



**HAL**  
open science

# Etude fonctionnelle et structurale de la protéine d'immunité Llp du phage T5

Séraphine Crassac Degroux

► **To cite this version:**

Séraphine Crassac Degroux. Etude fonctionnelle et structurale de la protéine d'immunité Llp du phage T5. Biologie structurale [q-bio.BM]. Université Grenoble Alpes [2020-..], 2022. Français. NNT : 2022GRALV073 . tel-04166889

**HAL Id: tel-04166889**

**<https://theses.hal.science/tel-04166889>**

Submitted on 20 Jul 2023

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

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

THÈSE

Pour obtenir le grade de

**DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES**

École doctorale : CSV- Chimie et Sciences du Vivant

Spécialité : Biologie Structurale et Nanobiologie

Unité de recherche : Institut de Biologie Structurale

**Etude fonctionnelle et structurale de la protéine d'immunité Llp du phage T5**

**Functional and structural study of Llp, the immunity protein of Phage T5**

Présentée par :

**Séraphine CRASSAC DEGROUX**

Direction de thèse :

**Cécile BREYTON**  
COMITE DE THESE CBS2, CNRS

Directrice de thèse

Rapporteurs :

**Pascale BOULANGER**  
DIRECTRICE DE RECHERCHE, CNRS délégation Ile-de-France Sud  
**Sophie ZINN-JUSTIN**  
INGENIEUR HDR, CEA centre de Paris-Saclay

Thèse soutenue publiquement le **29 novembre 2022**, devant le jury composé de :

<b>Cécile BREYTON</b> DIRECTRICE DE RECHERCHE, CNRS délégation Alpes	Directrice de thèse
<b>Pascale BOULANGER</b> DIRECTRICE DE RECHERCHE, CNRS délégation Ile-de-France Sud	Rapporteuse
<b>Sophie ZINN-JUSTIN</b> INGENIEUR HDR, CEA centre de Paris-Saclay	Rapporteuse
<b>Christelle BRETON</b> PROFESSEUR DES UNIVERSITES, Université Grenoble Alpes	Présidente
<b>Adeline GOULET</b> CHARGE DE RECHERCHE, CNRS délégation Provence et Corse	Examinatrice
<b>Eva PEBAY-PEYROULA</b> PROFESSEUR EMERITE, Université Grenoble Alpes	Examinatrice







*A Marie-Josèphe MICHAUD, Mam dans nos cœurs.*





# Remerciement

Une thèse n'est pas une longue course solitaire mais un chemin pour grandir rempli de rencontres et de guides.

J'aimerais remercier chaleureusement Cécile Breyton pour m'avoir fait confiance et m'avoir soutenue d'abord en master puis en thèse pendant ces 4 années qui sont passées très vite finalement. Je la remercie pour la formation reçue, la liberté qu'elle m'a donnée tout en me guidant, les discussions scientifiques autour de mes résultats et sa bienveillance au quotidien. Travailler dans son équipe a été un réel plaisir et m'a fait découvrir la recherche publique que j'ai trouvé passionnante.

Je remercie le groupe MSP pour m'avoir accueilli chaleureusement, en particulier Franck Fieschi et Christine Ebel qui m'a encadrée pour mon M1.

Je remercie Pascale Boulanger et Sophie Zinn pour avoir accepté d'être mes rapportrices et d'évaluer ce travail. Je remercie également Adeline Goulet, Eva Pebay-Peyroula, et Christelle Breton d'avoir accepté d'être jury de ma soutenance.

Je voudrais remercier Hugues Nury, Paulo Tavares et Philippe Mas pour leurs soutiens, leurs écoutes et leurs propositions au cours de mes CSI.

Merci à Loïc Salmon et Emeline Mestdach pour les enregistrements de RMN et de m'avoir appris quelques trucs au passage. Un merci particulier à Torsten Herrmann pour la résolution de la structure de Llp par RMN.

J'aimerais également remercier Corinne Deniaud pour son aide précieuse pour la biologie moléculaire, et l'étude des interactions. Un remerciement également pour les nombreuses relectures très efficaces de ma thèse. Merci également pour ta chaleur et ton soutien dans les moments durs de la thèse.

Merci à Aline Le Roy pour tous les moments passés ensemble. Les conseils pratiques au labo, les dépannages de la Biorad qui n'en fait qu'à sa tête et les astuces de technicienne qui sont toujours très précieuses au labo. Merci également pour les pauses café et le soutien quand je n'avais pas le moral. Et les manips de SEC-MALS et d'AUC.

Je remercie Michel Thépaut et Claudine Darnault pour les purifications et les manipulations sur le complexe FhuA-Llp.

