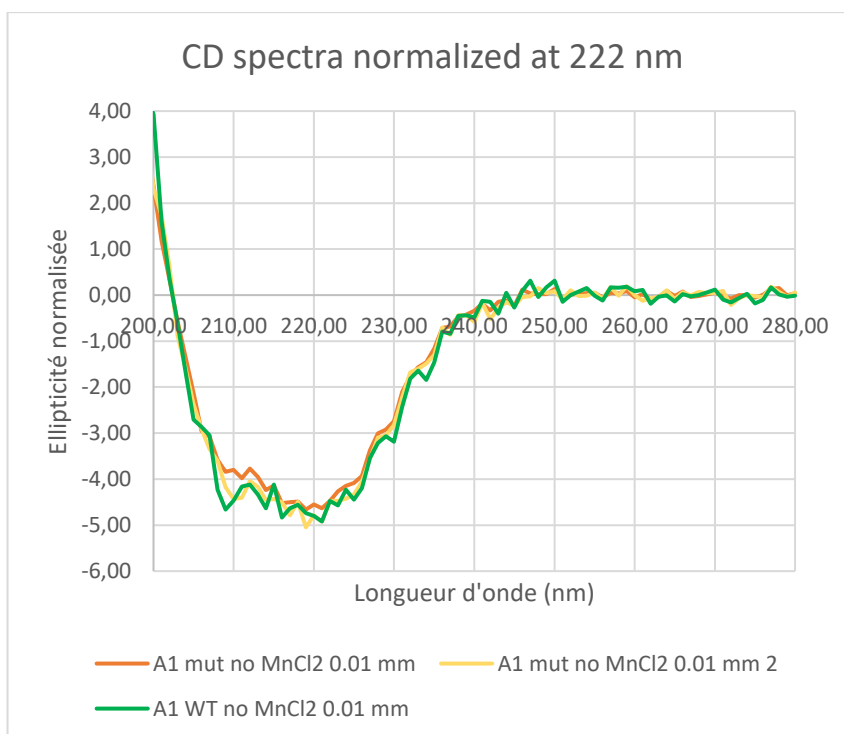
## 1.4 Rôle de l'activité DNase de A1 dans l'infection : Tests de complémentation avec A1

### 1.4.1 Dichroïsme circulaire

Pour pouvoir tirer des conclusions sur des tests de complémentation comparant les versions sauvage et mutante de la protéine A1, il faut s'assurer que le mutant testé ne présente pas de défaut de repliement. En effet, une protéine déstructurée ne devrait pas pouvoir compléter. Les expériences de microscopie suggéraient déjà que la structure des mutants D304S et HD397LV n'était pas significativement modifiée par rapport à la structure de la protéine sauvage (Paragraphe 1.3.3).

Les expériences de dichroïsme circulaire qui ont servi à déterminer les proportions d'hélices  $\alpha$  et de feuillets  $\beta$  de A1 (voir paragraphe 1.1.4) permettent également de comparer la structure secondaire des versions sauvage et mutante de A1. J'ai analysé la protéine sauvage et le mutant A1 HD397LV. Pour pouvoir comparer les protéines, j'ai fait une normalisation des spectres de CD obtenus sur A1 sauvage et mutante. La valeur à 222 nm de A1 mutante sans  $MnCl_2$  est prise comme référence. Seules les valeurs avec un trajet optique de 0,01 mm sont utilisées.

Les courbes sont globalement superposées à partir de 200 nm (**Figure 83**). A1 WT et A1 HD397LV ont donc bien des structures secondaires voisines. Il est donc possible de faire des tests de complémentation comparant la protéine A1 sauvage et le mutant HD397LV.

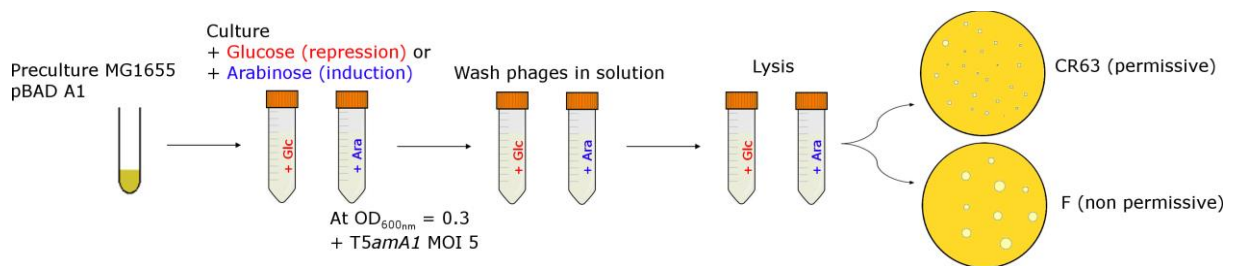


**Figure 83. Spectres de dichroïsme circulaire de A1 sauvage et mutante HD397LV.** Une normalisation des courbes à 222 nm a été faite.

### 1.4.2 Test de complémentation

Pour tester si l'activité DNase de A1 est essentielle pour l'infection par le bactériophage T5, j'ai effectué des tests de complémentation en utilisant des bactériophages portant une mutation *ambre* sur le gène *A1*. Pour la complémentation fonctionnelle, j'ai utilisé des plasmides portant la version sauvage du gène *A1* dans ses deux versions possibles (longue et courte, voir paragraphe I.2.3.2) et un plasmide portant le gène mutant *A1*H<sub>397</sub>LV inactif pour sa fonction DNase. Toutes les expériences ont été faites en triplicata (**Figure 84**).

Le schéma de la **Figure 98** récapitulant le principe de l'expérience est recopié ici :

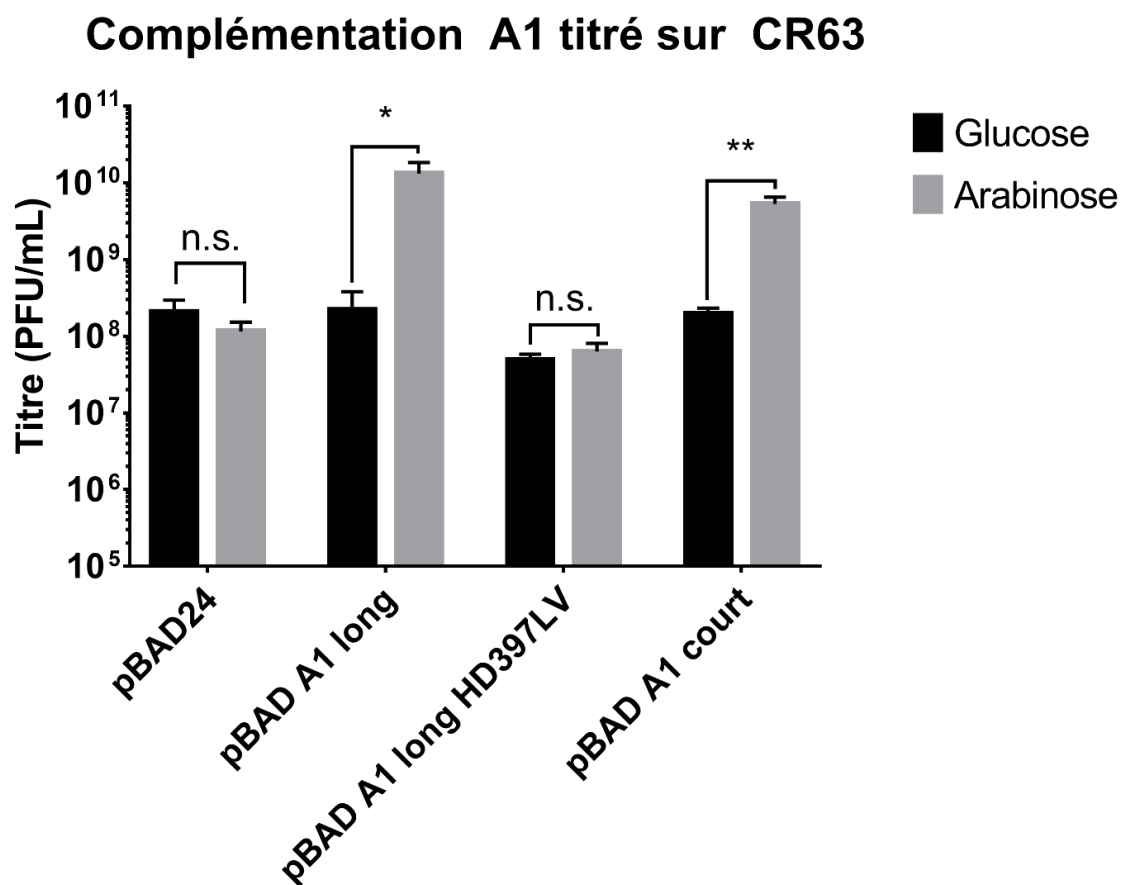


Les phages obtenus après amplification sur la souche MG1655 sont titrés sur une souche CR63 suppressive des codons ambre, ce qui permet de compter les phages mutants et sauvages obtenus (**Figure 84**). On observe que la complémentation avec les versions sauvages longue et courte de *A1* donne des titres moyens plus élevés respectivement d'un facteur 60 et 27 lorsque l'expression du gène a été induite en présence d'Arabinose que lorsqu'elle a été réprimée en présence de Glucose. Ce n'est pas le cas des autres constructions testées. Des tests statistiques sont nécessaires pour vérifier si cette variation est significative.

Analyses statistiques des résultats : En supposant que le nombre de plages de lyses obtenues suive une loi normale, on peut effectuer un test de Student pour vérifier que les moyennes sont différentes. Comme je n'ai fait que des triplicats, je ne peux pas faire le test non-paramétrique de Mann-Whitney car la p-valeur ne pourrait jamais être inférieure à 10 %. Les données sont appariées, ce qui augmente la puissance du test. J'ai calculé les moyennes des titres lors de la répression en présence de glucose et lors de l'induction en présence d'arabinose. Pour les comparer, j'ai posé comme hypothèse ( $H_1$ ) que la véritable moyenne de l'échantillon après induction de l'expression du gène est supérieure à la moyenne de l'échantillon après répression. En effet, si la protéine testée permet la complémentation, il doit y avoir plus de plages de lyse mesurées lorsque l'expression du gène est induite en arabinose que lorsqu'elle est réprimée en glucose.

Pour la version longue de A1, on trouve une P-valeur de 2,55 % et pour la version courte 0,842 %. Ces moyennes sont donc significativement différentes si on accepte une erreur de 5 % : il faudrait faire plus d'essais pour augmenter la puissance du test. En revanche, ni le contrôle négatif pBAD24, ni le mutant ne donnent une p-valeur indiquant une variation significative de la moyenne. Ce résultat indique que les deux versions testées de A1 WT complémentent, alors que le mutant ne peut pas compléter.

Pour vérifier que les résultats obtenus sont bien dus à une complémentation fonctionnelle, nous devons nous assurer que le taux de recombinaison entre le gène sauvage du plasmide et le gène muté du phage est faible.



**Figure 84. Tests de complémentation avec la protéine A1 sur le phage T5amH27.** Ce phage produit une protéine A1 non fonctionnelle car tronquée au niveau de la Glutamine 238 qui est remplacée par un codon Stop. Le test de complémentation est effectué sur une souche MG1655 *recA*<sup>-</sup> contenant un plasmide pBAD24 exprimant ou non le gène *A1*. Les phages obtenus sont titrés sur une souche CR63, qui possède un ARN de transfert suppresseur des codons ambré et permet au phage d'exprimer une protéine complète. Cette souche permet de titrer tous les phages obtenus, même s'ils sont mutants. Le glucose réprime l'expression du gène cloné dans pBAD, l'arabinose induit cette expression. n = 3. \* = p-valeur < 5 %. \*\* = p-valeur < 1 %. n.s. = non significatif.

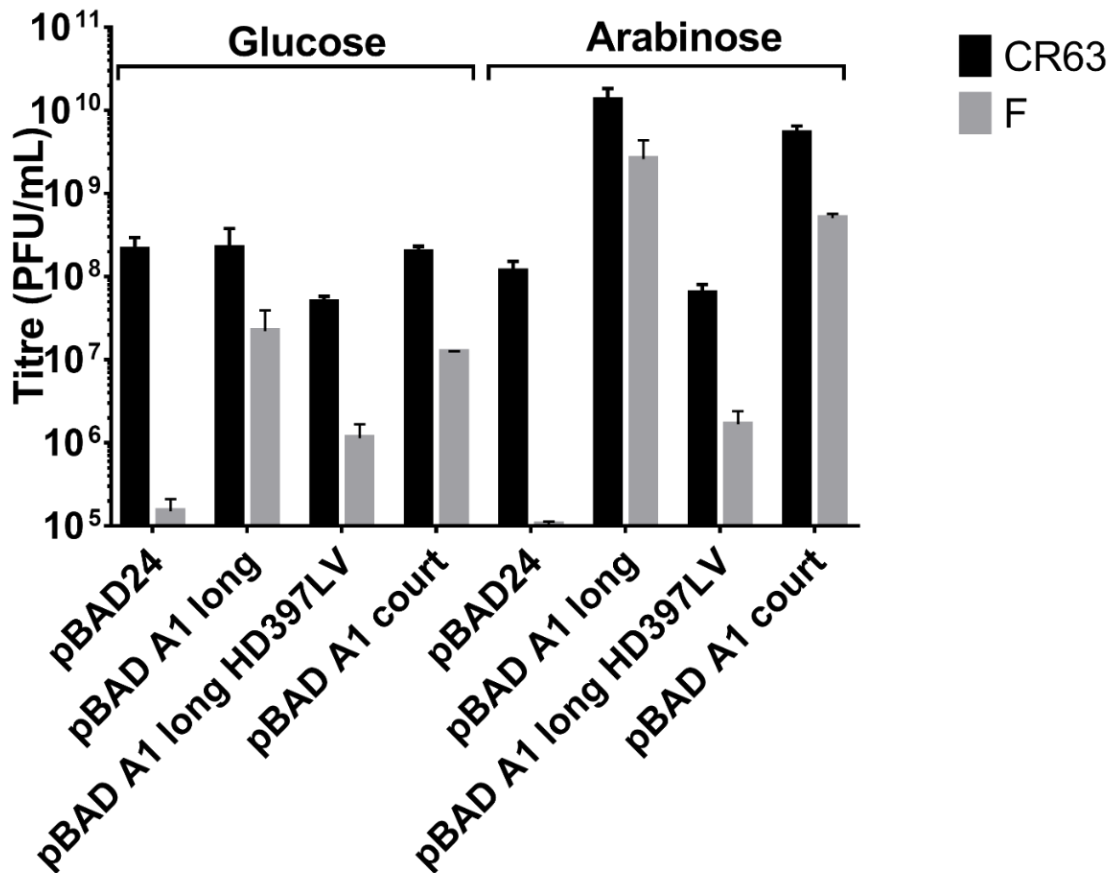
### 1.4.3 Mise en évidence d'une activité recombinase importante

J'ai titré les phages obtenus par complémentation sur les deux souches de *E.coli* CR63 (suppressive des mutations ambre) et F (non suppressive) afin d'évaluer l'importance de la recombinaison génétique dans les résultats obtenus (**Figure 85**). La souche F ne permet de titrer que les phages sauvages et le titre obtenu lorsque l'expérience est faite avec un plasmide vide indique le taux de réversion. Lorsque l'expérience est faite avec un plasmide portant A1, si le titre sur F est supérieur au taux de réversion, alors il y a eu une recombinaison. Si c'est le cas, je cherche à déterminer si l'activité recombinase peut expliquer à elle seule les résultats des expériences de complémentation décrites dans le paragraphe précédent.

Les mêmes tests statistiques que précédemment sont appliqués : on suppose que la distribution du nombre de plages de lyse suit une loi normale pour toutes les constructions. Cette fois, j'ai posé comme hypothèse ( $H_1$ ) que la véritable moyenne de l'échantillon titré sur F est inférieure à la valeur de comparaison qui est l'échantillon titré sur CR63. En effet, nous nous attendons à obtenir une minorité de phages sauvages (issus de la recombinaison, titrés sur la souche F) dans la population totale de phages titrés sur la souche CR63 (T5 sauvages + mutants).

On trouve une p-valeur inférieure à 2 % lorsqu'on compare les moyennes des titres en arabinose sur CR63 et sur F, pour toutes les constructions portant A1. Ces moyennes sont donc bien différentes : le nombre de phages présentant un phénotype sauvage est donc significativement inférieur au nombre total de phages après complémentation. La complémentation observée ne peut donc pas être expliquée par une recombinaison, ce qui valide les résultats précédents. Je confirme donc que l'activité DNase de A1 est essentielle pour l'infection par le bactériophage T5.

## Recombinaison observée en titrant sur F



**Figure 85.** Évaluation de l'activité de recombinaison par titration sur CR63 et F des phages obtenus par complémentation (voir Figure 84). F est une souche non permissive pour l'infection par un phage T5 mutant ambré sur la protéine A1. n = 3.

Les titres observés sur la souche F sont plus élevés avec les versions WT de A1 par rapport à ce qu'on observe avec un plasmide vide. La réversion seule ne peut pas expliquer ces résultats : une recombinaison entre le gène *A1* sur le plasmide et sur le génome du phage a probablement eu lieu.