Merci à mes stagiaires Perrine Guenneau, Joanna Gottret, Hêvi Sahin et Edine Hamouche qui ont tous fait de leur mieux pour travailler sur les sujets que je leur ai donnés et qui ont tous contribué à un petit bout de ce travail.

Merci à Quentin Durieux, Clara Delaunay, Perrine Rochas et Ilham Bouchemal pour tous les moments de rigolade ou d'entraide ou de soutien mutuel qui a permis d'avoir une super ambiance de travail. Mais aussi pour les repas et le temps d'un café bien mérité après un dur labeur. Merci aussi à Aimé Berwa, Victor Simon, et Adèle Renier qui ont rejoint notre bureau au cours de ma rédaction et qui ont toujours accepté de me donner leurs avis sur des tournures de phrases un peu compliquées.

Je tiens également à remercier chaleureusement tout le personnel support de l'IBS qui rend ce lieu de travail confortable. Sylvie Gusella et Sindy Pesavento pour la laverie et l'autoclave, Alain Grunberg pour son accueil matinal et les pains au chocolat, Nathalie Cardesie pour avoir pris soin de ma santé et pour ses conseils de sécurité, Auriane Denis-Meyere pro de la qualité et des cahiers de labo, Cynthia David et Olivia Ravet qui ont toujours répondu à toutes les questions de ressources humaines, Isabelle Astiere fée de la propreté, Alain Roux qui m'a sauvé une ou deux manip en me dépannant du matériel, et tous les autres que je n'ai pas nommés qui travaillent pour que l'IBS tourne sans accros.

Je remercie également toutes les personnes qui ont corrigé l'orthographe de cette thèse avec une mention spéciale pour ma maman qui a corrigé le matériel et méthode, mais aussi Martine Degroux qui a été d'une grande efficacité, Aline Le Roy et Clara Delaunay.

Enfin je veux remercier ma famille et mes amis. En particulier mon mari Rémy Degroux qui m'a encouragé à faire cette thèse et soutenue durant ces 4 ans, mes deux garçons Augustin et Célestin qui me donnent leur amour tous les jours et même parfois un peu la nuit, et qui m'ont apporté un équilibre de vie important pour survivre à une thèse. A Dad qui m'a donnée quelques articles à lire pour pas que je m'ennuie pendant ma thèse. A mon père qui m'a fait si forte. A mes amis qui me comprennent plus ou moins mais qui sont toujours là années après années Lucie, Marine & Emma & toute la petite famille, Pierre & Estelle, Anne-Céline & Florent, Alïson & Adrien, Florent & Virginie, Célestin & Lucile, ... Enfin, merci à tous ceux qui ont été présents durant ces 4 années intenses et à tous ceux que je n'ai pas cités mais qui sont dans mon cœur.

# Table des matières

## Table des matières

Remerciement.....	v
Table des matières .....	vii
Liste des abréviations.....	xiii
Liste des figures.....	xv
Liste des tableaux.....	xix
1 Introduction.....	21
1.1 Les bactériophages.....	22
1.1.1 Historique : de 1915 à nos jours.....	22
1.1.1.1 Découverte des bactériophages.....	22
1.1.1.2 Bactériophages et biologie moléculaire .....	23
1.1.1.3 L'étude des bactériophages aujourd'hui.....	24
1.1.2 Classification, morphologie et structure .....	25
1.1.3 Les applications biotechnologiques liées aux phages .....	27
1.1.3.1 Phagothérapie médicale & agro-alimentaire .....	27
1.1.3.1.1 L'antibiorésistance .....	27
1.1.3.1.2 Les solutions contre l'antibiorésistance.....	27
1.1.3.2 Applications technologiques & médicales :.....	29
1.1.4 Relations phage - bactérie et cycle de vie des phages .....	30
1.1.4.1 Les cycles de vie des bactériophages .....	30
1.1.4.2 Relation phage - bactérie, la course à l'armement .....	33
1.1.5 Phénomène d'exclusion de surinfection .....	36
1.1.6 Le phage T5.....	39
1.1.6.1 Génome .....	39
1.1.6.2 Aspect structuraux .....	42
1.1.6.3 Cycle d'infection .....	44
1.2 Présentation des protéines .....	46
1.2.1 La protéine FhuA .....	46
1.2.2 La protéine pb5 .....	48
1.2.3 Le complexe FhuA-pb5 .....	49
1.2.4 La protéine Llp .....	50
2 Matériels et méthodes .....	53
2.1 Biologie moléculaire.....	54