Lors de la répression de A1 par le glucose, on observe une augmentation significative du titre sur F pour tous les plasmides contenant une version de A1, même mutante. Cette augmentation reflète le taux de recombinaison de base dans la souche MG1655 *recA*<sup>-</sup>, indépendamment de l'expression de *A1*. Une augmentation de la recombinaison d'un facteur 117 et 41 est observée respectivement pour les variantes de A1 sauvage longue et courte lors de l'induction en Arabinose par rapport à la répression en Glucose (p-valeur = 6,67 % pour A1 long, p-valeur = 0,286 % pour A1 court). Au contraire, aucune augmentation significative de l'activité de recombinaison n'est observée pour le mutant *A1*HD397LV. Ces résultats suggèrent que A1

participe à cette activité de recombinaison et que son activité DNase est essentielle pour cette fonction.

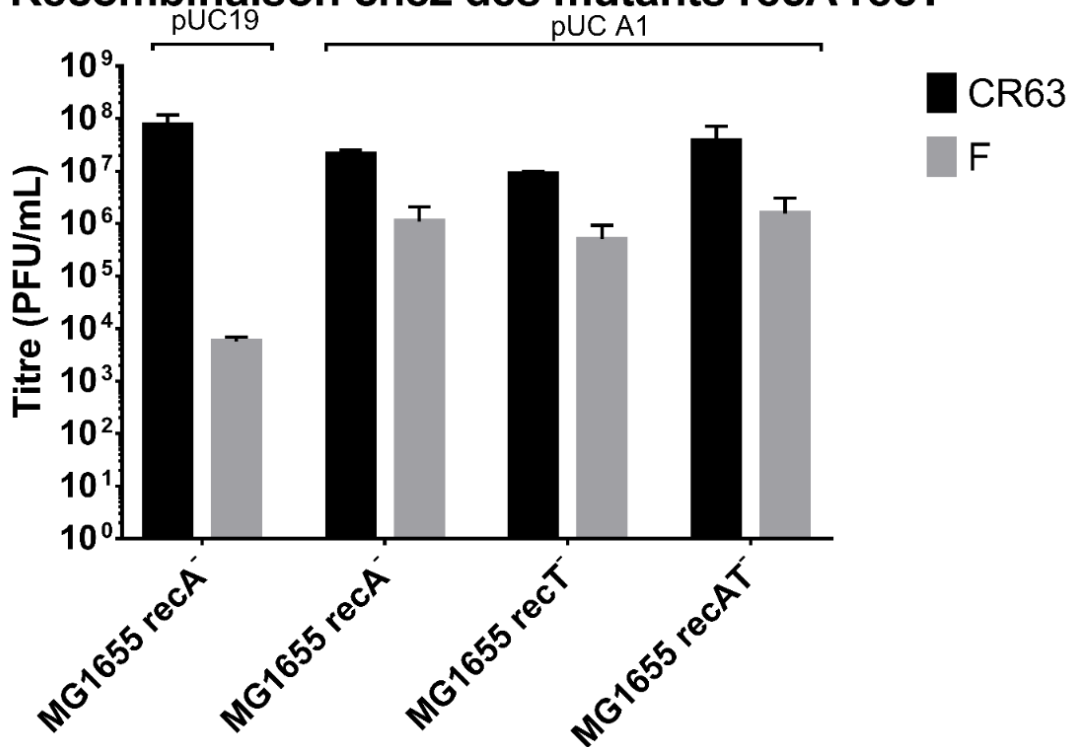
#### 1.4.4 *Recombinaison chez des mutants recA recT*

Nous avons testé si des recombinaisons bactériennes pouvaient être impliquées dans cette recombinaison. Pour tester cette hypothèse, j'ai effectué un essai de complémentation en utilisant un plasmide pUC19 à haut nombre de copies dans lequel j'ai cloné le gène A1 sauvage dans le sens inverse du sens de transcription pour empêcher toute expression de A1. Ces expériences ont été faites dans des bactéries MG1655 portant des délétions simples ou une double délétion du gène de recombinaison bactérienne *recA* et du gène de recombinaison *recT* provenant d'un prophage (**Figure 86**) alors que les expériences précédentes étaient faites avec MG1655 *recA*<sup>-</sup> uniquement.

Dans tous les cas, on observe un taux de recombinaison de base similaire à ce qui était obtenu en glucose dans l'expérience précédente. La recombinaison n'est donc pas due à RecA ou à RecT. L'activité recombinaison provient donc probablement d'une protéine du FST (8% du génome injectés en premier) du phage, car la deuxième étape du transfert (SST) ne peut pas avoir lieu sans une protéine A1 fonctionnelle.

Puisque cette activité peut être observée sans les recombinaisons majeures de la bactérie, il faut vérifier que la recombinaison a effectivement eu lieu pour confirmer qu'il n'existe pas un autre moyen d'expliquer ces résultats.

## Recombinaison chez des mutants *recA recT*



**Figure 86.** Recombinaison chez des mutants *recA recT* en présence d'un plasmide pUC19-A1inv à haut nombre de copies. n = 2.

### 1.4.5 Confirmation de l'activité recombinase par séquençage :

Pour prouver qu'une recombinaison a lieu, il faut montrer qu'une séquence est transférée depuis l'ADN du plasmide vers celui du phage pendant l'expérience de complémentation. J'ai construit un plasmide portant 2 mutations synonymes sur les acides aminés Q237 et K239 autour du codon stop Q238Stop du phage T5amH27 (**Figure 87**) et j'ai fait le test de de complémentation avec ce phage.

T5amH27			pUCA1 synonyme		
237	238	239	237	238	239
Q	Stop	K	Q	Q	K
<u>CAG</u>	<u>TAG</u>	<u>AAA</u>	<u>CAA</u>	<u>CAG</u>	<u>AAG</u>

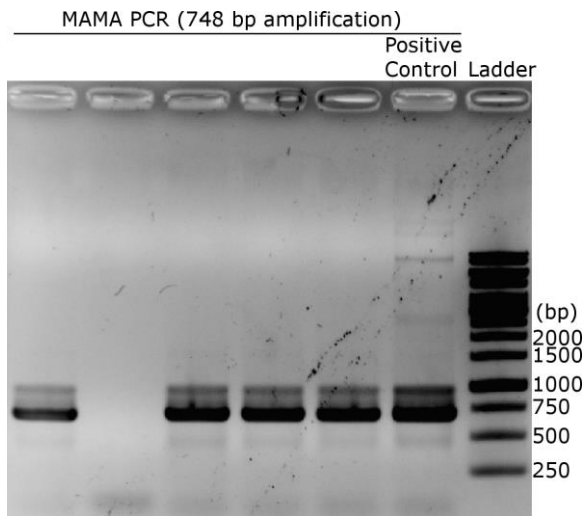
**Figure 87.** Acides aminés 237 à 239 de A1 chez T5amH27 et chez le plasmide pUC19 muté pour tester la recombinaison. Les bases mutées sont indiquées en rouge.

Pour tester si les mutations synonymes sont transférées sur les phages, j'ai utilisé deux méthodes : la MAMA PCR et le séquençage.

La MAMA PCR consiste à faire une PCR en utilisant une amorce qui peut s'hybrider avec la séquence de A1 du mutant synonyme, mais pas avec celle du phage portant le codon ambre<sup>129</sup>. Des plages de lyses produites sur la souche F ont été testées par MAMA PCR. Sur 10 plages de



lyse testées, 9 étaient positives, ce qui indiquait que le phage possédait les mutations synonymes (**Figure 88**).



**Figure 88. Analyse par MAMA PCR des phages après complémentation.** Des plages de lyse sont obtenues après une complémentation de phages mutants pour A1 par un plasmide portant des mutations synonymes au niveau du codon Stop. Le gène A1 du phage est amplifié par PCR avec des amorces complémentaires de la séquence synonyme mais non-complémentaires de la séquence originale. L'amplification est vérifiée par migration sur gel d'agarose.

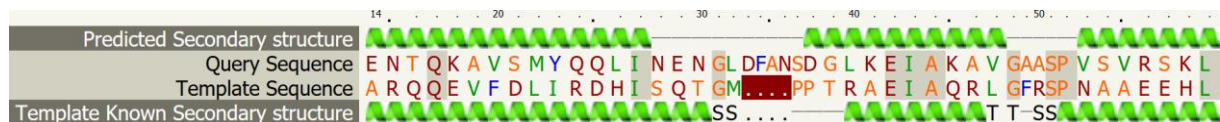
Nous avons amplifié par PCR la région contenant le codon stop de deux phages positifs en MAMA PCR. Le séquençage de ces ADN de phages capables de lyser sur F a démontré la présence des mutations synonymes et le remplacement du codon stop 238 par une glutamine, ce qui confirme que nous observons bien une recombinaison.



### 2.1.1.2 Prédiction de structure sous Phyre

Le serveur Phyre permet de faire des prédictions de structure secondaire à partir de la séquence de la protéine, puis de comparer cette structure avec celles de protéines connues. La meilleure correspondance obtenue est le domaine N-terminal de liaison à l'ADN de la protéine LexA, qui est un répresseur transcriptionnel bactérien de la superfamille des "winged helix DNA-binding domain". Il s'agit là encore d'un régulateur transcriptionnel. Le domaine N-terminal de A2 est aligné sur le motif Helix-Turn-Helix (HTH) de LexA avec un "confidence score" de 57,8 et 31 % d'identité de séquence (**Figure 90**).

Nous avons donc mis en évidence une ressemblance probable de A2 avec des régulateurs transcriptionnels, basée sur une analogie de structure avec une région Helix Turn Helix.



**Figure 90.** Alignement de la protéine A2 avec la protéine LexA obtenu avec le serveur Phyre.

### 2.1.2 Dichroïsme Circulaire

Pour obtenir des indices sur la structure secondaire de A2, nous avons fait des expériences de dichroïsme circulaire.

Les données de BestSel indiquent 41,1 % d'hélices  $\alpha$  et 7,7 % de feuillets  $\beta$  avec un très bon ajustement des données reconstruites aux données expérimentales (**Figure 91**). Le serveur Dichroweb avec le logiciel CDSSTR indique 61 % d'hélices pour 10 % de feuillets. Les prédictions de structure secondaire de A2 indiquent donc qu'elle est composée en majorité d'hélices  $\alpha$ .

# SINGLE SPECTRUM ANALYSIS RESULTS

A. Micsonai, F. Wien, L. Kernya, Y-H. Lee, Y. Goto, M. Réfrégiers and J. Kardos. PNAS 112:E3095-103 (2015)

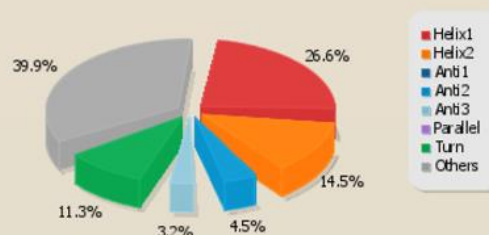
Header

Title	<b>A2</b>	Wavelength range	<b>190-250 nm</b>
Date/time	<b>15-01-2018 11:08 (GMT+1:00)</b>	Scale factor	<b>1</b>

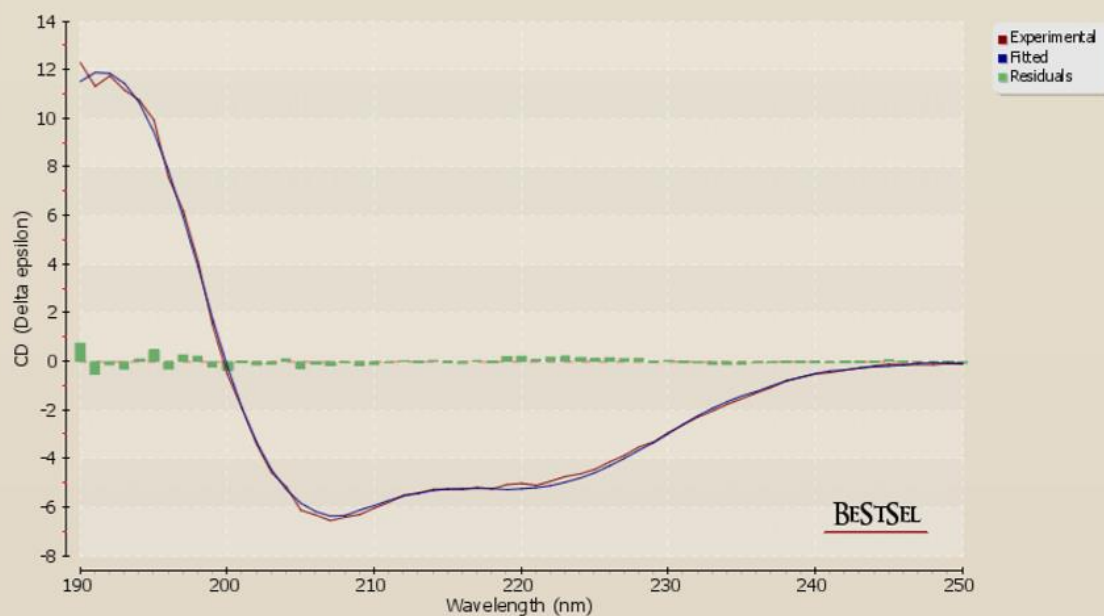
Secondary structure

Estimated secondary structure content (%)

Helix	<b>41.1</b>	Helix1 (regular)	<b>26.6</b>
		Helix2 (distorted)	<b>14.5</b>
Antiparallel	<b>7.7</b>	Anti1 (left-twisted)	<b>0.0</b>
		Anti2 (relaxed)	<b>4.5</b>
Parallel	<b>0.0</b>	Anti3 (right-twisted)	<b>3.2</b>
Turn	<b>11.3</b>		
Others	<b>39.9</b>		



Fitting results



**Figure 91.** Analyse des données de Dichroïsme circulaire obtenues pour A2 par le serveur BestSel.

### 2.1.3 Analyses de SAXS

A2 étant très soluble et n'étant pas toxique, nous avons pu en obtenir des quantités suffisantes pour faire des études de SAXS et ainsi étudier l'enveloppe de A2 en solution. Les principales caractéristiques extraites de la courbe SAXS de A2 sont résumées dans le **Tableau 7** page 139 donné dans la section Matériel et Méthodes.

La fonction de distribution des distances extraite de la courbe de SAXS est caractéristique d'un objet allongé (**Figure 92 A**). Les représentations de Kratky et de Porod (**Figure 92 B et C**) suggèrent que A2 est très flexible. En effet, un objet ayant une surface rigide bien définie a une valeur  $I(q)q^4$  en fonction de  $q$  plutôt constante alors que cette valeur est croissante pour notre objet. Les prédictions de structure secondaire (**Figure 92 D**), de désordre (**Figure 92 E**) et la prédiction de Phyre qui aligne le domaine N-terminal de A2 sur le motif HTH du répresseur LexA (**Figure 90**) permettent de déduire que la dimérisation a probablement lieu au niveau des hélices du domaine C-terminal. La protéine est prédite comme étant flexible entre les résidus 60 et 80, ce qui pourrait correspondre à la présence d'un linker entre les domaines N- et C-terminaux de A2. Cette flexibilité permet la mobilité entre les deux domaines N-ter qui contiennent un motif HTH et peuvent interagir avec l'ADN. A2 présente une structure dimérique caractéristique de nombreuses protéines se liant à l'ADN (**Figure 92 F**). Les résultats présentés correspondent à la protéine A2 marquée avec un His tag du côté C-terminal, mais des résultats similaires ont été obtenus pour la protéine A2 marquée avec un His tag et avec un site de clivage Factor Xa du côté N-terminal. Les courbes de diffusion expérimentale et ajustée sont présentées en **Figure 92 G**.



#### 2.1.4 Essais de cristallogénèse

A2 est une protéine très soluble, ce qui a empêché sa cristallisation. Nous avons essayé de cristalliser deux variantes de A2 : A2 Factor Xa His Nter et A2 His Cter. Nous avons fait des tests avec une protéine préparée dans des tampons à différents pH : Borate de Sodium 50 mM pH 9, NaCl 100 mM ; Acétate de Sodium 50 mM pH 4, NaCl 100 mM ou Tris 50 mM pH 7,2, NaCl 100 mM. Les tests ont été fait sur les kits Classics, PEGs I et PEGs II. Très peu de conditions ont permis d'observer de la précipitation, même lors d'un essai avec une protéine concentrée à 213,2 mg/mL.

Quelques aiguilles ont été obtenues pour A2 factor Xa His Nter à 4°C et nous avons essayé d'optimiser deux conditions :

- 4°C en sodium cacodylate 0,1 M pH 6,5 + tri Na citrate 0,2 M + isopropanol 25 à 35 %.
- 4°C en Hepes 0,1 M pH 7,5 + MgCl<sub>2</sub> 0,2 M + isopropanol 25 à 35 %.

Aucune de ces conditions n'a permis d'obtenir de cristal exploitable. Il est donc nécessaire d'utiliser d'autres méthodes pour résoudre la structure de A2. La masse moléculaire de A2 est de 28 kDa, ce qui est suffisamment petit pour pouvoir la résoudre par RMN.

#### 2.1.5 Étude structurale de A2 par RMN : spectre HSQC

Nous avons fait des essais préliminaires de RMN pour voir si la résolution de la structure de A2 serait possible de cette manière. J'ai produit une version de A2 marquée en <sup>15</sup>N pour faire un spectre HSQC (Heteronuclear Single Quantum Coherence) qui permet de faire apparaître les liaisons N-H. On observe donc les pics correspondant aux liaisons N-H du squelette de la protéine, soit un pic par acide aminé (à l'exception des prolines) et également des pics correspondant aux liaisons N-H des chaînes latérales des lysines, arginines, glutamines et asparagines. La position de chaque pic dépend de l'environnement atomique de chaque liaison N-H qui est modifié par l'acide aminé et ses voisins proches.

Les pics sur le spectre RMN sont globalement dispersés et assez fins, ce qui indique que la protéine A2 est majoritairement repliée et structurée. En revanche, certains pics sont plus étalés, ce qui est caractéristique de domaines flexibles. La résolution de la structure de A2 par RMN devrait être possible (**Figure 93**).

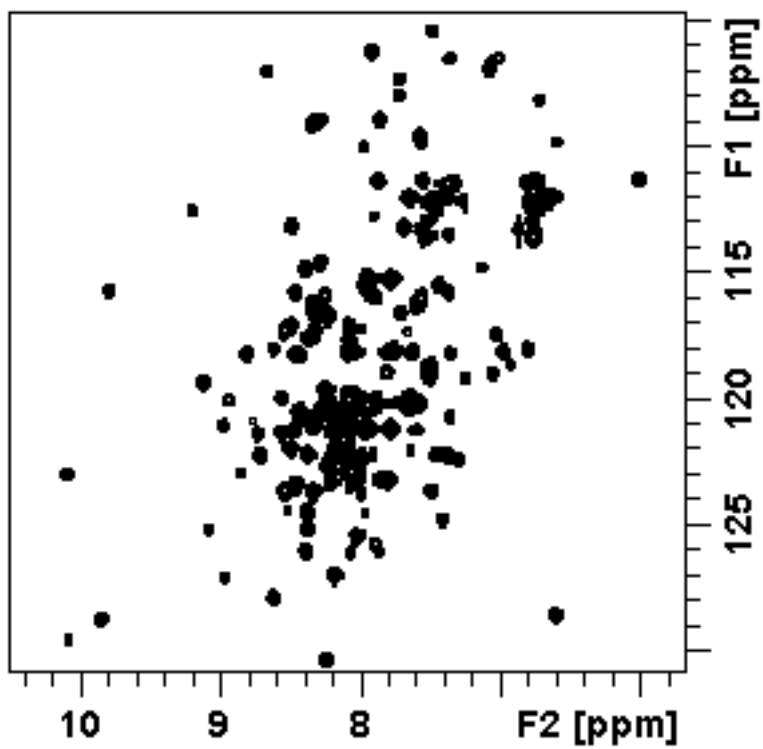


Figure 93. Spectre HSQC de la protéine A2 obtenu par RMN.





---

## III. Discussion et perspectives

---

## Discussion

La stratégie originale utilisée par le phage T5 pour prendre le contrôle des fonctions de sa bactérie hôte reste encore énigmatique. Le mécanisme de transfert de l'ADN du phage T5 en deux étapes n'est pas encore compris, particulièrement parce que les protéines A1 et A2 responsables de la reprise du transfert sont très peu connues. De plus, une des étapes précoces les plus marquantes de l'infection est la dégradation massive de l'ADN bactérien, attribuée à A1 sans que son rôle dans la dégradation de l'ADN n'ait été élucidé.

Nous avons cloné, produit, purifié et résolu la structure de la protéine toxique A1 pour laquelle très peu d'informations en dehors de la masse molaire du monomère n'avaient été obtenues depuis plus de 50 ans.

Nous avons démontré l'activité DNase de A1 alors qu'aucun article dans la littérature n'a mis en évidence une activité DNase pour l'ensemble des gènes précoces codés par le fragment FST<sup>88</sup>. Nous avons résolu la structure atomique de A1 qui montre clairement l'existence d'une activité DNase. J'ai caractérisé la spécificité de cette activité DNase *in vitro*, identifié des résidus essentiels pour cette activité et montré que cette activité est essentielle pour l'infection par le bactériophage T5. Par ailleurs, cette activité DNase est impliquée dans une activité recombinase exercée par le phage. Nous avons étudié l'équilibre d'oligomérisation de A1 entre une forme monomérique et une forme octamérique et montré que cette forme octamérique existe *in vivo*.

J'ai également purifié et caractérisé les propriétés biochimiques de A2, la deuxième protéine précoce du phage T5 qui est nécessaire à la reprise du transfert de l'ADN. Ces travaux permettront de débiter son étude fonctionnelle et structurale.

### Structure originale de la protéine A1.

Le premier objectif de ma thèse était d'obtenir des informations sur la structure de A1 à haute résolution. Plusieurs méthodes nous ont permis d'atteindre cet objectif : nous avons mené en parallèle des études de cryo-microscopie électronique et de cristallographie aux rayons X. L'analyse des données obtenues en cristallographie est encore en cours et a permis d'obtenir des données à 3,3 Å de résolution. Les meilleurs résultats ont été obtenus grâce à la cryo-microscopie électronique qui a permis la résolution atomique à 3,0 Å de résolution de la structure de la protéine A1. Les deux modèles obtenus se superposent, ce qui indique qu'ils sont fiables : la comparaison de chaque chaîne d'un tétramère du modèle obtenu en cristallographie

avec les chaînes moyennées du modèle obtenu en cryo-microscopie donne des RMSD de 1,22, 1,19, 1,15 et 1,19 Å. Les modèles se superposent donc avec 1,22 Å ou moins d'écart moyen entre les positions des carbones  $\alpha$ . Ce qui est observé et discuté sur le modèle de cryo-microscopie est donc valide par rapport au modèle de cristallographie.

Le fait que A1 forme un octamère est surprenant, cette structure originale pour une DNase semble unique : rien de semblable n'est décrit dans la littérature. A1 suit une symétrie D<sub>4</sub> : de manière remarquable, A1 utilise deux tétramères symétriques pour son activité alors que rien ne semble exclure qu'un tétramère seul puisse être actif. Des interactions sont observées entre les sites actifs de chaque monomère au niveau de la symétrie 4. Ces interactions sont inexistantes chez des DNases similaires comme Mre11 qui forme un dimère impliqué dans la recombinaison et la réparation de l'ADN ou la DNA pol II qui forme un dimère avec une activité exonucléase de type "proofreading". Ces DNases sont précises et coupent un nombre limité de nucléotides alors que la protéine A1 est processive. Les interactions inter-chaîne entre les sites actifs sont-elles responsables de la grande processivité de A1 ? Globalement, la séquence de A1 est très peu conservée par rapport aux DNases homologues. Seuls les acides aminés du site actif sont conservés. A1 a donc évolué vers une forme plus processive.

Les DNases en anneau qui enserrant l'ADN sont souvent très processives, comme par exemple l'exonucléase du bactériophage  $\lambda$  qui forme un tétramère<sup>130</sup>. A1 présente ce type de structure dans laquelle l'ADN est enfermé pour être dégradé : cette organisation pourrait être responsable de la processivité de A1. Les charges majoritairement négatives localisées à la surface du tétramère et les charges positives concentrées à l'entrée du canal sont très nettement séparées chez A1, ce qui est inhabituel : dans les structures de protéines, les charges électrostatiques sont généralement plus régulièrement réparties sur l'ensemble de la molécule. L'entrée du canal chargée positivement pourrait porter une activité "unwindase" qui, par compétition avec les liaisons électrostatiques qui stabilisent l'ADN double brin, agirait en séparant les brins d'ADN sans activité ATPase. En effet, tous les tests d'activité *in vitro* ont été faits sans addition d'ATP et aucun site de fixation à l'ATP n'a été révélé par la structure de A1. Cette activité de dissociation des brins d'ADN pourrait être facilitée en présence de substrats présentant une région d'ADN simple-brin et expliquer la dégradation complète des substrats d'hélicase que nous avons observée.

La structure du monomère de A1 montre un domaine DNase actif fixant deux ions Mn<sup>2+</sup> et un domaine DNase inactif du côté N-terminal de la protéine. Ce deuxième domaine résulte

probablement d'une duplication du domaine actif qui aurait perdu les acides aminés participant à l'activité DNase au cours de l'évolution. De nombreux acides aminés de ce domaine participent à l'oligomérisation au niveau de l'axe 4 de symétrie, ce domaine apparaît donc essentiel pour la formation de l'octamère.

Nous avons obtenu une résolution de 3,0 Å en cryo-microscopie électronique, ce qui est très bon. Cependant, cette résolution est un peu faible pour résoudre tous les détails de la structure : il n'y a pas de doute possible sur la chaîne principale, mais de nombreux rotamères ne sont pas fiables. Nous ne pouvons pas visualiser directement les sites métalliques et il a été nécessaire de déduire l'existence de deux ions  $Mn^{2+}$  dans le site actif. La couche de solvation qui correspond aux molécules d'eau les plus proches de la protéine n'est pas non plus visible à cette résolution. Par ailleurs, le domaine 1-89 et la boucle 138-148 sont flexibles et n'ont pas pu être résolus, que ce soit par microscopie ou par cristallographie.

Les travaux précédents dans la littérature ne donnaient aucun indice sur la structure de la protéine A1. Notre travail a permis de comparer la structure de A1 avec la structure d'autres protéines dont la fonction est connue, ce qui nous a convaincus de son activité DNase que nous avons confirmé par des tests biochimiques et ce qui nous a donné des indices très forts sur son rôle dans l'infection.

### Spécificité de l'activité DNase de la protéine A1

Nos essais préliminaires montraient que la protéine A1 dégradait l'ADN *in vitro* en présence de  $MnCl_2$ . Notre objectif était alors de caractériser la spécificité de cette activité DNase.

Nous avons exploré de nombreux aspects de l'activité DNase de A1. Nous avons comparé son activité sur plusieurs types de substrats dans des conditions variées et caractérisé des protéines A1 mutantes inactives.

A1 est plus efficace pour dégrader l'ADN simple-brin que l'ADN double-brin, ce qui est à priori surprenant, étant donné son rôle supposé dans la dégradation massive du chromosome de l'hôte. Nous avons montré qu'elle est active dans les deux sens  $3' \rightarrow 5'$  et  $5' \rightarrow 3'$ , comme l'Exonucléase VII<sup>125</sup>. Cette activité bidirectionnelle est-elle importante pour la processivité de A1 ? Dans le cas où les tétramères agiraient de manière indépendante, A1 pourrait dégrader un même brin d'ADN des deux côtés en même temps. Des résultats très récents obtenus par Jean Lepault en cryo-microscopie électronique ont corroboré cette hypothèse : ils ont mis en évidence que l'octamère de A1 peut fixer deux brins d'ADN différents à chacune de ses

extrémités lorsque l'ADN est en excès. Par ailleurs, il n'est pas nécessaire que l'extrémité 5' possède un groupement phosphate pour être reconnue et dégradée par A1.

Bien que A1 dégrade plus efficacement l'ADN simple-brin, elle peut dégrader *in vitro* de l'ADN chromosomique extrait des bactéries et, avec une moindre efficacité, de l'ADN plasmidique. De plus, des expériences faites au laboratoire par Ombeline Rossier ont montré que l'expression ectopique de A1 induisait la dégradation du chromosome bactérien *in vivo*. Ces résultats tendent à confirmer le rôle direct de A1 dans la dégradation du chromosome. Une question importante liée à l'activité DNase de A1 est d'élucider le mécanisme qui protège l'ADN du phage T5, qui ne porte pas de modifications chimiques, contre cette activité. Nous avons montré que A1 est active sur un substrat d'hélicase : pourrait-elle exclusivement envahir les fourches de réplication sur le chromosome de l'hôte pendant l'infection, ce qui protégerait les autres formes d'ADN de la dégradation ? L'ADN du FST du phage est le seul fragment d'ADN viral exposé à A1 dans la première étape de l'injection. Il n'est pas répliqué et il serait alors protégé de l'activité DNase de A1. Cependant, dès que la deuxième étape du transfert a lieu, l'ADN polymérase du phage est rapidement synthétisée pour initier la réplication de l'ADN viral et celui-ci ne serait plus protégé à moins que A1 soit rapidement inactivée ou dégradée. Par ailleurs, nous avons observé *in vitro* la dégradation de différents plasmides purifiés qui sont principalement dans un état surenroulé au début du test. Une activité plus spécifique de A1 pour des fourches de réplication ne suffirait donc pas à expliquer tous les résultats observés, ce qui affaiblit cette hypothèse. Un autre mécanisme de protection possible pourrait venir d'une "titration" de A1 par l'ADN chromosomique : lorsqu'elle est produite, A1 pourrait s'associer rapidement à la masse importante d'ADN bactérien (environ  $4,5 \times 10^6$  pb) et la quantité de A1 disponible pour dégrader un exemplaire du fragment FST dont la taille est de 10 kpb serait alors très réduite. La dissociation des octamères de A1 pourrait être suffisamment lente pour préserver l'ADN FST et ensuite l'ADN SST. Ce schéma n'exclut pas une inactivation ou dégradation de A1 qui aurait lieu rapidement après le deuxième transfert. Des bactéries sans chromosome, les minicells<sup>131</sup>, pourraient être utilisées pour tester cette hypothèse : l'ADN de phage serait alors majoritaire et devrait être dégradé. Par ailleurs, l'ADN de l'hôte étant déjà dégradé, A1 pourrait ne pas être nécessaire et un phage mutant pour A1 pourrait infecter des minicells. Pour tester le rôle des autres gènes précoces dans cette protection de l'ADN de phage, des phages mutants ambre pourraient être construits afin de vérifier si l'ADN de phage est dégradé en l'absence de certains de ces gènes.

Nos expériences de complémentation ont permis de démontrer que l'activité DNase de A1 est essentielle pour l'infection par le bactériophage T5. L'activité DNase est-elle essentielle parce qu'elle permet le transfert en deux étapes ou est-elle liée à un autre aspect crucial de l'infection ? Certains phages comme SPP1 n'ont pas besoin de dégrader l'ADN de l'hôte <sup>132</sup>, ou d'autres le dégradent lentement et recyclent les nucléotides produits pour la synthèse de leur génome <sup>90,91</sup>. Pourquoi le phage T5 a-t-il conservé au cours de l'évolution un mécanisme d'infection qui dégrade totalement le génome bactérien, ne réincorpore pas les nucléotides issus de cette dégradation <sup>98</sup> lors de la réplication de son ADN et assure leur excrétion rapide dans le milieu extérieur ? Cette dégradation pourrait être une étape clé de l'arrêt de la transcription des gènes de défense de l'hôte. La compréhension de cette stratégie nécessite certainement d'explorer non seulement la fonction de l'ensemble des protéines précoces de T5 mais également d'autres étapes du processus d'infection comme le mécanisme de réplication de l'ADN.

L'observation de l'activité DNase de A1 est une avancée importante par rapport aux précédents résultats de la littérature. En effet, aucune activité DNase n'avait été prédite pour les gènes précoces de T5 <sup>88</sup>. L'interaction spécifique de A1 avec l'ARN polymérase bactérienne avait été proposée sur la base d'une expérience de co-immunoprécipitation par des anticorps spécifiques dirigés contre l'ARN polymérase de *E. coli* <sup>101</sup>. D'autres travaux avaient montré que le phage T5 sauvage inhibe rapidement la production des ARN messagers des gènes précoces du fragment FST et celle des ARN messagers bactériens <sup>133,134</sup>, ce qui n'était pas observé lors de l'infection par les mutants de A1. Sur la base de ces résultats, leurs auteurs ont postulé que A1 était probablement un régulateur transcriptionnel précoce de T5, capable de rediriger l'ARN polymérase bactérienne vers les promoteurs intermédiaires et tardifs de T5 (ceux du fragment SST). Cette activité de A1 peut sembler peu compatible avec son activité DNase. De plus, ces travaux de 1968 et 1970 sont anciens et ces résultats n'ont pas été reproduits avec des méthodes récentes. L'interaction directe entre A1 et l'ARN polymérase qui n'a pas été démontrée jusqu'ici devra être explorée par des méthodes biochimiques et génétiques (purification TAP-Tag, double hybride) qui n'existaient pas lorsque ces données ont été publiées. Si cette interaction est confirmée, on ne peut pas exclure qu'elle inhibe la fonction DNase de A1 et modifie la spécificité de l'ARN polymérase. Il faut noter que l'arrêt rapide de la synthèse des gènes bactériens associé à la production de A1 peut être la conséquence directe de la dégradation du génome de l'hôte, mais le FST doit rester intact pour que le génome entier puisse être répliqué : A1 n'a probablement pas de rôle dans l'arrêt de l'expression des gènes du fragment FST.

## Assemblage et dissociation de l'octamère

Pour mieux comprendre le rôle de l'octamère de A1 lors de l'infection, nous voulions étudier l'équilibre de formation de l'oligomère. Nous avons identifié plusieurs conditions permettant l'oligomérisation de A1 *in vitro* grâce à la microscopie électronique et nous avons étudié la cinétique de formation de l'oligomère. Nous avons confirmé l'existence de l'octamère *in vivo*. L'oligomérisation de A1 a également pu être observée par des méthodes de SEC-MALLS et de SAXS.

Trois conditions au moins permettent la formation de l'octamère : l'ajout de  $MnCl_2$  provoque l'oligomérisation lente sur plusieurs heures, l'ajout d'ADN permet une oligomérisation rapide en quelques minutes et la mutation de certains résidus du site actif provoque l'oligomérisation de la protéine. Ce résultat est à priori surprenant car ces conditions sont très différentes. Chaque ion  $Mn^{2+}$  apporte 2 charges positives dans le site actif, l'ADN n'apporte que des charges négatives et les mutations D304S ou HD397LV retirent respectivement 1 ou 2 charges négatives.