2.1.1	Clonage des différentes formes de Llp .....	54
2.1.2	Mutagenèse dirigée : .....	57
2.1.3	Cellule compétente pour la biologie moléculaire : .....	60
2.2	Tests de surexpression : .....	60
2.3	Production et purification des protéines .....	60
2.3.1	Llp acylée .....	60
2.3.1.1	Surexpression .....	60
2.3.1.2	Solubilisation .....	61
2.3.1.3	Purification .....	61
2.3.2	Llp fusionnée a la MBP .....	62
2.3.2.1	Surexpression .....	62
2.3.2.2	Marquage isotopique .....	62
2.3.2.3	Purification .....	62
2.3.3	FhuA.....	63
2.3.3.1	Surexpression .....	63
2.3.3.2	Solubilisation différentielle des membranes.....	64
2.3.3.3	Purification de FhuA .....	64
2.3.3.4	Échange de détergent .....	64
2.3.3.5	FhuA-L10-Llp.....	64
2.3.4	Production de protéines utilitaires.....	65
2.3.4.1	TEV.....	65
2.3.4.2	MSP.....	65
2.4	Caractérisations biochimiques .....	66
2.4.1	Détermination de la concentration de protéine: .....	66
2.4.2	SDS-PAGE.....	66
2.4.3	SEC-MALS.....	67
2.4.4	Dichroïsme circulaire.....	67
2.4.5	Température de dénaturation par la nano-DSF .....	68
2.5	Études fonctionnelles et interactions.....	68
2.5.1	Test de fonctionnalité.....	68
2.5.1.1	Test de sensibilité in vivo.....	68
2.5.1.2	Test d'absorption de phage.....	69
2.5.1.3	Test d'interaction fonctionnelle FhuA-Ac-Llp in vitro .....	69
2.5.2	Mesure de l'interaction entre FhuA et Llp .....	69
2.5.2.1	SPR et BLI .....	69
2.5.2.2	MST (MicroScale Thermophoresis) : .....	71

2.5.2.3	AUC.....	72
2.5.2.4	Fidabio.....	72
2.6	Études structurales.....	73
2.6.1	RMN.....	73
2.6.1.1	Tests préliminaires et attribution.....	73
2.6.1.2	Information structurale 3D.....	75
2.6.1.3	Détermination de structure par RMN.....	75
2.6.1.4	Dynamique de Llp : relaxation de spin.....	76
2.6.1.5	Exploration de la surface d'interaction entre Llp et FhuA.....	78
2.6.2	Microscopie électronique.....	79
2.6.3	Cristallographie.....	80
3	Contexte & Objectifs de thèse.....	81
	Résultats.....	83
4	Chapitre I : Structure du complexe FhuA-pb5.....	85
4.1	Introduction.....	86
4.2	Formation du complexe FhuA-pb5.....	87
4.2.1	Production de FhuA.....	87
4.2.2	Production de pb5.....	88
4.2.3	Formation du complexe FhuA-pb5.....	89
4.3	Structure du complexe FhuA-pb5 : Article.....	90
5	Chapitre II : Structure de la protéine Sol-Llp.....	111
5.1	Introduction.....	112
5.2	Mise au point de la production et purification de Sol-Llp.....	112
5.2.1	Construction du plasmide pour la surexpression de Llp sous forme soluble.....	112
5.2.2	Production de la protéine Sol-Llp.....	113
5.2.2.1	Culture bactérienne, expression et localisation.....	113
5.2.2.2	Mise au point de la purification.....	116
5.3	Caractérisation biochimique de Sol-Llp et validation pour la RMN.....	118
5.3.1	Vérification de l'intégrité des protéines par spectrométrie de masse.....	118
5.3.2	Stabilité thermique.....	119
5.3.3	Comparaison de structure de Ac-Llp et Sol-Llp par dichroïsme circulaire et RMN.....	121
5.4	Résolution de la structure de Sol-Llp.....	123
5.4.1	Détermination des conditions de mesure :.....	123
5.4.1.1	Spectre 2D HSQC.....	123
5.4.1.2	Effet de la température.....	124
5.4.1.3	Stabilité de la protéine.....	125

5.4.2	Attribution du squelette carboné.....	126
5.4.3	La structure de Sol-Llp.....	128
5.5	Discussion.....	129
6	Chapitre III : Etude du complexe FhuA-Llp.....	133
6.1	Introduction.....	134
6.2	Mise au point de la production et purification de Ac-Llp.....	134
6.2.1	Construction des plasmides de surexpression de Ac-Llp.....	134
6.2.2	Production de la protéine Ac-Llp.....	135
6.2.2.1	Culture bactérienne, expression et solubilisation.....	135
6.2.2.2	Mise au point de la purification.....	139
6.3	Caractérisation biochimique de Ac-Llp.....	143
6.3.1	Spectrométrie de masse.....	143
6.3.2	Stabilité thermique.....	144
6.3.3	Étude de l'état d'oligomérisation par SEC-MALS.....	145
6.4	Interactions FhuA et Llp.....	147
6.4.1	Interactions <i>in vivo</i> .....	147
6.4.2	Interaction <i>in vitro</i> .....	149
6.4.2.1	SEC-MALS.....	149
6.4.2.2	Nano-DSF.....	150
6.4.2.3	Protection du phage contre FhuA par Llp, détection de l'éjection d'ADN.....	151
6.4.2.4	Protection du phage contre FhuA par Llp, détection par inactivation du phage....	153
6.4.3	Détermination du Kd de l'interaction entre FhuA et Ac-Llp.....	154
6.4.3.1	Résonance de plasmonique de Surface, SPR.....	154
6.4.3.2	Micro Scale Thermophoresis MST.....	159
6.4.3.3	AUC.....	160
6.4.3.4	Autres méthodes testées.....	161
6.4.3.5	Synthèse des différentes méthodes utilisée :.....	162
6.4.4	Exploration de la surface d'interaction côté FhuA.....	162
6.4.4.1	Mutants de FhuA.....	162
6.4.4.2	Variations naturelles de FhuA dans des souches de Shigella et de Salmonella.....	167
6.4.5	Exploration de la surface d'interaction coté Llp.....	168
6.4.5.1	RMN.....	168
6.4.5.2	Exploration de la surface d'interaction de Sol-Llp par mutagenèse.....	173
6.5	Études structurales du complexe FhuA-Llp.....	175
6.5.1	Cristallogénèse.....	175
6.5.1.1	Macro cristaux.....	175

6.5.1.2	Microcristaux.....	175
6.5.2	Formation de nanodisques contenant FhuA-Ac-Llp pour la microscopie électronique 176	
6.5.3	Formation de fusion FhuA-Llp pour la résolution de structure.....	177
6.6	Discussion .....	178
7	Conclusion et perspective .....	181
	Bibliographies.....	185
	ANNEXES.....	195
	Annexe 1 : Fiches plasmides.....	196
	Annexe 2 : Table des entrées pour le calcul de structure et la caractérisation de la structure RMN minimisée en énergie pour Sol-Llp.....	206
	Annexe 3 : vérification de la présence de l'insert dans les clonages. ....	208
	Annexe 4 : Test d'interaction BLI et Fidabio.....	209
	Résumés .....	212
	Abstract .....	212





# Liste des abréviations

- Ac-Llp** : Llp sous forme acylé
- Acrs** : anti-CRISPR proteins
- ADN / sb ou db** : Acide DésoxyriboNucléique / simple brin ou double brin
- AHT** : anhydrotetracycline
- ANSM** : Agence Nationale de Sécurité du Médicament et des produits de santé
- ARN / sb ou db** : Acide RiboNucléique / simple brin ou double brin
- AUC** : Analytical UltraCentrifugation
- BCA** : Bicinchoninic Acid
- BHP** : Base plate Hub Protein
- BLI** : BioLayer Interferometry
- CD** : Circular Dichroism
- CFP** : Central Fibre Protein
- CMC** : concentration micellaire critique
- CRISPR-Cas**: Clustered Regularly Interspaced Short Palindromic Repeats - CRISPR Associated
- Cryo-EM** : Cryo Electronique Microscopie
- DDM** : Dodécyl-β-D-maltoside
- DNase** : Deoxyribonuclease
- DSF (nano-DSF)** : differential scanning fluorimetry
- DTP** : Distal Tail Protein
- EDC** : N-ethyl-N'-(dimethylaminopropyl) carbodiimide
- EDTA** : EthyleneDiamineTetraAcetic acid
- ESRF** : European Synchrotron Radiation Facility
- Fc** : Flow cell
- FhuA** : Ferric hydroxamate uptake protein A
- FIDA** : Flow Induced Dispersion Analysis
- FITC** : Isothiocyanate de fluorescéine
- FNIII** : domains fibronectins
- FST** : first step transfer
- Gmr** : glucose-modified restriction
- HMC / glu-HMC** : hydroxymethylcytosine / glucose- hydroxymethylcytosine
- hsr** : host receptor specificity
- HSQC** : Heteronuclear Single Quantum Coherence
- ICTV** : International Committee on Taxonomy of Viruses
- IMAC** : immobilized metal affinity chromatography
- IPI** : Internal Protein
- IPTG** : isopropyl β-D-1-thiogalactopyranoside
- ISBG** : Integrated Structural Biology Grenoble
- K<sub>d</sub>** : constante de dissociation
- LB** : Lysogeny Broth
- LDAO** : Lauryldimethylamine N-oxide
- LED** : light-emitting diode
- Lgt** : phosphatidylglycerol-prolipoprotein diacylglyceryl transferase
- Llp** : Lytic conversion lipoprotein
- LMNG** : Lauryl Maltose Neopentyl Glycol
- Lnt** : apolipoprotein n-acyltransferase
- Lsp** : lipoprotein signal peptidase
- LPS** : Lipo-PolySacharide
- MBP**: Maltose Binding Protein
- MBP-Llp** : protéine de fusion MNB et Llp
- MCS** : Multiple Cloning Site
- MD** : Molecular dynamics
- MDS** : modification-dependent systems
- MSP** : Membrane Scaffold Protein
- MST** : Microscale thermophoresis
- NHS** : N-hydroxysuccinimide
- Ni-NTA** : Nickel - acide nitrilotriacétique
- NOE** : Nuclear Overhauser Effect
- NOESY** : Nuclear Overhauser Effect Spectroscopy