Il est toutefois possible d'expliquer ces résultats : les interactions entre les résidus du site actif sont très défavorables en absence de  $Mn^{2+}$  car ces résidus sont sur des boucles différentes et portent la même charge électrostatique. L'ajout de  $Mn^{2+}$  permet l'interaction entre tous ces résidus, ce qui peut apporter un changement de conformation suffisant pour déclencher l'oligomérisation. Le retrait d'une ou de plusieurs charges négatives dans le site actif chez les formes mutées de A1 va dans le même sens. L'ajout d'ADN permet probablement l'oligomérisation de manière totalement différente, avec des interactions au niveau du canal formé par l'octamère.

A1 peut former des oligomères sur des ADN très courts de 8 bases, ce qui montre que l'ADN n'a pas besoin d'interagir sur toute la longueur du canal pour permettre l'oligomérisation de la protéine.

Nous avons observé la formation de structures cristallines en microscopie, uniquement en présence d'oligonucléotides simple-brin, dans les conditions dans lesquelles on observe la dégradation. De manière générale, nous observons une précipitation rapide de la protéine A1 en présence d'ADN simple ou double-brin. Cette précipitation a-t-elle un rôle *in vivo*, par exemple dans la reprise du transfert de l'ADN ? L'arrêt du transfert de l'ADN de T5 au début de l'infection ne requiert la synthèse d'aucune protéine de l'hôte mais nécessite un facteur bactérien inconnu <sup>73,80,81</sup>. A1, en formant des complexes de très haut poids moléculaire avec



l'ADN de T5, pourrait déstabiliser son interaction avec ce facteur bactérien impliqué dans le blocage de l'ADN et permettre la reprise du transfert. Il est assez improbable que A1 tire mécaniquement l'ADN car ce type d'activité mécanique requiert de l'énergie, or nous n'avons identifié aucun site ATPase chez A1, même dégénéré.

Les expériences de microscopie électronique à coloration négative présentent des limites pour analyser en détail ces complexes ADN-protéines. Les essais sont effectués sur une grille de microscopie avec des échantillons fixés, ce qui peut générer des artefacts. La fixation de la protéine sur la grille dépend de ses paramètres biochimiques et les gros agrégats ne se fixent pas sur la grille. Il est possible que le monomère se fixe mieux sur la grille que l'octamère ou inversement, ce qui modifie artificiellement l'équilibre observé. Si l'observation en coloration négative est plus rapide qu'en cryo-microscopie, elle est aussi moins précise. Dans l'idéal, toutes les observations faites en coloration négative devraient être confirmées en cryo-microscopie électronique.

Les travaux antérieurs qui visaient à caractériser l'état d'oligomérisation *in vivo* ont conclu à la formation d'un oligomère de 244 kDa <sup>134</sup>, valeur qui s'approche de la masse d'un tétramère que nous n'avons jamais observé. Notre étude est la première à avoir caractérisé la forme octamérique existant *in vivo* ainsi que les conditions permettant l'assemblage et la dissociation de l'octamère *in vitro*.

### A1, une possible DNase-Recombinase ?

Nous voulions savoir si l'activité DNase de A1 était essentielle pour l'infection par le bactériophage T5. Nous avons réussi à identifier un mutant de A1 sans activité DNase dont la structure n'est pas modifiée par rapport à celle de la protéine sauvage et qui ne complémente pas. Nous confirmons donc que l'activité DNase de A1 est indispensable.

Les expériences de complémentation ont permis de montrer que l'activité DNase de A1 est liée à une activité recombinase élevée. A1 possède un domaine DNase très similaire à celui de Mre11, une enzyme impliquée dans la recombinaison homologue avec son partenaire Rad50 et la réparation de l'ADN <sup>108</sup>. Ce domaine DNase ressemble également à celui de la nucléase SbcD qui est responsable de la recombinaison homologue avec son partenaire SbcC <sup>135,136</sup>. On ne sait pas actuellement si des processus de réparation de l'ADN et de recombinaison sont impliqués pendant les étapes précoces de l'infection. Cependant, les résultats que nous avons obtenus sont

une indication que A1 pourrait avoir un rôle similaire à celui de ses homologues dans des processus de recombinaison et de réparation de l'ADN.

Pour approfondir cette question, il sera nécessaire de répéter les essais de complémentation et d'obtenir de meilleures statistiques pour l'analyse des données. Il faudra également tester un nombre plus important de mutants de A1 inactivés pour leur fonction DNase et vérifier leur capacité de complémentation et recombinaison. Nos essais montrent que la protéine A1 a une activité de recombinaison sur son propre gène, ce qui pourrait introduire un biais sur l'estimation de son taux de recombinaison. Les mêmes expériences de complémentation/recombinaison seront effectuées avec un phage muté sur la protéine A2 et des plasmides contenant le gène A2.

Des essais de complémentations avaient déjà été faits par Lanni en 1968<sup>73</sup>. Ces tests étaient très différents des nôtres : des bactériophages mutants étaient utilisés pour infecter différentes souches de *E. coli* et la complémentation consistait en une surinfection avec un autre phage mutant ou avec un phage sauvage. Ces expériences ont permis de déterminer quel gène était muté chez les phages testés alors que le séquençage n'était pas possible. Nos tests de complémentation sont les premiers à montrer une activité recombinase liée aux gènes du FST de T5.

On trouve des protéines similaires à A1 dans de nombreux autres phages qui ne font pas partie de la famille *T5virus* : l'étude approfondie de cette protéine pourra donner des indices son rôle dans des systèmes différents.

### Caractérisation biochimique de la protéine A2 vers son étude structurale

Notre objectif était d'obtenir des informations sur la structure et la fonction de A2 qui pourraient nous éclairer sur son rôle pendant l'infection. Nous avons confirmé l'organisation dimérique de A2 mise en évidence par Snyder en 1981<sup>137</sup> qui avait extrait la protéine à partir de bactéries infectées. Les analyses bioinformatiques prédisent que A2 contient un domaine Helix-Turn-Helix du côté N-terminal que l'on trouve très fréquemment dans les protéines qui se fixent à l'ADN et indiquent une forte ressemblance avec plusieurs régulateurs transcriptionnels. Cette possible fonction de A2 n'avait encore jamais été proposée dans la littérature.

A2 est une protéine difficilement cristallisable. Nous n'avons donc pas pu faire son étude structurale par cristallographie, mais sa production ayant été optimisée en milieu riche et

minimum, nous pouvons débiter son étude fonctionnelle ainsi que son étude structurale par RMN.

Des résultats très récents obtenus en RMN en collaboration avec l'équipe de Sophie Zinn-Justin ont confirmé la présence de trois hélices  $\alpha$  dans le domaine N-terminal (acides aminés 3-80), ce qui valide les prédictions bioinformatiques. Des expériences préliminaires de QRT-PCR faites en collaboration avec l'équipe de Mireille Ansaldi sur des bactéries infectées par T5 sauvage ou par un mutant de A2 ont montré une baisse significative de l'expression des gènes du FST de T5 en présence de la protéine A2. Ces résultats tendent à confirmer les prédictions qui indiquent un rôle possible de A2 dans la régulation transcriptionnelle.

## Perspectives

### Poursuite de l'étude structurale de la protéine A1

Pour compléter l'étude structurale et fonctionnelle de A1, nous devons déterminer précisément quels acides aminés sont impliqués dans l'interaction avec l'ADN et donc résoudre la structure de A1 en présence d'ADN. L'acquisition d'images de la protéine A1 en présence d'ADN par cryo-microscopie vient d'être faite par Jean Lepault et la reconstruction de la structure de ce complexe a débuté. Pour pouvoir étudier le rôle du domaine N-terminal 1-89 manquant dans la structure, des vecteurs d'expression permettant de produire ce domaine isolé ainsi que la protéine tronquée A1(90-557) viennent d'être construits. Ces deux domaines séparés vont être purifiés et leur structure et fonction étudiées. Il sera alors possible de comparer l'organisation oligomérique et l'activité DNase de la forme tronquée avec celles de la protéine sauvage. La structure de cette forme sera aussi explorée en espérant que des cristaux qui permettent de résoudre la structure à plus haute résolution que 3,0 Å puissent être obtenus. Pour résoudre la structure du domaine N-terminal manquant, des essais seront menés en parallèle par cristallographie et RMN. Le SAXS était un travail initiateur qui a permis d'observer la transition monomère/octamère en présence de  $MnCl_2$ . Ces expériences n'ont pas été poursuivies à la suite des résultats obtenus par cristallographie aux rayons X et de microscopie électronique. De nouvelles expériences de SAXS sur le monomère seront cependant très utiles pour analyser les propriétés du domaine N-terminal.

### Exploration de l'activité DNase de A1

L'activité très efficace de A1 sur l'ADN bactérien doit être confirmée dans des conditions contrôlées. Ces tests d'activité DNase ont été effectués sur des échantillons d'ADN chromosomique non traités à la RNase. Par ailleurs, même si nous avons utilisé des kits d'extraction commerciaux normalement optimisés, nous ne pouvons pas totalement exclure la présence de protéines bactériennes pouvant altérer l'activité de A1. Il est donc nécessaire de tester l'activité de la protéine A1 sur un ADN chromosomique extrait en phénol/chloroforme et traité avec la protéinase K et une RNase pour écarter cette possibilité. De manière générale, les tests d'activité DNase doivent être répétés sur un nombre plus important de lots différents de A1 (réplicats biologiques) pour confirmer les résultats obtenus et déterminer une activité spécifique pour A1. Pour pouvoir comparer ces activités dans les meilleures conditions, il faut

développer une méthode standard de conservation de la protéine A1 qui devra être mise en place immédiatement après sa purification. Un protocole de stockage rapide à  $-80^{\circ}\text{C}$  et en présence de glycérol sera établi et appliqué immédiatement après chaque purification pour obtenir une activité spécifique sur chaque lot. L'activité doit être testée sur des ADN linéaires de même longueur que les plasmides testés pour vérifier si l'état circulaire ou linéaire de l'ADN a un effet sur l'activité de A1. Il faut également tester l'activité de A1 sur plusieurs types d'ARN, car l'absence d'activité sur des ARN ribosomiques ne signifie pas que d'autres types d'ARN ne peuvent pas être ciblés. Il serait intéressant d'identifier d'autres résidus essentiels pour l'activité en produisant des protéines mutées dont le repliement devra être contrôlé par CD et microscopie. L'identification d'un mutant inactif pouvant fixer les ions  $\text{Mn}^{2+}$  et l'ADN pourrait permettre de contourner les problèmes de toxicité pour des études de structure plus poussées et pourrait permettre de caractériser la protéine A1 en présence d'ADN et de  $\text{MnCl}_2$ , en absence d'activité DNase.

Pour confirmer que le manganèse est le cation divalent le plus adapté pour caractériser l'activité de la protéine A1, une étude systématique de formation des oligomères couplée à un test d'activité DNase en utilisant les autres cations susceptibles d'activer A1 ( $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Zn}^{2+}$  ...) pourrait être faite. Pour ces expériences, il faudrait optimiser nos méthodes de visualisation de l'ADN pour pouvoir mieux observer l'oligomère en présence d'ADN

Enfin, une question qui reste intrigante est le rôle de A1 dans le déblocage du transfert de l'ADN après l'expression des gènes précoces. L'exploration des conditions permettant la formation des structures cristallines par microscopie électronique pourrait donner des informations supplémentaires sur la façon dont A1 permet la reprise de transfert de l'ADN du phage pendant l'infection. Jusqu'ici, cette cristallisation a été vue au cours de la digestion d'ADN simple brin. Est-elle observée avec de l'ADN double brin linéaire et dans ce cas est-elle associée à la digestion de l'ADN ? D'autre part, est-elle observée en présence de A2 qui est la deuxième protéine requise pour la reprise du transfert de l'ADN ? La réponse à ces questions pourrait fournir des indices sur le mécanisme d'injection en deux étapes.

### Complémentation fonctionnelle et recombinaison

L'éventuelle fonction recombinaise de A1 doit être explorée. Des tests de complémentation faits avec d'autres mutants de T5 affectés dans l'activité DNase de A1 sont nécessaires pour confirmer qu'il n'existe aucun mutant de l'activité DNase capable de compléter un phage mutant pour A1. Ces essais pourraient révéler un mutant présentant des caractéristiques

intermédiaires, c'est-à-dire une activité DNase plus lente associée à une complémentation faible, mais supérieure au contrôle négatif : ce serait une preuve que l'activité DNase et l'efficacité de complémentation sont liées. Nous avons démontré que l'activité DNase de A1 est essentielle pour l'infection par le bactériophage T5 : serait-il possible de compléter un phage mutant pour l'activité DNase de A1 avec un plasmide exprimant une autre DNase ?

Pour confirmer la forte activité recombinase observée *in vivo*, il faudra reproduire des tests de complémentation/recombinaison en utilisant des phages T5 mutés sur d'autres gènes que A1 et le plasmide produisant la protéine sauvage correspondante. Ces tests seront faits selon le protocole utilisant la restauration du phénotype sauvage avec des mutations synonymes. La complémentation de mutants T5 $amA2^-$  produisant une protéine tronquée avec un plasmide exprimant A2 sera testée, ce qui permettra de vérifier l'implication de cette protéine dans l'activité recombinase. Si A2 participe à l'activité recombinase, alors une différence élevée dans les titres obtenus sur des souches permissive et non-permissive pour les mutations ambre devrait être observée. Si A2 ne participe pas à cette activité, une forte activité recombinase devrait être observée dans tous les cas car chez le mutant T5 $amA2^-$  la protéine A1 est normalement produite.

### Caractérisation fonctionnelle et structurale de A2

Nous avons vu que la RMN est probablement la méthode la mieux adaptée pour déterminer la structure atomique de A2 entière. A2 est organisée en deux domaines séparés par une courte région désordonnée. La production des deux domaines indépendants sera tentée. Si elle est possible, de nouveaux essais de cristallisation seront faits sur ces domaines.

Des mesures d'affinité de A2 pour l'ADN ont été effectuées en collaboration avec l'équipe de Sophie Zinn-Justin en utilisant la technologie switchSENSE commercialisée par Dynamic Biosensors. Elles ont permis de déterminer une constante de dissociation  $K_D = 12,8 \pm 1,6$  nM pour l'interaction de A2 avec un ADN de 48 pb, ce qui montre la forte affinité de A2 pour l'ADN. Cette propriété va permettre de former des complexes stables A2/ADN. Leur cristallisation pourra être tentée avec plus de chances de succès que celle de A2 seule et permettre de déterminer la structure de A2 en présence d'ADN.

## Conclusion

L'ensemble des travaux sur la structure de A1 et son activité DNase font l'objet d'un article en préparation, pour lequel des expériences complémentaires sur l'activité DNase sont en cours.

Mes travaux sur l'activité recombinaison vont être poursuivis au laboratoire et devront faire l'objet d'une publication ultérieure lorsque cette activité sera caractérisée de manière plus approfondie.

---

## IV. Matériel et Méthodes

---



## 1) Bactéries et phages utilisés

### 1.1 Souches bactériennes d'*Escherichia coli* utilisées

- **XL1-Blue** (Agilent Technologies) : bactéries super-compétentes utilisées pour les clonages
- **BL21(DE3)** : Souche *E. Coli* BL21 contenant une copie chromosomique du gène de l'ARN polymérase du phage T7 sous contrôle du promoteur LacUV5 et une copie du répresseur lacI. Cette souche est optimisée pour la surexpression plasmidique de gènes placés sous le contrôle du promoteur T7
- **BLIS** : Souche BL21(DE3) contenant le plasmide pLysS permettant l'expression constitutive du lysozyme de T7, inhibiteur de l'ARN polymérase de T7, qui réduit l'expression basale des gènes recombinants sous le contrôle du promoteur T7 en inhibant l'expression basale de l'ARN polymérase de T7.
- **DH5 $\alpha$**  : bactéries optimisées pour la production de plasmides
- **CR63** : souche dérivée de *E. Coli* K12, suppressive pour les codons ambre (supD60 : attribution d'une sérine face à un codon ambre)
- **F** : souche possédant un LPS dont la région O-antigène est spécifiquement reconnue par les fibres longues de T5, ce qui permet une adsorption très efficace du phage T5
- **MG1655** : souche *E. Coli* K12 sauvage de référence, dont le génome a été entièrement séquencé <sup>138</sup>, utilisée pour les essais de complémentation

### 1.2 Phages utilisés

- **T5 WT** : phage T5 sauvage
- **T5 $amH27$**  : phage T5 mutant ambre dans le gène A1 : Q238Stop.
- **M13** : phage filamenteux possédant un ADN simple-brin, extrait et utilisé pour des tests d'activité de A1

### 1.3 Mutations ambre et souches suppressives

Le phage T5 $amH27$  est un mutant ambre qui produit une protéine A1 tronquée qui ne permet pas une infection productive dans un hôte classique. Le codon stop ambre (TAG) remplace le codon de la Glutamine 238 dans le gène A1.

CR63 est une souche suppressive pour les codons ambre. Elle porte le gène *SupD* codant pour un ARN de transfert (ARNt) muté dans la région anticodon, capable d'insérer une Sérine au

niveau d'un codon ambre situé dans une séquence codante, ce qui permet l'expression du gène complet. Le phage T5amH27 peut être produit sur une bactérie CR63.