**oad** : o-antigène dépendant

**OG** : Octyl-beta-Glucoside Detergent

**OMPs** : outer membrane protein

**OMVs** : outer membrane vesicles / vésicules de membranes externes

**OPOE** : Octyl polyoxyethylene

**ORF** : Open Reading Frame

**OTG** : Octyl Thio Glucoside

**PCR** : Polymerase Chain Reaction

**RBP** : Recepteur Binding Protein

**RF** : radiofréquence

**R/M** : restriction–modification system

**RMN** : Résonance Magnétique Nucléaire

**Rmsd** : Root-Mean-Square Deviation

**Rpm** : Rotation par minute

**SANS** : Small Angle Neutron Scattering

**SDS** : Sodium Dodecyl Sulfate

**SEC-MALS** : size exclusion chromatography - Multi-Angle Light Scattering

**Sie** : SuperInfection exclusion protein

**Soc** : Super Optimal broth with Catabolite repression

**SOFAST** : selective optimized flip angle short transient

**Sol-Llp** : Llp sous forme soluble

**SPR** : Surface Plasmon Resonance

**PBS** : phosphate-buffered saline

**PCR** : Polymerase chain reaction

**TAE** : Tris Acetate EDTA buffer

**TB** : Terrific Broth

**TBDTs** : transport actif TonB dépendant

**TBS** : Tris-Buffered Saline

**TCP** : Tail Completion Protein

**TEMED** : Tetramethylethylenediamine

**TEV** : Tobacco Etch Virus Protease

**Tm** : Thermal melting point

**TrP** : protéine de terminaison

**TTC** : Tail Tip Complex

**TTP** : Tail Tube Protein

**TMP** : Tail Measure Protein / protéine vernier

**UV** : ultraviolets

**VHH** : camelid heavy-chain variable domains

# Liste des figures

Figure 1-1 : Chronologie des découvertes des mécanismes moléculaires de la vie.....	23
Figure 1-2 : Classification des bactériophages.....	26
Figure 1-3 : Cycle d'infection possible pour les bactériophages.....	32
Figure 1-4 : Cycle d'infection des bactériophages et mécanismes de résistance des bactéries.....	33
Figure 1-5 : différents mécanismes de prévention d'absorption de phage sur leur récepteur.....	34
Figure 1-6 : La course aux armements entre les bactéries et les phages.....	35
Figure 1-7 : Blocage de l'injection de l'ADN dans la bactérie par le système Sie du coliphage T4. ...	37
Figure 1-8 : Avantages et inconvénients liés au Sie en fonction des mécanismes évolutifs de compétition ou de coopération.....	38
Figure 1-9 : Structure du génome du bactériophage T5.....	41
Figure 1-10 : Représentation schématique des siphophages.....	43
Figure 1-11 : Disposition générale des gènes de queue de siphophages.....	43
Figure 1-12 : Différentes étapes de l'infection par le phage T5.....	44
Figure 1-13 : Mécanismes proposés pour le déclenchement de l'infection, l'ancrage à la membrane et la formation de canaux par le phage T5.....	45
Figure 1-14 : Structure de FhuA en complexe avec le fer ferrichrome.....	47
Figure 1-15 : Structure des enveloppes de pb5 à partir des données SANS.....	48
Figure 1-16 : structure basse résolution de FhuA, pb5 et FhuA-pb5 en coloration négative.....	49
Figure 1-17 : Expression des protéines du phage T5 pendant l'infection.....	50
Figure 1-18 : Mécanisme de transport de Llp par la bactérie.....	51
Figure 1-19 : Structure générale des domaines des lipoprotéines (haut) et séquence du précurseur de la protéine Llp (bas).....	52
Figure 4-1 : Étapes de solubilisation différentielle de FhuA sur gel d'acrylamide 10%, coloré au bleu de Coomassie.....	87
Figure 4-2 : comparaison des profile d'éluion de FhuA sur une colonne échangeuse d'anions avec ou sans delipidation.....	88
Figure 4-3 : Chromatogramme NiNTA de pb5.....	88
Figure 4-4 : A : Gel SDS (10 % d'acrylamide) de l'éluion NiNTA de pb5 à faible pH et EDTA puis avec du NaCl (B) colorée au bleu de Coomassie.....	89
Figure 4-5 : FhuA-pb5, FhuA et pb5 sur gel SDS 10 % coloré avec du bleu de Coomassie.....	89
Figure 5-1 : courbe de croissance bactérienne de souche BL21 et Shuffle transformées avec les plasmides Sol-Llp ou MBP-Llp.....	113
Figure 5-2 : Tests d'expression de la protéinée Sol-Llp (A) ou MalE-Llp (B) dans les souches BL21(DE3).....	114
Figure 5-3 : test phénotypique par sensibilité au phage T5 de la souche BL21 contenant ou non le plasmide d'expression codant pour Sol-Llp.....	115
Figure 5-4 : Vérification de la localisation subcellulaire de la protéine MBP-Llp.....	115
Figure 5-5 : Résultat de la première purification de MBP-Llp.....	116
Figure 5-6 : comparaison de la purification de Sol-Llp avant et après optimisation.....	117
Figure 5-7 : spectrométrie de masse sur la protéine Sol-Llp après enregistrement des spectres RMN.....	118
Figure 5-8 : courbe de suivi de dénaturation de la protéine Sol-Llp par le nano-DSF.....	119
Figure 5-9 : courbe de suivi de dénaturation de la protéine Sol-Llp à différent pH.....	120
Figure 5-10 : Comparaison des spectres de dichroïsme circulaire de Ac-Llp ou Sol-Llp.....	121

Figure 5-11 : TROSY-methyl-HSQC sur les méthyles de Sol-Llp (rouge) et de Ac-Llp en OG (bleu).	122
Figure 5-12 : Spectre <sup>1</sup> H, <sup>15</sup> N-HSQC de Sol-Llp.....	123
Figure 5-13 : Investigation de l'effet de température sur les spectres <sup>1</sup> H, <sup>15</sup> N-HSQC de Sol-Llp.....	124
Figure 5-14 : Spectres <sup>1</sup> H, <sup>15</sup> N-HSQC de Sol-Llp au cours du temps acquis à 303K à 600 MHz. ....	125
Figure 5-15: Noms des atomes de carbone sur la structure d'un acide aminé général. ....	126
Figure 5-16 : Stratégie d'attribution de la protéine Sol-Llp pour les résidus S50 à N56. ....	126
Figure 5-17 : Spectre HSQC de Sol-Llp avec l'attribution. ....	127
Figure 5-18 : structure RMN moyenne de Sol-Llp.....	128
Figure 5-19 : structure RMN des 20 conformères les plus probables. ....	129
Figure 5-20 : Alignement des protéines Cor et Llp-like avec Llp de T5. ....	130
Figure 5-21 : Superposition des prédictions de structure des protéine Cor et Llp-like par rapport à la structure de Llp obtenue en RMN.....	131
Figure 5-22 : Comparaisons structurales entre Sol-Llp et des lipoprotéines aux repliments similaires. .....	132
Figure 6-1 : courbe de croissance bactérienne de souches BL21(DE3) et C43(DE3) transformée avec le plasmide codant pour Ac-Llp.....	135
Figure 6-2 : Test d'expression de la protéine Ac-Llp dans les souches BL21(DE3) et C43(DE3). ....	136
Figure 6-3 : Test de solubilisation de la protéine Ac-Llp suivi par électrophorèse sur gel.....	137
Figure 6-4 : Séparation des membranes de C43(DE3), transformées ou non, 37°C ou 20 °C, en présence ou en absence d'ITPG sur gradient de saccharose et analyse sur gel. ....	138
Figure 6-5 : Solubilisation des membranes issues d'une culture à 20°C avec ou sans induction ....	138
Figure 6-6 : Vérification de la purification de la protéine Ac-Llp sur colonne d'affinité au Nickel, éluée à 250 mM Imidazol. ....	139
Figure 6-7 : Vérification de la purification de la protéine Ac-Llp sur colonne échangeuse d'ions... ..	140
Figure 6-8 : Vérification de l'état d'oligomérisation de la protéine Ac-Llp. ....	141
Figure 6-9 : Chromatogramme d'éluion de la protéine Ac-Llp sur colonne d'affinité au nickel.....	142
Figure 6-10 : Contrôle de la pureté de Ac-Llp après purification sur une colonne d'affinité au nickel avec une éluion à l'EDTA.....	142
Figure 6-11 : Caractérisation par spectrométrie de masse de la protéine Ac-Llp après purification. .....	143
Figure 6-12 : Courbe de suivi de dénaturation de la protéine Ac-Llp avec 5% ou sans β-mercaptoéthanol par nano-DSF. ....	144
Figure 6-13 : Résultats du SEC-MALS sur la protéine Ac-Llp dans différents détergents.....	146
Figure 6-14 : Test phénotypique de sensibilité au phage T5 des souches F, BL21(DE3) et C43(DE3). .....	147
Figure 6-15 : Test phénotypique de sensibilité au phage T5 de la souche C43 contenant ou non le plasmide d'expression codant pour AtpF. ....	148
Figure 6-16 : test phénotypique par sensibilité au phage T5 des souches BL21 (DE3) et C43 (DE3) contenant ou non le plasmide d'expression codant pour Ac-Llp.....	148
Figure 6-17 : Expérience de SEC-MALS réalisées sur des mélanges FhuA-Ac-Llp et FhuA-Sol-Llp à un ratio 1/8. ....	149
Figure 6-18 : Courbe de suivi de dénaturation de la protéine Ac-Llp, FhuA et du complexe FhuA : Ac-Llp. ....	150
Figure 6-19 : Courbe de suivi de dénaturation de la protéine Ac-Llp, FhuA et du complexe FhuA : Ac-Llp-expérimental ou calculée. ....	150
Figure 6-20 : Principe du test d'interaction in vitro entre FhuA et Ac-Llp avec T5 comme rapporteur de l'interaction.....	151