## 2) Clonages et mutagenèses

Au début de ma thèse, la version longue du gène A1 était déjà clonée dans le plasmide pBAD24 et la version courte dans le plasmide pCDF. Le gène A2 était déjà cloné dans les plasmides pBAD24, pET28b et pASK.

### 2.1 Plasmides utilisés

- **pBAD24** : contrôle de l'expression du gène cloné par le glucose (répression) et l'arabinose (induction). Contient une cassette de résistance à l'ampicilline.
- **pCDF2-EK\_LIC** : Le gène est systématiquement cloné avec un His tag (6 Histidines) et un site de coupure entérokinase du côté N-terminal. L'expression du gène cloné est sous contrôle du promoteur T7 inductible par l'IPTG. Contient une cassette de résistance à la streptomycine.
- **pET28b** : Un site His-tag avec ou sans site de coupure TEV peut être intégré en C-ter ou en N-ter. L'expression du gène cloné est sous contrôle du promoteur T7 inductible par l'IPTG. Contient une cassette de résistance à la kanamycine.
- **pUC19** : Ce plasmide a été utilisé comme substrat pour les tests d'activité DNase de A1 et pour des essais de recombinaison. Contient une cassette de résistance à l'ampicilline.
- **pASK** : L'induction de l'expression du gène cloné est déclenchée par ajout de Tétracycline. Contient une cassette de résistance à l'ampicilline.

### 2.2 Clonages conventionnels avec enzymes de restriction

Le gène à cloner a été amplifié par PCR avec des amorces comprenant des sites de restriction. L'insert obtenu et le vecteur de destination ont été digérés avec les mêmes enzymes de restriction (EcoRI et XbaI). L'insert obtenu a été introduit dans le plasmide par une étape de ligation. Une étape supplémentaire de digestion a été faite par une enzyme de restriction (BamHI) dont le site est présent sur le plasmide original et absent sur le plasmide reconstruit pour éliminer le plasmide de départ. Le plasmide reconstruit a été transformé dans la souche *E. coli* DH5 $\alpha$  qui a été cultivée pendant la nuit (overnight O/N) dans un milieu contenant l'antibiotique adéquat. Les plasmides ont été extraits avec le kit "Thermo Fisher GeneJet Plasmid Miniprep kit" et le succès du clonage a été vérifié par PCR sur colonie, puis par séquençage.

Cette technique a permis d'obtenir toutes les constructions non mutées utilisées dans ce manuscrit de thèse.

## 2.3 Mutagenèse dirigée

Le gène A1 a été muté par amplification du plasmide par PCR avec deux amorces complémentaires portant la mutation désirée. Les amplifications avec chaque amorce sont faites séparément pour éviter une hybridation entre les amorces, puis les produits de réaction sont mélangés et hybridés. Les produits hybridés sont traités avec l'enzyme DpnI qui dégrade spécifiquement le plasmide matrice méthylé sans affecter le plasmide amplifié par PCR qui est non méthylé. Le produit de réaction est transformé dans des bactéries XL1-Blue qui sont cultivées O/N dans un milieu LB + antibiotique. Les plasmides sont extraits avec le kit "Thermo Fisher GeneJet Plasmid Miniprep kit" et la mutation est vérifiée par séquençage.

## 3) Production et purification des protéines A1 et A2

### 3.1 Production

Les protéines sont produites à partir de bactéries *E. coli* BL21(DE3) transformées avec le plasmide exprimant la protéine d'intérêt :

- **pCDF ou pET28b pour A1.** Culture O/N dans un milieu LB + Streptomycine 100 mg/mL (pCDF) ou Kanamycine 50 mg/mL (pET28b) + glucose 0,5 % à 37°C. La préculture obtenue est utilisée pour inoculer 3 Litres de culture dans un milieu LB + Antibiotique + Glucose 0,5 %. Lorsque l'absorbance à 280 nm atteint environ 0,9, 0,4 mM d'IPTG sont ajoutés pour induire l'expression de la protéine A1 pendant 90 minutes.
- **pET28b ou pASK pour A2.** Culture O/N dans un milieu LB + Kanamycine 50 mg/mL (pET28) ou Ampicilline 100 mg/mL (pASK) à 37°C. La préculture obtenue est utilisée pour inoculer 1,5 Litres de culture dans un milieu LB + Antibiotique. Lorsque l'absorbance à 280 nm atteint environ 0,4, 0,4 mM d'IPTG (pET28b) ou 0,2 µg/µL de Tétracycline anhydre (pASK) sont ajoutés pour induire l'expression de la protéine A2 pendant 3 heures.

Les cellules sont centrifugées à 11000 g pendant 10 min à 4°C puis congelées dans de l'azote liquide et conservées à -20°C.

## 3.2 Purification

### 3.2.1 Extraction des protéines

Les cellules sont resuspendues dans 20 mL de tampon A d'équilibration d'une colonne His Trap FF (GE Healthcare) (voir **Tableau 6**), avec un cocktail d'inhibiteurs de protéases (cOmplete ULTRA Tablets, Sigma-Aldrich) et 1 mM d'EDTA.

Les cellules sont cassées à la presse de French à une pression de 1100 PSI : environ 760 bars. Une ultracentrifugation à 60000 g pendant 45 min à 4°C permet d'éliminer les débris cellulaires et les vésicules lipidiques. Le surnageant correspondant à la fraction cytoplasmique contient les protéines à purifier.

### 3.2.2 Purification par chromatographie d'affinité sur colonne de métal immobilisé (His trap)

La première étape de purification est faite sur une colonne His trap FF (GE Healthcare). Il s'agit d'une colonne chargée en ions  $Ni^{2+}$  qui interagissent avec les noyaux imidazole des histidines des protéines. La protéine à purifier est fusionnée avec une étiquette de 6 histidines, elle est donc retenue par la colonne. Du  $MgCl_2$  est ajouté à la fraction à purifier à une concentration finale de 1 mM pour chélater l'EDTA présent dans le tampon de "cassage" et empêcher l'élution des ions Nickel de la colonne. Un gradient linéaire de 0 à 100 % de tampon B contenant 1 M d'imidazole permet l'élution de la protéine par compétition.

La purification est faite avec un système de FPLC (Fast Protein Liquid Chromatography) et l'élution des protéines est suivie par mesure des absorbances à 260 et 280 nm en sortie de colonne pour évaluer la concentration des protéines purifiées.

Pour la protéine A1, une colonne His trap FF de 1 mL est utilisée. La protéine est injectée à 1 mL/min, puis la colonne est lavée avec le tampon A pendant 15 volumes de colonne et l'élution a lieu au cours d'un gradient de 0 à 100 % de tampon B en 8 volumes de colonne.

Pour la protéine A2, une colonne His trap FF de 5 mL est utilisée. La protéine est injectée à 1 mL/min, lavée avec le tampon A pendant 8 volumes de colonne dont 3 volumes de colonne avec un tampon Tris 50 mM pH 8,0, NaCl 1 M pour éliminer l'ADN associé à la protéine. L'élution a lieu de 0 à 100 % de tampon B en 10 volumes de colonne.

Protéine	Tampon A	Tampon B
A1	Tris 50 mM pH 7,4, NaCl 1 M, Imidazole 20 mM	Tris 50 mM pH 7,4, NaCl 1 M, Imidazole 1 M
A2	Tris 50 mM pH 8,0, NaCl 150 mM, Imidazole 20 mM	Tris 50 mM pH 8,0, NaCl 150 mM, Imidazole 1 M

**Tableau 6. Tampons utilisés pour la purification des protéines sur colonne His trap.**

### 3.2.3 *Deuxième étape de purification de la protéine A1 par chromatographie d'affinité (colonne Héparine)*

La deuxième étape de purification de la protéine A1 est faite sur une colonne Héparine HP (GE Healthcare) de 1 mL qui retient les protéines interagissant avec l'ADN. Les fractions contenant la protéine A1 obtenues après la purification sur His Trap FF sont groupées et diluées à un facteur 1/5 avec un tampon Tris 50 mM pH 8,0 ; NaCl 62,5 mM afin d'atteindre une concentration finale en NaCl de 250 mM. Le mélange est fait à l'aide d'une pipette, de manière à ne pas provoquer une chute locale trop rapide de la concentration en NaCl qui déclencherait la précipitation de A1.

Un gradient linéaire de 0 à 100 % de tampon B contenant 1 M de NaCl permet l'élution de la protéine par compétition.

**Tampon A** : Tris 50 mM pH 8,0, NaCl 250 mM. **Tampon B** : Tris 50 mM pH 8,0 ; NaCl 1 M.

La protéine est injectée à 0,5 mL/min, la colonne est ensuite lavée avec le tampon A pendant 12 volumes de colonne et l'élution est obtenue par un gradient de 0 à 100 % de tampon B en 13,4 volumes de colonne.

### 3.2.4 *Purification par chromatographie d'échange de cations de la protéine A2*

La deuxième étape de purification de la protéine A2 est faite sur une colonne d'échange de cations HiTrap Sp HP (GE Healthcare) de 1 mL qui retient les protéines chargées positivement.

Un gradient linéaire de 0 à 100 % de tampon B contenant 1 M de NaCl permet l'élution de la protéine par compétition.

**Tampon A** : Phosphate 25 mM pH 7,0. **Tampon B** : Phosphate 25 mM pH 7,0, NaCl 1 M.

Les fractions de la première étape de purification contenant la protéine A2 sont injectées à 1 mL/min. La colonne est lavée avec le tampon A pendant 10 volumes de colonne et l'élution de A2 a lieu au cours d'un gradient de 0 à 100 % de tampon B en 10 volumes de colonne.

### 3.2.5 *Chromatographie d'exclusion par la taille*

La Chromatographie d'exclusion par la taille permet d'estimer la masse moléculaire pour déterminer l'état d'oligomérisation de la protéine, ou elle permet d'obtenir une fraction monodisperse pour des études structurales. Les protéines sont séparées en fonction de leur taille et de leur forme. L'échantillon subit une dilution d'un facteur environ 10 en sortie de colonne,

il faut donc parfois concentrer la protéine avant son injection. Lorsque c'est nécessaire, la protéine est concentrée sur un centricon Vivaspin (Sartorius) avec un seuil de poids moléculaire de 3000 Da pour la protéine A2 (14 kDa) et de 30000 Da pour la protéine A1 (62 kDa).

A1 est purifiée sur colonne Superdex 200 10/300 GL. A2 est purifiée sur colonne Superdex 75 10/300 GL (GE Healthcare).

Tampon d'élution de A1 : Tris 50 mM pH 8,0, NaCl 500 mM.

Tampon d'élution de A2 : Tris 50 mM pH 7,2, NaCl 100 mM.

### *3.2.6 Vérification des protéines obtenues par migration sur gel d'acrylamide*

Les protéines purifiées sont déposées sur gel dénaturant d'acrylamide 12 % (SDS-PAGE). La migration a lieu pendant 45 minutes à 30 mA. Les protéines sont révélées par coloration avec du bleu de Coomassie.

L'identité des protéines peut être vérifiée par Western Blot avec des anticorps anti-A1 ou des anticorps anti-A2.

## **4) Tests de toxicité**

La toxicité de la protéine A1 a été testée en comparant la croissance de différentes cultures en milieu liquide. Les mesures ont été faites sur un Tecan Infinite F200 Pro.

Des bactéries BL21(DE3) sont transformées avec chaque plasmide à tester et cultivées O/N dans un milieu LB + Antibiotique + Glucose 0,5 %. Pour obtenir des cellules en phase exponentielle pour le test, des cultures de 50 µL pour chaque condition testée sont effectuées en milieu LB + Antibiotique + Glucose 0,5 % dans une plaque 96 puits pendant 1 H à 37°C avec agitation à 220 rpm.

10 µL de bactéries en phase exponentielle sont utilisées pour inoculer 190 µL de milieu de culture avec du LB + Antibiotique + répresseur/inducteur pour chaque condition testée sur une nouvelle plaque 96 puits. La plaque est insérée dans le Tecan et les cultures sont faites à 37°C O/N avec agitation à 220 rpm.

Les résultats obtenus sont analysés avec le programme Multi\_assays\_graph.py sur le site <https://github.com/RayCharlZ/PhageT5> dont le fonctionnement est expliqué dans le fichier readme.txt (L. Ramirez et L. Zangelmi).

## 5) SEC-MALLS

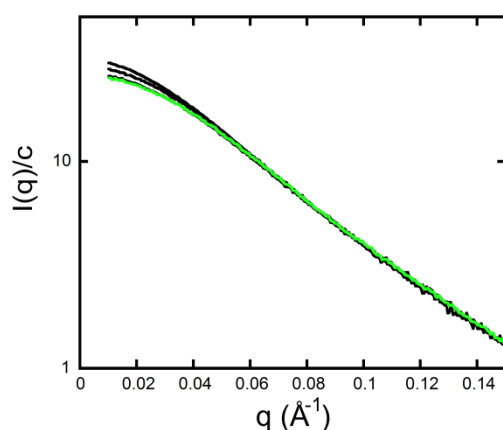
Les échantillons protéiques ont été analysés par Chromatographie d'exclusion par la taille (SEC) couplée avec des analyses de diffusion de lumière laser sous multiples angles (MALLS). La protéine A1 a été chargée sur une colonne Superdex 200 10/300 GL increase (GE Healthcare) équilibrée avec un tampon Tris 50 mM pH 8,0, NaCl 500 mM à 0,5 mL/min sur un système HPLC (Shimadzu). La diffusion de lumière statique a été mesurée avec un MiniDAWN TREOS (Wyatt Technology), la concentration de la protéine a été déterminée en utilisant un réfractomètre Optilab T-rEX (Wyatt Technology). Les données ont été analysées sur le logiciel ASTRA® 6.1.

## 6) Diffusion des rayons X aux petits angles (SAXS)

Les expériences de SAXS sur les protéines A1 et A2 ont été réalisées sur la ligne BM29 de l'ESRF (Grenoble, France). Le détecteur a été placé à 2867 mm de l'échantillon pour collecter les données dans la gamme  $0,005 \text{ \AA}^{-1} < q < 0,5 \text{ \AA}^{-1}$  à une longueur d'onde des rayons X  $\lambda = 0,992 \text{ \AA}$ . L'échantillon circule dans un capillaire de diamètre 1,8 mm.

La protéine A1 a été chargée sur une colonne Superose 6 increase 10/300 (GE Healthcare) équilibrée avec un tampon Tris 50 mM pH 7,5, NaCl 700 mM à 0,5 mL/min sur un système HPLC. La protéine éluée a été analysée sur une cellule de mesure UV puis sur une cellule de mesure SAXS. L'analyse des courbes de SAXS est explicitée dans le texte, et résumée dans le **Tableau 7**.

La protéine A2 a été injectée directement dans la cellule de mesure SAXS, à différentes concentrations comprises entre 4,3 et 0,93 mg/ml. On observe une légère évolution avec la concentration et la courbe  $I(q)$  analysée est une courbe obtenue en extrapolant à concentration nulle (**Figure 94**).



**Figure 94.** Superposition des courbes de diffusion expérimentales normalisées à la concentration (courbes noires), mesurées pour les trois concentrations 4,30, 2,07 et 0,93 mg/mL et de la courbe extrapolée à concentration nulle (courbe verte).

	A1 sans MnCl <sub>2</sub>	A1 avec MnCl <sub>2</sub>	A2
<b>Paramètres d'acquisition</b>			
Instrument	BM29 (ESRF)		
Détecteur	Pilatus 1M		
Taille du faisceau	0,7mm x 0,7mm		
Longueur d'onde [Å]	0,992		
Domaine en q [Å <sup>-1</sup> ]	0,004 < q < 0,50		
Temps d'exposition [s]	1		
Température [K]	293		
HPLC en ligne <sup>a</sup>	oui	oui	non <sup>b</sup>
<b>Paramètres structuraux</b>			
R <sub>g</sub> Guinier [Å] <sup>c</sup>	33,5 ± 0,4	60,0 ± 0,6	28,3 ± 0,3
R <sub>g</sub> p(r) [Å]	35,3 ± 0,4	60,7 ± 0,6	29,2 ± 0,3
D <sub>Max</sub> [Å]	136 ± 5	204 ± 10	110 ± 5
<b>Détermination de la masse</b>			
MM <sub>sequence</sub> [kDa] <sup>d</sup>	61,7	494 (octamère)	30,3 (dimère)
MM <sub>Scatter</sub> [kDa] <sup>e</sup>	63,0	529	30,1
<b>Analyse <i>ab initio</i> (GASBOR)</b>			
Symétrie	non	P42	P2
Nombre de modèles <sup>f</sup>	10	10	10
χ <sup>2</sup>	0,6-0,7	1,1-1,2	2,1-2,5

**Tableau 7. Principales caractéristiques des courbes de SAXS des protéines A1 et A2.**

<sup>a</sup> Colonne de gel filtration utilisée : Superose 6 Increase (10x300 mm) ; flux=0.5ml/mn

<sup>b</sup> Gamme de concentration : 0.93 – 4.0 mg/ml

<sup>c</sup> R<sub>g</sub> obtenu à partir de l'approximation de Guinier dans le domaine qR<sub>g</sub><1.2.

<sup>d</sup> Masse calculée à partir de la séquence

<sup>e</sup> Masse molaire déduite de la courbe expérimentale à l'aide du programme ScÅtter3 disponible sur <https://bl1231.als.lbl.gov/scatter/>. Cette détermination utilise la méthode du volume de corrélation et est indépendante de la concentration. Elle est intégrée dans le module HPLC du programme US-SOMO.