<b>Figure 6-21 : Courbe d'étalonnage de l'intensité de Fluorescence en fonction de la quantité d'ADN.</b>	151
<b>Figure 6-22 : Suivi de l'éjection de l'ADN du phage T5 en présence de FhuA avec ou sans Ac-Llp.</b>	152
<b>Figure 6-23 : Principe du test in vitro d'interaction entre FhuA et Ac-Llp avec T5 comme reporteur de l'interaction.</b>	153
<b>Figure 6-24 : Pourcentage de phages T5 inactivés en fonction de la concentration en FhuA.</b>	153
<b>Figure 6-25 : Test fonctionnel d'interaction entre FhuA et Ac-Llp contrôlé par la protection du phage T5 contre l'inactivation par FhuA.</b>	154
<b>Figure 6-26 : Interaction du LMNG ou du DMM avec FhuA, Ac-Llp ou Sol-Llp immobilisées.</b>	155
<b>Figure 6-27 : Interaction de Ac-Llp avec FhuA immobilisée, en DDM ou en LMNG.</b>	156
<b>Figure 6-28 : Interaction de Ac-Llp avec FhuA immobilisée, en DDM.</b>	158
<b>Figure 6-29 : Structure Sol-Llp.</b>	158
<b>Figure 6-30 : Courbe d'affinité de Sol-Llp pour FhuA.</b>	159
<b>Figure 6-31 : Courbes d'affinité de Ac-Llp et du DDM pour FhuA.</b>	160
<b>Figure 6-32 : Coefficient de sédimentation de Ac-Llp, Sol-Llp, FhuA et de complexe FhuA : Llp à différents ratios molaires.</b>	161
<b>Figure 6-33 : Localisation des mutants de FhuA étudiés.</b>	163
<b>Figure 6-34 : Vérification de l'expression des gènes de FhuA mutés dans la souche AW740 en absence de bipyridyl.</b>	163
<b>Figure 6-35 : Comparaison des profils d'élution des mutants de FhuA sur NiNTA, purifiés sur la plateforme M3P de l'IBS.</b>	165
<b>Figure 6-36 : Courbe de suivi de dénaturation de FhuA mutant et des complexes FhuA : Ac-Llp.</b>	166
<b>Figure 6-37 : Titration du phage T5 sur les souches Salmonella enterica indica en absence ou en présence de Ac-Llp.</b>	168
<b>Figure 6-38 : 1H,15N-SOFAST-HSQC sur Llp.</b>	169
<b>Figure 6-39 : Investigation de l'influence du DDM (a &amp; b) ou de FhuA (c &amp; d) sur Sol-Llp par l'analyse des pics du spectre 1H, 15N-SOFAST-HSQC.</b>	169
<b>Figure 6-40: 1H,15N-SOFAST-HSQC de Sol-Llp.</b>	171
<b>Figure 6-41 : Variations de déplacements chimiques et d'intensités reportées sur la structure de Llp.</b>	172
<b>Figure 6-42 : Vérification de l'expression des mutants de Ac-Llp dans la souche C43.</b>	173
<b>Figure 6-43 : Influence des mutations de Ac-Llp sur la protection des souches C43 portant les plasmides des mutants de Llp face au phage T5.</b>	174
<b>Figure 6-44 : Synthèse des acides aminés pouvant être impliqués dans l'interaction avec FhuA sur la structure de Sol-Llp.</b>	174
<b>Figure 6-45 : Exemple de cristaux obtenus sur la plateforme du HTX-lab.</b>	175
<b>Figure 6-46 : Exemple de cristaux obtenus à la main.</b>	176
<b>Figure 6-47 : Observation des microcristaux au microscope électronique : Cristaux du complexe FhuA : Ac-Llp en DDM avec un ratio 1 : 4 à une concentration de 8 mg. A gauche l'image d'un microcristal observé au microscope électronique à droite diffraction électronique de ce cristal.</b>	176
<b>Figure 6-48 : Suivi de la purification des nanodisques contenant FhuA et Ac-Llp sur colonne de NiNTA.</b>	176
<b>Figure 6-49 : Vérification de la formation des nanodisques de FhuA : Llp ratio 1 : 1 par microscopie électronique.</b>	177
<b>Figure 6-50 : Structure du complexe FhuA-Llp.</b>	180
<b>Figure 7-0-1 : Digestion enzymatique par BamHI des constructions Sol-Llp (A) et MalE-Llp (B).</b>	208
<b>Figure 7-2: Digestion enzymatique par BamHI des constructions Ac-Llp (WT), et PuIA-Ac-Llp (CDNS).</b>	208