<sup>f</sup> Les différents modèles sont comparés à l'aide de la suite de programmes DAMAVER<sup>139</sup>. Le plus typique est présenté dans les figures.

## 7) Dichroïsme Circulaire (CD)

Les expériences de CD ont été réalisées à l'Institut des Neurosciences Paris-Saclay sur un spectropolarimètre Jasco J-810 à 20°C en tampon Tris 50 mM pH 8,0 NaCl 600 mM MnCl<sub>2</sub> 1 mM. Les échantillons ont été scannés entre 185 et 260 nm à la vitesse de 100 nm/min. 10 accumulations ont été faites sur un trajet optique de 0,01 mm.

## 8) Microscopie électronique à coloration négative

Les expériences de microscopie électronique ont été réalisées à la plateforme du département de Virologie de l'I2BC. L'assemblage de la protéine A1 sous forme d'oligomère est obtenu par incubation de A1 à une concentration de 11 μM en présence de MnCl<sub>2</sub> 1 mM ou d'ADN simple-



brin ou double-brin de différentes tailles. La protéine est déposée sur une grille de microscopie électronique hydrophile recouverte d'un film de carbone à une concentration finale de 0,6  $\mu\text{M}$  dans un tampon Tris 50 mM pH 7,4, NaCl 100 mM, +/-  $\text{MnCl}_2$  1 mM. La grille est colorée dans de l'acétate d'Uranyle 1 %. Les observations ont été faites sur un microscope électronique à transmission JEM-1400Plus (JEOL) à 80 kV avec une caméra à détection directe d'électrons K2 Summit GATAN.

## 9) Cryo-microscopie électronique

La protéine A1 a été incubée à 11  $\mu\text{M}$  dans un tampon Tris 50 mM pH 8,0, NaCl 500 mM,  $\text{MnCl}_2$  1 mM à 4°C O/N. Avant d'être déposée sur une grille d'observation, la protéine a été diluée à une concentration finale de 1,5  $\mu\text{M}$  dans un tampon Tris 50 mM pH 7,4, NaCl 100 mM,  $\text{MnCl}_2$  1 mM. La congélation de l'échantillon est faite dans de l'éthane liquide à -160°C.

La structure 3D de la protéine A1 a été résolue par cryo-microscopie électronique par Jean Lepault, Stéphane Roche et Rafael Navaza.

L'observation des échantillons et l'acquisition des images a été faite à l'Institut de Biologie Structurale à Grenoble sur un microscope Tecnai POLARA à 300 kV avec une caméra à détection directe d'électrons K2 Summit GATAN.

La carte de densité électronique a été construite en moyennant 40000 images, sur le logiciel Relion.

## 10) Cristallographie aux rayons X

Les cristaux de la protéine A1 ont été analysés aux synchrotrons SOLEIL et ESRF. Quatre jeux de données anormales ont été collectés sur quatre cristaux différents, trois sur la ligne PROXIMA-1 et un à PROXIMA-2. Les collectes anormales ont été faites à la longueur d'ondes  $\lambda_1 = 1,892 \text{ \AA}$  au seuil de la diffusion anormale des ions  $\text{Mn}^{2+}$ . Ensuite, plusieurs jeux de données natifs ont été collectés aux longueurs d'ondes standard  $\lambda_2 = 0,9786 \text{ \AA}$  et  $0,9763 \text{ \AA}$  à PROXIMA-1 et ID23-1. Le meilleur cristal a diffracté à 3,3  $\text{Å}$  de résolution.

La résolution et l'affinement de la structure sont faits par Inès Gallay dans l'équipe Fonction et Architecture des Assemblages Macromoléculaires (FAAM) à l'Institut de Biologie Intégrative de la Cellule (I2BC). Les données ont été intégrés avec le programme XDS (voir statistiques dans le **Tableau 8**). Le programme SHELX a été utilisé pour déterminer les positions des atomes lourds : 4 positions par unité asymétrique pour les ions  $\text{Mn}^{2+}$ . Les positions ont ensuite été affinées avec le programme PHASER et les phases initiales ont été modifiées par

	Native	Mn derivative
<b>Data collection</b>		
X-ray source	PX1, SOLEIL	PX1, SOLEIL
Wavelength (Å)	0.979	1.892
Space group	P6 <sub>1</sub> 22	P6 <sub>1</sub> 22
Cell dimensions		
<i>a, b, c</i> (Å)	230.4, 230.4, 314.1	230.4, 230.4, 314.1
Resolution (Å)	48.52–3.3 (3.5–3.3)	49.9–3.71 (3.81–3.71)
Total reflections	833,616	4,500,869
Unique reflections	75,220	98,678
<i>R</i> <sub>meas</sub>	0.17 (2.35)	0.36 (2.98)
<i>I</i> /σ <i>I</i>	14.8 (1.22)	12.18 (1.01)
Completeness (%)	99.6 (98.1)	99.9 (99.8)
Redundancy	11.1 (10.9)	45.6 (36.0)
CC <sub>1/2</sub> (%)	99.9 (53.3)	99.7 (42.9)
Anomal Corr*		41 (14)
<b>Refinement</b>		
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> (%)	23.3/24.5	
No. atoms		
Protein	14448	
Ligand/ion	4	
<i>B</i> -factors (Å <sup>2</sup> )		
from Wilson plot	138	
mean B value	131	
Root mean square deviations		
Bond lengths (Å)	0.008	
Bond angles (°)	1.19	
Ramachandran (%)		
Favored	85.6	
Allowed	8.2	
Outliers	6.2	

**Tableau 8. Récapitulatif des paramètres utilisés en cristallographie aux rayons X.** \* Entre parenthèses, valeur de Corrélation Anomale à 5 Å.

aplatissement du solvant avec l'aide du programme PARROT afin de rendre la carte de densité électronique mieux interprétable. Les statistiques concernant l'affinement de la structure cristalline indiquées dans le Tableau 8 reflètent l'état actuel de la structure et peuvent évoluer car le processus de reconstruction et l'affinement de la structure cristalline sont encore en cours.

## 11) Tests d'activité DNase

### 11.1 ADN utilisés

Les ADN marqués avec une molécule fluorescente d'Isothiocyanate de fluorescéine (FITC) utilisés pour les tests d'activité DNase ont tous une taille de 75 bases.

Des versions double-brin ont été construites en incubant 10  $\mu$ M d'ADN marqué au FITC avec 12  $\mu$ M d'ADN complémentaire non marqué dans un tampon Tris 10 mM pH 8,3, KCl 50 mM, de 98°C à 20°C à -1°C / 5min. Le marquage FITC est surligné en vert dans les oligonucléotides présentés.

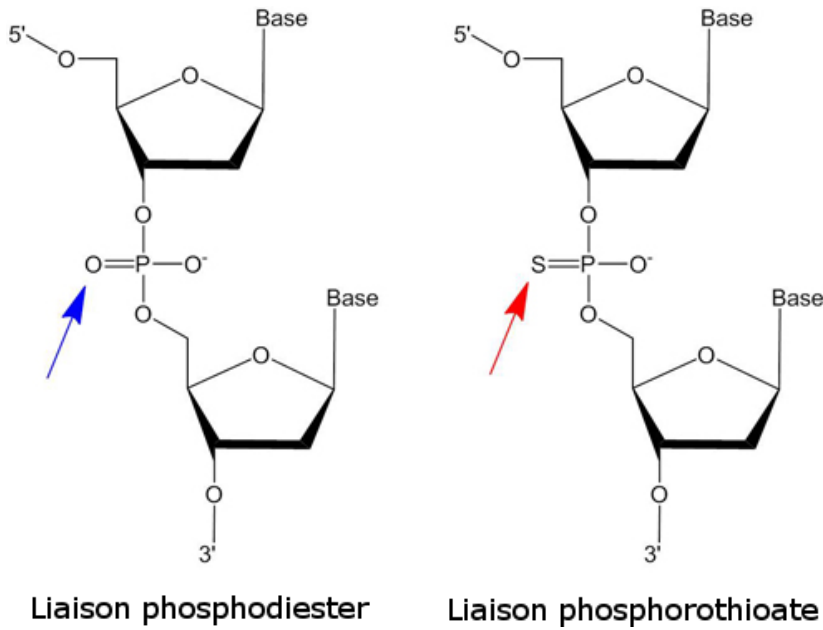
```
FITC-5' 5' FITC - ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACTTTTATAAACTGC 3'
FITC-3' 5' GCAGTTTATAAAAGTTACTATTTTGGGCTATTCAAACAATTATATCCGAACGCTGATTTTACAATGATTGATGAT - FITC 3'
Mid-FITC 5' ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTG - FITC - TTTGAATAGCCCAAAATAGTAACTTTTATAAACTGC 3'
No-FITC 5' ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACTTTTATAAACTGC 3'
```

L'échelle d'ADN utilisée dans les gels avec des ADN marqués au FITC est constituée de l'ADN FITC-3' (75 bases), des ADN marqués au FITC en 5' de 38 bases, 20 bases ou 5 bases et de l'ADN FITC-3' dégradé par ExoI.

L'ADN Phosphorothioate (PTO) contient un Soufre non-pontant à la place d'un Oxygène non-pontant sur le phosphate entre chaque base (**Figure 95**). La liaison PTO est non-hydrolysable par la plupart des DNases. Les liaisons PTO sont surlignées en jaune dans les oligonucléotides suivants :

```
PTO 3' 5' ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACTTTTATAAAC TGC 3'
PTO 5' 5' ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACTTTTATAAACTGC 3'
PTO 3' 5' 5' ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACTTTTATAAAC TGC 3'
PTO 25-25-25 5' ATCATCAATCATTGTAAAAATCAGCG TTCGGATATAATTGTTTGAATAGCC CAAAATAGTAACTTTTATAAACTGC 3'
PTO 15-25-35 5' ATCATCAATCATTGT AAAATCAGCGTTCGGATATAATTGT TTTGAATAGCCCAAAATAGTAACTTTTATAAACTGC 3'
PTO 35-35-15 5' ATCATCAATCATTGTAAAAATCAGCGTTCGGATATA ATTGTTTGAATAGCCCAAAATAGTAACTTTTATAAACTGC 3'
```

L'échelle d'ADN utilisée dans les gels avec des ADN PTO est constituée de l'ADN PTO 15-25-35 (75 bases, 170 fmol déposés sur gel), d'ADN classiques de 62 (210 fmol), 40 (330 fmol) et 35 bases (330 fmol) et d'ADN composés uniquement de liaisons PTO de 25 (560 fmol) et 15 bases (830 fmol).



**Figure 95. Modification de l'ADN par une liaison phosphorothioate.**

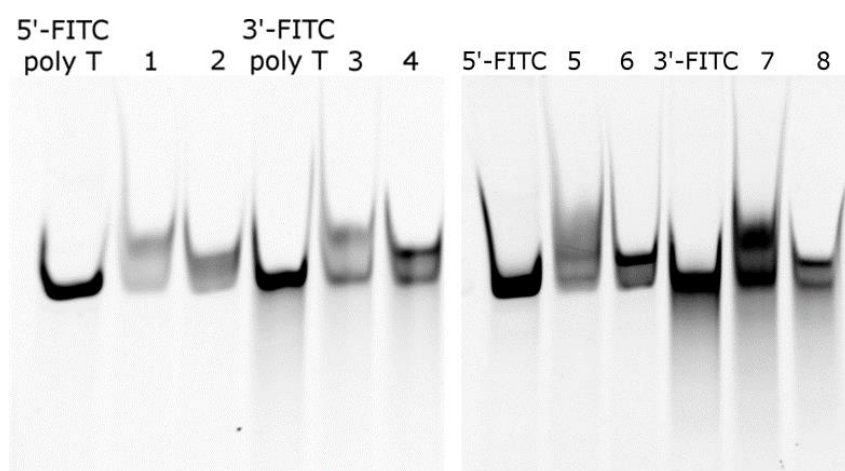
Des ADN marqués au FITC partiellement hybridés avec des ADN non marqués ont été construits dans les mêmes conditions que les autres ADN double-brin. Les parties complémentaires sont indiquées en vert, les parties non complémentaires sont indiquées en rouge. Le marquage FITC est surligné en vert.

1	5'	FITC	-	ATCATCAATCATTGTAAAAAT (T) 55	3'
	3'			TAGTAGTTAGTAACATTTTA (T) 55	5'
2	5'	FITC	-	ATCATCAATCATTGTAAAAAT (T) 55	3'
	3'			TAGTAGTTAGTAACATTTTA	5'
3	5'			55 (T) ATTTTACAATGATTGATGAT - FITC	3'
	3'			55 (T) TAAAAATGTTACTAECTACTA	5'
4	5'			55 (T) ATTTTACAATGATTGATGAT - FITC	3'
	3'			TAAAAATGTTACTAECTACTA	5'
5	5'	FITC	-	ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACCTTTATAAAGTGC	3'
	3'			TAGTAGTTAGTAACATTTTACAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACCTTTATAAAGTGC	5'
6	5'	FITC	-	ATCATCAATCATTGTAAAAATCAGCGTTCGGATATAATTGTTTGAATAGCCCAAAATAGTAACCTTTATAAAGTGC	3'
	3'			TAGTAGTTAGTAACATTTTA	5'
7	5'	GCAGTTTATAAAAGTTACTATTTTGGGCTATTCAAACAATTATATCCGACGCTGATTTTACAATGATTGATGAT - FITC			3'
	3'	GCAGTTTATAAAAGTTACTATTTTGGGCTATTCAAACAATTATATCCGACGCTGAAAAATGTTACTAECTACTA			5'
8	5'	GCAGTTTATAAAAGTTACTATTTTGGGCTATTCAAACAATTATATCCGACGCTGATTTTACAATGATTGATGAT - FITC			3'
	3'	TAAAAATGTTACTAECTACTA			5'

Les ADN sont vérifiés par migration sur gel d'acrylamide natif 16 % (**Figure 96**). Le tampon de migration est un tampon TBE 1 X : 100 mM Tris pH 8,0, 100 mM acide borique, 2 mM EDTA.

Composé	Volume
Acrylamide 40 % 29 : 1	2,82 mL
Tampon TBE 5 X	1,5 mL
Persulfate d'Ammonium 10 %	37,5 µL
TEMED	3,75 µL

**Tableau 9. Composition du gel d'acrylamide non-dénaturant pour la migration d'ADN.**



**Figure 96. Vérification des ADN hybridés par migration sur gel d'acrylamide natif 16 %, 250 V 35 min. Il existe une fraction non hybridée dans tous les cas.**

Des essais ont également été faits sur l'ADN de M13mp18 RFI (NEB) ou sur différents plasmides.

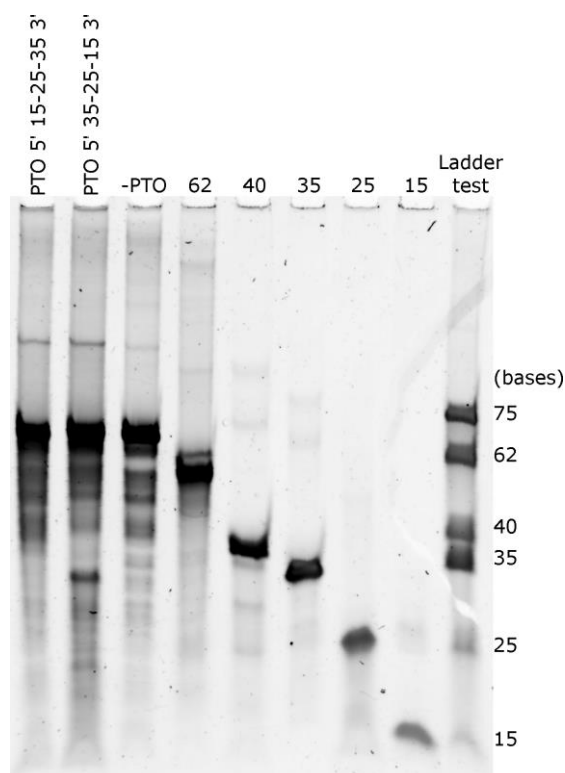
## 11.2 Test du tampon de charge

Les tests d'activité ont été faits en incubant différents substrats d'ADN simple ou double brin en présence de la protéine A1 dans des conditions variables. Les résultats de la digestion nucléasique ont été analysés sur gel d'acrylamide 16 % urée 8 M (**Tableau 10**). Au cours des premiers essais, l'ADN simplement incubé avec A1 était absent du le gel, même en absence de cation divalent. Nous avons fait l'hypothèse que A1 formait un complexe avec l'ADN qui ne peut pas être prélevé par "pipetage". Pour récupérer l'ADN agrégé avec A1, j'ai mis au point un protocole d'arrêt de la réaction et de solubilisation de l'échantillon avec un tampon de charge chauffé à 98°C contenant 8 M d'urée, 0,75 % de SDS, 40 % de formamide et 1,5 mM d'EDTA, coloré avec 0,1 % de bleu de bromophénol (Tampon USFE).

Composé	Quantité
Acrylamide 40 % 29 : 1	2,82 mL
Urée	2,7 g
Tampon TBE 5 X	1,5 mL
Persulfate d'Ammonium 10 %	37,5 µL
TEMED	3,75 µL

**Tableau 10. Composition du gel d'acrylamide dénaturant pour la migration d'ADN.**

Pour vérifier que ce tampon ne modifie pas le profil de migration de l'ADN simple-brin, un test de migration a été fait avec des ADN connus (**Figure 97**).



**Figure 97. Migration de différents ADN simple brin de tailles connues dans un gel acrylamide 16 %, Urée 8M, TBE 1,5 X. Dénaturation des échantillons à 98°C pendant 5 min dans le tampon USFE (Urée 8 M, SDS 0,75 %, formamide 40 %, EDTA 1,5 mM, bleu de bromophénol 0,1 %), puis refroidissement dans la glace avant dépôt sur gel.**

Les ADN traités au tampon USFE sont bien visibles aux positions attendues, identiques à leurs positions quand ils sont non traités. Ce résultat indique que ce tampon très dénaturant ne dégrade pas les ADN. J'ai donc utilisé ce tampon de charge dans toutes les expériences avec des petits oligonucléotides.

### 11.3 Test d'activité

- A1 aliquotée et conservée à -20°C dans le tampon Tris 50 mM pH 8,0, NaCl 500 mM est diluée à une concentration de 1,1 µM dans un tampon Tris 185 mM pH 7,4, NaCl 370 mM, DTT 3,70 mM. Concentration finale de 0,3 µM.
- L'ADN est dilué dans de l'eau distillée (qualité Milli-Q) à 10 µM. Concentration finale de 1 µM.

- A1 et l'ADN sont mélangés en premier. L'ajout de  $\text{MnCl}_2$  provoque le démarrage de la réaction ( $t = 0$ ). Concentration finale de 1 mM.
- Le tampon de réaction contient : Tris 50 mM pH 7,4, NaCl 100 mM, DTT 1 mM,  $\text{MnCl}_2$  1 mM.

La plupart des expériences ont été effectuées avec un rapport molaire [monomère A1]/[pb ADN] = 1/250 dans un thermocycleur réglé à la température d'incubation désirée.

La réaction est stoppée avec 2 volumes de tampon USFE et par une dénaturation de 5 minutes à 98°C, puis refroidissement dans la glace. Le tampon ne contient pas de bleu de bromophénol lorsque des ADN fluorescents sont utilisés à cause d'artefacts visibles sur le gel.

Pour les tests d'activité faits sur les ADN circulaires double-brin, les mêmes concentrations de A1 et  $\text{MnCl}_2$  sont utilisées. L'ADN est utilisé à 75  $\mu\text{M}$  de paires de bases dans un volume réactionnel final de 10  $\mu\text{L}$ . La réaction est stoppée par l'ajout de 2  $\mu\text{L}$  d'un tampon de charge contenant 30 % de glycérol, 0,6 % de bleu de bromophénol et 60 mM d'EDTA.

#### 11.4 Migration sur gel et révélation

La migration a lieu dans un gel acrylamide 16 %, Urée 8M, TBE 1,5 X. Les ADN PTO sont dilués au 1/5 et 3  $\mu\text{L}$  sont déposés sur gel (0,6 pmol). Les ADN fluorescents ne sont pas dilués et 10  $\mu\text{L}$  sont déposés sur gel (10 pmol). Une migration préliminaire a lieu sans ADN à 250 V pendant 15 minutes, puis les ADN sont déposés et la migration est effectuée à 250 V pendant 35 minutes. Les ADN non fluorescents sont colorés avec une solution aqueuse de SYBR Gold (Invitrogen) à une concentration finale de 1 X. Les ADN fluorescents peuvent être directement observés.

Les gels colorés sont observés avec un appareil LAS-4000 mini émettant de la lumière bleue et équipé d'une plaque filtrant les UV. Les gels sont exposés 15 secondes au paramètre "high resolution".

Pour les essais sur les ADN circulaires double-brin, les produits de réaction sont déposés sur gel d'agarose 1%. La migration a lieu à 135 V pendant 25 min. Les gels sont colorés au Bromure d'Ethidium (BET) et révélés sous lumière UV.

## 12) Tests de complémentation

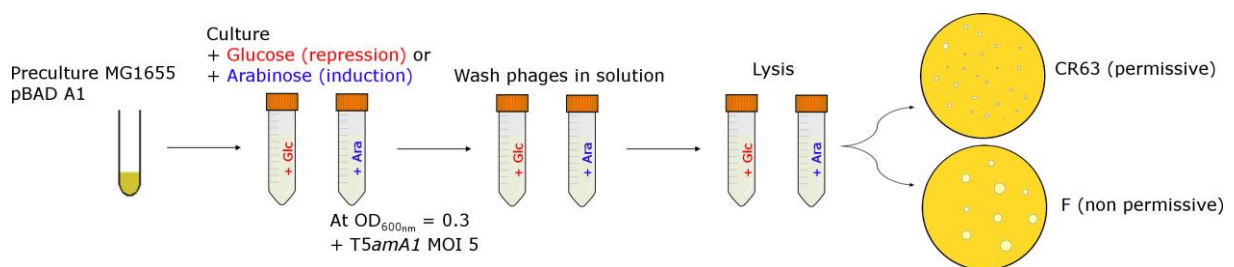
Les tests de complémentation fonctionnelle des mutants de phages avec la protéine recombinante A1 n'ont pas pu être effectués par observation des plages de lyse sur un tapis

bactérien exprimant A1 sur boîte de Pétri en raison de la toxicité de A1. En effet, l'induction de la production de A1 par l'arabinose dans des bactéries transformées avec pBAD A1 inhibe la croissance bactérienne en quelques heures. Pour contourner le problème de toxicité de A1, un test de complémentation en milieu liquide a été développé. Les phages sont produits en milieu liquide sur une souche MG1655 portant la construction à tester, puis ils sont titrés en milieu solide sur une souche CR63 permettant le développement des mutants ou une souche F ne permettant pas leur développement.

Les bactéries MG1655 contenant le plasmide pBAD24 exprimant la version de la protéine A1 à tester sont cultivées dans 15 mL de milieu LB + Antibiotique + CaCl<sub>2</sub> 1 mM + MgCl<sub>2</sub> 1 mM +/- Glucose 0,2 % à 37°C. L'expression de la protéine clonée est induite par l'ajout d'arabinose 0,2 % et réprimée par l'ajout de Glucose 0,2 %.

Lorsque la densité optique de la culture à 600 nm (OD<sub>600nm</sub>) atteint 0,3, 0,2 % d'arabinose sont ajoutés dans les tubes sans Glucose pour induire l'expression de la protéine, puis les bactéries sont cultivées 10 minutes supplémentaires à 37°C. Le phage T5 mutant (T5amH27, mutant ambre dans A1 pour la Glutamine Q238) est ajouté à une Multiplicité d'infection (MOI) de 5, 37°C 10 min. La culture bactérienne ainsi infectée est lavée pour être débarrassée des phages non adsorbés, par 3 étapes de centrifugation 8 min à 8000 g et resuspension dans 15 mL de milieu de culture.

La lyse est suivie pendant 90 minutes à 37°C par mesure de DO à 600 nm puis les débris cellulaires sont éliminés par centrifugation 8 min à 8000 g. Les phages récoltés dans le surnageant sont dilués dans un tampon Tris 10 mM pH 7,2, NaCl 100 mM, CaCl<sub>2</sub> 1 mM, MgCl<sub>2</sub> 1 mM puis titrés en milieu solide sur la souche *E. Coli* indicatrice CR63 ou F.



**Figure 98. Schéma récapitulatif du protocole de complémentation en milieu liquide.**

### 13) Résonance magnétique nucléaire (RMN)

La protéine A2 a été produite en milieu minimum M9 enrichi en <sup>15</sup>N avec du <sup>15</sup>NH<sub>4</sub>Cl marqué. La production a eu lieu dans les mêmes conditions que pour la production en milieu LB. La



protéine a été dialysée dans un tampon phosphate 25 mM pH 6,6, NaCl 100 mM et concentrée à 195  $\mu$ M sur un centricon Vivaspin (Sartorius). La pureté a été vérifiée par migration sur gel d'acrylamide 15 % et coloration au bleu de Coomassie.

Le spectre HSQC de RMN a été effectué au Laboratoire de Biologie Structurale et Radiobiologie du CEA de Saclay en collaboration avec Sophie Zinn-Justin.

---

## V. Références

---

1. Breitbart, M. & Rohwer, F. Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology* **13**, 278–284 (2005).
2. Sullivan, M. B., Weitz, J. S. & Wilhelm, S. Viral ecology comes of age: Crystal ball. *Environmental Microbiology Reports* **9**, 33–35 (2017).
3. Wigington, C. H. *et al.* Re-examining the relationship between virus and microbial cell abundances in the global oceans. (2015). doi:10.1101/025544
4. Suttle, C. A. Marine viruses — major players in the global ecosystem. *Nature Reviews Microbiology* **5**, 801–812 (2007).
5. Hatfull, G. F. & Hendrix, R. W. Bacteriophages and their Genomes. *Curr Opin Virol* **1**, 298–303 (2011).
6. Ackermann, H.-W. Bacteriophage Taxonomy. *Microbiology Australia* 90–94 (2011).
7. De Smet, J., Hendrix, H., Blasdel, B. G., Danis-Wlodarczyk, K. & Lavigne, R. Pseudomonas predators: understanding and exploiting phage–host interactions. *Nature Reviews Microbiology* **15**, 517–530 (2017).
8. Yap, M. L. & Rossmann, M. G. Structure and function of bacteriophage T4. *Future Microbiol* **9**, 1319–1327 (2014).
9. Tolstoy, I., Kropinski, A. M. & Brister, J. R. Bacteriophage Taxonomy: An Evolving Discipline. *Methods Mol. Biol.* **1693**, 57–71 (2018).
10. Kruger, D., Schneck, P. & Gelderblom, H. Helmut Ruska and the visualisation of viruses. *The Lancet* **355**, 1713–1717 (2000).
11. Luria, S. E. & Anderson, T. F. The Identification and Characterization of Bacteriophages with the Electron Microscope. *Proceedings of the National Academy of Sciences* **28**, 127-130.1 (1942).
12. Luria, S. E., Delbrück, M. & Anderson, T. F. Electron Microscope Studies of Bacterial Viruses1. *J Bacteriol* **46**, 57–77 (1943).
13. Ventola, C. L. The Antibiotic Resistance Crisis. *P T* **40**, 277–283 (2015).
14. *Antibiotic Resistance Threats in the United States*. (Centers for Disease Control and Prevention, 2013).
15. Carlet, J. & Le Coz, P. Rapport du groupe de travail spécial pour la préservation des antibiotiques. Paris : Ministère des Affaires sociales, de la Santé et des Droits des femmes (2015).
16. Barbu, E. M., Cady, K. C. & Hubby, B. Phage Therapy in the Era of Synthetic Biology. *Cold Spring Harbor Perspectives in Biology* **8**, a023879 (2016).

17. Ando, H., Lemire, S., Pires, D. P. & Lu, T. K. Engineering Modular Viral Scaffolds for Targeted Bacterial Population Editing. *Cell Syst* **1**, 187–196 (2015).
18. Parasion, S., Kwiatek, M., Gryko, R., Mizak, L. & Malm, A. Bacteriophages as an alternative strategy for fighting biofilm development. *Pol. J. Microbiol.* **63**, 137–145 (2014).
19. Comeau, A. M. Phage-Antibiotic Synergy (PAS): b-Lactam and Quinolone Antibiotics Stimulate Virulent Phage Growth. *PLoS ONE* **4** (2007).
20. Eaton, M. D. & Bayne-Jones, S. BACTERIOPHAGE THERAPY: REVIEW OF THE PRINCIPLES AND RESULTS OF THE USE OF BACTERIOPHAGE IN THE TREATMENT OF INFECTIONS. *Journal of the American Medical Association* **103**, 1769 (1934).
21. Smith, H. W. & Huggins, M. B. Successful treatment of experimental Escherichia coli infections in mice using phage: its general superiority over antibiotics. *J. Gen. Microbiol.* **128**, 307–318 (1982).
22. Smith, H. W. & Huggins, M. B. Effectiveness of Phages in Treating Experimental Escherichia coli Diarrhoea in Calves, Piglets and Lambs. *Microbiology* **129**, 2659–2675 (1983).
23. Rhoads, D. D. *et al.* Bacteriophage therapy of venous leg ulcers in humans: results of a phase I safety trial. *Journal of Wound Care* **18**, 237–243 (2009).
24. Dufour, N. & Debarbieux, L. La phagothérapie - Une arme crédible face à l'antibiorésistance. *Med Sci (Paris)* **33**, 410–416 (2017).
25. Nechaev, S. & Severinov, K. The Elusive Object of Desire - Interactions of Bacteriophages and Their Hosts. *Curr Opin Microbiol* **11**, 186–193 (2008).
26. Scott, A. E. *et al.* Genome dynamics of Campylobacter jejuni in response to bacteriophage predation. *PLoS Pathog.* **3**, e119 (2007).
27. Brussow, H., Canchaya, C. & Hardt, W.-D. Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiology and Molecular Biology Reviews* **68**, 560–602 (2004).
28. Lindell, D. *et al.* Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* **449**, 83–86 (2007).
29. Soni, K. A., Nannapaneni, R. & Hagens, S. Reduction of *Listeria monocytogenes* on the Surface of Fresh Channel Catfish Fillets by Bacteriophage Listex P100. *Foodborne Pathogens and Disease* **7**, 427–434 (2010).

30. Carter, C. D. *et al.* Bacteriophage cocktail significantly reduces Escherichia coli O157. *Bacteriophage* **2**, 178–185 (2012).
31. Woolston, J. *et al.* Bacteriophages lytic for Salmonella rapidly reduce Salmonella contamination on glass and stainless steel surfaces. *Bacteriophage* **3**, (2013).
32. Sarhan, W. A. & Azzazy, H. M. Phage approved in food, why not as a therapeutic? *Expert Review of Anti-infective Therapy* **13**, 91–101 (2015).
33. Beke, G., Stano, M. & Klucar, L. Modelling the interaction between bacteriophages and their bacterial hosts. *Mathematical Biosciences* **279**, 27–32 (2016).
34. Wittmann, J., Brancato, C., Berendzen, K. W. & Dreiseikelmann, B. Development of a tomato plant resistant to *Clavibacter michiganensis* using the endolysin gene of bacteriophage CMP1 as a transgene. *Plant Pathology* **65**, 496–502 (2016).
35. Bertani, G. & Weigle, J. J. HOST CONTROLLED VARIATION IN BACTERIAL VIRUSES. *J Bacteriol* **65**, 113–121 (1953).
36. Luria, S. E. & Human, M. L. A NONHEREDITARY, HOST-INDUCED VARIATION OF BACTERIAL VIRUSES<sup>1</sup>. *J Bacteriol* **64**, 557–569 (1952).
37. Smith, H. O. & Welcox, K. W. A Restriction enzyme from Hemophilus influenzae. *Journal of Molecular Biology* **51**, 379–391 (1970).
38. Weiss, B. & Richardson, C. C. Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from Escherichia coli infected with T4 bacteriophage. *Proc Natl Acad Sci U S A* **57**, 1021–1028 (1967).
39. Collins, J. & Hohn, B. Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage lambda heads. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4242–4246 (1978).
40. Luria, S. E. & Delbrück, M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics* **28**, 491–511 (1943).
41. Hershey, A. D. & Chase, M. INDEPENDENT FUNCTIONS OF VIRAL PROTEIN AND NUCLEIC ACID IN GROWTH OF BACTERIOPHAGE. *J Gen Physiol* **36**, 39–56 (1952).
42. Crick, F. H. C., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General Nature of the Genetic Code for Proteins. *Nature* **192**, 1227–1232 (1961).
43. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).
44. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Current Opinion in Microbiology* **37**, 67–78 (2017).

45. Strotskaya, A. *et al.* The action of *Escherichia coli* CRISPR–Cas system on lytic bacteriophages with different lifestyles and development strategies. *Nucleic Acids Research* gkx042 (2017). doi:10.1093/nar/gkx042
46. Rath, D., Amlinger, L., Rath, A. & Lundgren, M. The CRISPR-Cas immune system: Biology, mechanisms and applications. *Biochimie* **117**, 119–128 (2015).
47. Zhu, B. Bacteriophage T7 DNA polymerase – sequenase. *Front Microbiol* **5**, (2014).
48. Fiers, W. *et al.* Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature* **260**, 500–507 (1976).
49. Sanger, F. *et al.* Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* **265**, 687–695 (1977).
50. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. & Petersen, G. B. Nucleotide sequence of bacteriophage  $\lambda$  DNA. *Journal of Molecular Biology* **162**, 729–773 (1982).
51. Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317 (1985).
52. Henry, K. A., Arbabi-Ghahroudi, M. & Scott, J. K. Beyond phage display: non-traditional applications of the filamentous bacteriophage as a vaccine carrier, therapeutic biologic, and bioconjugation scaffold. *Front Microbiol* **6**, (2015).
53. Hamzeh-Mivehroud, M., Alizadeh, A. A., Morris, M. B., Bret Church, W. & Dastmalchi, S. Phage display as a technology delivering on the promise of peptide drug discovery. *Drug Discovery Today* **18**, 1144–1157 (2013).
54. Van Dorst, B., Mehta, J., Rouah-Martin, E., Blust, R. & Robbens, J. Phage display as a method for discovering cellular targets of small molecules. *Methods* **58**, 56–61 (2012).
55. Smith, H. O., Hutchison, C. A., Pfannkoch, C. & Venter, J. C. Generating a synthetic genome by whole genome assembly:  $\phi$ X174 bacteriophage from synthetic oligonucleotides. *Proc Natl Acad Sci U S A* **100**, 15440–15445 (2003).
56. Fogg, P. C. M., Colloms, S., Rosser, S., Stark, M. & Smith, M. C. M. New Applications for Phage Integrases. *J Mol Biol* **426**, 2703–2716 (2014).
57. Kwan, K.-M. Conditional alleles in mice: Practical considerations for tissue-specific knockouts. *genesis* **32**, 49–62 (2002).
58. Shis, D. L. & Bennett, M. R. Library of synthetic transcriptional AND gates built with split T7 RNA polymerase mutants. *Proc Natl Acad Sci U S A* **110**, 5028–5033 (2013).
59. Cenens, W., Makumi, A., Govers, S. K., Lavigne, R. & Aertsen, A. Viral Transmission Dynamics at Single-Cell Resolution Reveal Transiently Immune Subpopulations Caused by a Carrier State Association. *PLOS Genetics* **11**, e1005770 (2015).

60. Bailly-Bechet, M., Vergassola, M. & Rocha, E. Causes for the intriguing presence of tRNAs in phages. *Genome Res* **17**, 1486–1495 (2007).
61. Nolan, J. M., Petrov, V., Bertrand, C., Krisch, H. M. & Karam, J. D. Genetic diversity among five T4-like bacteriophages. *Virology Journal* **3**, 30 (2006).
62. Roucourt, B. & Lavigne, R. The role of interactions between phage and bacterial proteins within the infected cell: a diverse and puzzling interactome. *Environ. Microbiol.* **11**, 2789–2805 (2009).
63. Zivanovic, Y. *et al.* Insights into bacteriophage T5 structure from analysis of its morphogenesis genes and protein components. *J. Virol.* **88**, 1162–1174 (2014).
64. Majtan, T., Halgasova, N., Bukovska, G. & Timko, J. Transcriptional profiling of bacteriophage BFK20: Coexpression interrogated by “guilt-by-association” algorithm. *Virology* **359**, 55–65 (2007).
65. Leskinen, K., Blasdel, B. G., Lavigne, R. & Skurnik, M. RNA-Sequencing Reveals the Progression of Phage-Host Interactions between  $\phi$ R1-37 and *Yersinia enterocolitica*. *Viruses* **8**, 111 (2016).
66. Ukah, U. V. *et al.* Risk factors for acquisition of multidrug-resistant *Escherichia coli* and development of community-acquired urinary tract infections. *Epidemiology and Infection* **146**, 46–57 (2018).
67. Effantin, G., Boulanger, P., Neumann, E., Letellier, L. & Conway, J. F. Bacteriophage T5 Structure Reveals Similarities with HK97 and T4 Suggesting Evolutionary Relationships. *Journal of Molecular Biology* **361**, 993–1002 (2006).
68. Flayhan, A., Wien, F., Paternostre, M., Boulanger, P. & Breyton, C. New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its *Escherichia coli* receptor FhuA. *Biochimie* **94**, 1982–1989 (2012).
69. Arnaud, C.-A. *et al.* Bacteriophage T5 tail tube structure suggests a trigger mechanism for Siphoviridae DNA ejection. *Nature Communications* **8**, 1953 (2017).
70. Böhm, J. *et al.* FhuA-mediated phage genome transfer into liposomes: a cryo-electron tomography study. *Curr. Biol.* **11**, 1168–1175 (2001).
71. McCORQUODALE, D. J., Shaw, A. R., Shaw, P. K. & Chinnadurai, G. Pre-Early Polypeptides of Bacteriophages T5 and BF23. *J. VIROL.* **22**, 9 (1977).
72. Davison, J. Pre-early functions of bacteriophage T5 and its relatives. *Bacteriophage* **5**, e1086500 (2015).
73. Lanni, Y. T. First-Step-Transfer Deoxyribonucleic Acid of Bacteriophage T5. *BACTERIOL. REV.* **32**, 16 (1968).

74. Pfefferkorn, E. & Amos, H. Deoxyribonucleic acid breakdown and resynthesis in T5 bacteriophage infection. *Virology* **6**, 299–301 (1958).
75. Crawford, L. V. Nucleic acid metabolism in *Escherichia coli* infected with phage T5. *Virology* **7**, 359–374 (1959).
76. Warner, H. R., Drong, R. F. & Berget, S. M. Early events after infection of *Escherichia coli* by bacteriophage T5. Induction of a 5'-nucleotidase activity and excretion of free bases. *J Virol* **15**, 273–280 (1975).
77. Lanni, Y. T. Invasion by bacteriophage T5: I. Some basic kinetic features. *Virology* **10**, 501–513 (1960).
78. Lanni, Y. T., Mccorquodale, D. J. & Wilson, C. M. MOLECULAR ASPECTS OF DNA TRANSFER FROM PHAGE T5 TO HOST CELLS. II. ORIGIN OF FIRST-STEP-TRANSFER DNA FRAGMENTS. *J. Mol. Biol.* **10**, 19–27 (1964).
79. Breyton, C. *et al.* Assessing the Conformational Changes of pb5, the Receptor-binding Protein of Phage T5, upon Binding to Its *Escherichia coli* Receptor FhuA. *Journal of Biological Chemistry* **288**, 30763–30772 (2013).
80. Boulanger, P. *et al.* Purification and Structural and Functional Characterization of FhuA, a Transporter of the *Escherichia coli* Outer Membrane †. *Biochemistry* **35**, 14216–14224 (1996).
81. Plançon, L., Chami, M. & Letellier, L. Reconstitution of FhuA, an *Escherichia coli* Outer Membrane Protein, into Liposomes: BINDING OF PHAGE T5 TO FhuA TRIGGERS THE TRANSFER OF DNA INTO THE PROTEOLIPOSOMES. *Journal of Biological Chemistry* **272**, 16868–16872 (1997).
82. Davison, J. & Brunel, F. Restriction insensitivity in bacteriophage T5 I. Genetic characterization of mutants sensitive to EcoRI restriction. *J. Virol.* **29**, 11–16 (1979).
83. Krüger, D. H. & Bickle, T. A. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol Rev* **47**, 345–360 (1983).
84. Stern, A. & Sorek, R. The phage-host arms-race: Shaping the evolution of microbes. *Bioessays* **33**, 43–51 (2011).
85. Sakaki, Y. Inactivation of the ATP-dependent DNase of *Escherichia coli* After Infection with Double-Stranded DNA Phages. *J Virol* **14**, 1611–1612 (1974).
86. Wang, J. *et al.* Complete genome sequence of bacteriophage T5. *Virology* **332**, 45–65 (2005).



87. Beckman, L. D., Witonsky, P. & McCorquodale, D. J. Effect of rifampin on the growth of bacteriophage T5. *J. Virol.* **10**, 179–186 (1972).
88. Beckman, L. D., Hoffman, M. S. & McCorquodale, D. J. Pre-early proteins of bacteriophage T5: Structure and function. *Journal of Molecular Biology* **62**, 551–564 (1971).
89. Sayers, J. R. Bacteriophage T5. in *The Bacteriophages* (ed. Abedon, S. T.) 268–276 (Oxford University Press, USA, 2006).
90. Parson, K. A. & Snustad, D. P. Host DNA Degradation after Infection of Escherichia coli with Bacteriophage T4: Dependence of the Alternate Pathway of Degradation Which Occurs in the Absence of Both T4 Endonuclease II and Nuclear Disruption on T4 Endonuclease IV. *J Virol* **15**, 221–224 (1975).
91. Bacteriophage T3 and bacteriophage T7 virus-host cell interactions. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC281497/>. (Accessed: 16th September 2018)
92. Everett, R. D. DNA replication of bacteriophage T5. 3. Studies on the structure of concatemeric T5 DNA. *J. Gen. Virol.* **52**, 25–38 (1981).
93. Bourguignon, G. J., Sweeney, T. K. & Delius, H. Multiple origins and circular structures in replicating T5 bacteriophage DNA. *Journal of Virology* **18**, 245–259 (1976).
94. Young, R. Phage lysis: Three steps, three choices, one outcome. *Journal of Microbiology* **52**, 243–258 (2014).
95. Jarábková, V., Tišáková, L. & Godány, A. Phage Endolysin: A Way To Understand A Binding Function Of C-Terminal Domains A Mini Review. *Nova Biotechnologica et Chimica* **14**, 117–134 (2015).
96. Tran, T. A. T., Struck, D. K. & Young, R. The T4 RI Antiholin Has an N-Terminal Signal Anchor Release Domain That Targets It for Degradation by DegP. *Journal of Bacteriology* **189**, 7618–7625 (2007).
97. Moussa, S. H., Kuznetsov, V., Tran, T. A. T., Sacchetti, J. C. & Young, R. Protein determinants of phage T4 lysis inhibition. *Protein Sci.* **21**, 571–582 (2012).
98. Mozer, T. J., Thompson, R. B., Berget, S. M. & Warner, H. R. Isolation and characterization of a bacteriophage T5 mutant deficient in deoxynucleoside 5'-monophosphatase activity. *J Virol* **24**, 642–650 (1977).
99. Shaw, A. R., Lang, D. & McCorquodale, D. J. Terminally redundant deletion mutants of bacteriophage BF23. *J Virol* **29**, 220–231 (1979).

100. Blaisdell, P. W. & Warner, H. R. Cell-free transcription and translation of isolated restriction fragments localize bacteriophage T5 pre-early genes. *J Virol* **57**, 759–764 (1986).
101. McCORQUODALE, D. J., Chen, C. W., Joseph, M. K. & Woychik, R. Modification of RNA Polymerase from Escherichia coli by Pre-Early Gene Products of Bacteriophage T5. **40**, 5 (1981).
102. Lanni, Y. T. Functions of two genes in the first-step-transfer DNA of bacteriophage T5. *Journal of Molecular Biology* **44**, 173–183 (1969).
103. McCorquodale, D. J. & Lanni, Y. T. Patterns of protein synthesis in Escherichia coli infected by amber mutants in the first-step-transfer DNA of T5. *Journal of Molecular Biology* **48**, 133–143 (1970).
104. Snyder, C. E. & Benzinger, R. H. Second-Step Transfer of Bacteriophage T5 DNA: Purification and Characterization of the T5 Gene A2 Protein. *J. VIROL.* **40**, 10 (1981).
105. Snyder, C. E. Amino acid sequence of the bacteriophage T5 gene A2 protein. *Biochemical and Biophysical Research Communications* **177**, 1240–1246 (1991).
106. Snyder, C. E. Bacteriophage T5 gene A2 protein alters the outer membrane of Escherichia coli. *J Bacteriol* **160**, 1191–1195 (1984).
107. Delattre, H., Souiai, O., Fagoonee, K., Guerois, R. & Petit, M.-A. Phagonaute: A web-based interface for phage synteny browsing and protein function prediction. *Virology* **496**, 42–50 (2016).
108. Arthur, L. M. *et al.* Structural and functional analysis of Mre11-3. *Nucleic Acids Res* **32**, 1886–1893 (2004).
109. Teklemariam, T. A., Rivera, O. D. & Nelson, S. W. Kinetic Analysis of the Exonuclease Activity of the Bacteriophage T4 Mre11–Rad50 Complex. in *Methods in Enzymology* **600**, 135–156 (Elsevier, 2018).
110. Zimmermann, L. *et al.* A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of Molecular Biology* **430**, 2237–2243 (2018).
111. Gill, J. J. *et al.* The Caulobacter crescentus phage phiCbK: genomics of a canonical phage. *BMC Genomics* **13**, 542 (2012).
112. Novy, R. & Morris, B. Use of glucose to control basal expression in the pET System. *inNovations* **13**, 8–10
113. Xiong, S., Zhang, L. & He, Q.-Y. Fractionation of Proteins by Heparin Chromatography. in *2D PAGE: Sample Preparation and Fractionation* (ed. Posch, A.) **424**, 213–221 (Humana Press, 2008).

114. Bolten, S. N., Rinas, U. & Scheper, T. Heparin: role in protein purification and substitution with animal-component free material. *Appl Microbiol Biotechnol* (2018). doi:10.1007/s00253-018-9263-3
115. Vestergaard, B. Analysis of biostructural changes, dynamics, and interactions – Small-angle X-ray scattering to the rescue. *Archives of Biochemistry and Biophysics* **602**, 69–79 (2016).
116. Brookes, E., Vachette, P., Rocco, M. & Pérez, J. US-SOMO HPLC-SAXS module: dealing with capillary fouling and extraction of pure component patterns from poorly resolved SEC-SAXS data. *J Appl Crystallogr* **49**, 1827–1841 (2016).
117. Franke, D., Jeffries, C. M. & Svergun, D. I. Correlation Map, a goodness-of-fit test for one-dimensional X-ray scattering spectra. *Nat. Methods* **12**, 419–422 (2015).
118. Svergun, D. I., Petoukhov, M. V. & Koch, M. H. Determination of domain structure of proteins from X-ray solution scattering. *Biophys J* **80**, 2946–2953 (2001).
119. Mészáros, B., Erdos, G. & Dosztányi, Z. IUPred2A: context-dependent prediction of protein disorder as a function of redox state and protein binding. *Nucleic Acids Res.* **46**, W329–W337 (2018).
120. Micsonai, A. *et al.* BeStSel: a web server for accurate protein secondary structure prediction and fold recognition from the circular dichroism spectra. *Nucleic Acids Res* **46**, W315–W322 (2018).
121. Micsonai, A. *et al.* Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E3095–3103 (2015).
122. Vanecko, S. & Laskowski, M. Studies of the Specificity of Deoxyribonuclease I III. HYDROLYSIS OF CHAINS CARRYING A MONOESTERIFIED PHOSPHATE ON CARBON 5'. *J. Biol. Chem.* **236**, 3312–3316 (1961).
123. Gray, V. E., Hause, R. J. & Fowler, D. M. Analysis of Large-Scale Mutagenesis Data To Assess the Impact of Single Amino Acid Substitutions. *Genetics* **207**, 53–61 (2017).
124. Eckstein, F. Phosphorothioates, Essential Components of Therapeutic Oligonucleotides. *Nucleic Acid Therapeutics* **24**, 374–387 (2014).
125. Chase, J. W. & Richardson, C. C. Exonuclease VII of *Escherichia coli*. in *Molecular Mechanisms for Repair of DNA* (eds. Hanawalt, P. C. & Setlow, R. B.) 225–234 (Springer US, 1975). doi:10.1007/978-1-4684-2895-7\_30
126. Lehman, I. R. & Nussbaum, A. L. The Deoxyribonucleases of *Escherichia coli* V. ON THE SPECIFICITY OF EXONUCLEASE I (PHOSPHODIESTERASE). *J. Biol. Chem.* **239**, 2628–2636 (1964).

127. Marceau, A. H. *et al.* Structure of the SSB–DNA polymerase III interface and its role in DNA replication. *The EMBO Journal* **30**, 4236 (2011).
128. Anjem, A., Varghese, S. & Imlay, J. A. Manganese import is a key element of the OxyR response to hydrogen peroxide in *Escherichia coli*. *Mol Microbiol* **72**, 844–858 (2009).
129. Cha, R. S., Zarbl, H., Keohavong, P. & Thilly, W. G. Mismatch Amplification Mutation Assay (MAMA): Application to the c-H-ras Gene. *8* (1992).
130. Kovall, R. & Matthews, B. W. Toroidal structure of lambda-exonuclease. *Science* **277**, 1824–1827 (1997).
131. Farley, M. M., Hu, B., Margolin, W. & Liu, J. Minicells, Back in Fashion. *Journal of Bacteriology* **198**, 1186–1195 (2016).
132. Djacem, K., Tavares, P. & Oliveira, L. Bacteriophage SPP1 pac Cleavage: A Precise Cut without Sequence Specificity Requirement. *Journal of Molecular Biology* **429**, 1381–1395 (2017).
133. Sirbasku, D. A. & Buchanan, J. M. Patterns of Ribonucleic Acid Synthesis in T5-infected *Escherichia coli* II. SEPARATION OF HIGH MOLECULAR WEIGHT RIBONUCLEIC ACID SPECIES BY DISC ELECTROPHORESIS ON ACRYLAMIDE GEL COLUMNS. *J. Biol. Chem.* **245**, 2679–2692 (1970).
134. McCorquodale, D. J. & Buchanan, J. M. Patterns of Protein Synthesis in T5-infected *Escherichia coli*. *J. Biol. Chem.* **243**, 2550–2559 (1968).
135. Connelly, J. C., Kirkham, L. A. & Leach, D. R. F. The SbcCD nuclease of *Escherichia coli* is a structural maintenance of chromosomes (SMC) family protein that cleaves hairpin DNA. *Proc Natl Acad Sci U S A* **95**, 7969–7974 (1998).
136. Thoms, B., Borchers, I. & Wackernagel, W. Effects of Single-Strand DNases ExoI, RecJ, ExoVII, and SbcCD on Homologous Recombination of recBCD+ Strains of *Escherichia coli* and Roles of SbcB15 and XonA2 ExoI Mutant Enzymes. *Journal of Bacteriology* **190**, 179–192 (2008).
137. Snyder, C. E. & Benzinger, R. H. Second-step transfer of bacteriophage T5 DNA: purification and characterization of the T5 gene A2 protein. *J Virol* **40**, 248–257 (1981).
138. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
139. Volkov, V. V. *et al.* Uniqueness of ab initio shape determination in small-angle scattering. (2003). doi:10.1107/S0021889803000268