<b>Figure 7-0-3: Sensorgramme de l'interaction entre Ac-Llp et FhuA immobilisé, en DDM.....</b>	<b>209</b>
<b>Figure 7-0-4: : Courbe d'affinité de Ac-Llp pour FhuA par Fidabio. ....</b>	<b>210</b>
<b>Figure 7-5: Courbe d'affinité de Ac-Llp pour FhuA. ....</b>	<b>210</b>

# Liste des tableaux

Tableau 1-1 : Exemple de virus possédant un système d'exclusion de surinfection.....	36
Tableau 2-1 : Cycles de PCR1 utilisés pour amplifier les séquences du gène <i>lfp</i> .....	54
Tableau 2-2 : Cycle de PCR2 utilisé pour insérer les gènes dans le plasmide de sur expression.....	54
Tableau 2-3 : Amorces utilisées pour le clonage des différentes constructions.....	56
Tableau 2-4 : Cycle de PCR utilisé pour insérer les codons mutants.....	57
Tableau 2-5 : Amorces utilisées pour les mutations de <i>FhuA</i> .....	58
Tableau 2-6 : Amorces utilisées pour les mutations de <i>FhuA</i> .....	59
Tableau 2-7 : Conditions d'expression testées.....	60
Tableau 2-8 : Protéine et solvants testées en SEC-MALS.....	67
Tableau 2-9 : Gammes de concentration des protéines en flux pour la SPR et le BLI.....	70
Tableau 2-10 : Liste des échantillons testés en AUC.....	72
Tableau 2-11 : Titration <i>Lfp-FhuA</i> .....	79
Tableau 5-1 : Constructions et stratégie pour l'expression de <i>Lfp</i> sous forme soluble.....	112
Tableau 5-2 : Condition des tests d'expression pour la <i>Lfp</i> sous forme soluble.....	114
Tableau 5-3 : atomes de carbone observés sur les spectres pour chaque expérience 3D d'attribution.....	126
Tableau 6-1 : Construction et stratégie pour l'expression de <i>Lfp</i> acylée.....	134
Tableau 6-2 : Récapitulatif des données des mutants de <i>FhuA</i> .....	165
Tableau 6-3 : Concentration des différents partenaires pour chaque conditions mesurées en RMN. .....	170





# 1 Introduction

## 1.1 Les bactériophages

Les bactériophages sont des virus infectant les procaryotes et ce sont des parasites obligatoires<sup>1</sup>. Les phages peuvent être sous deux formes, soit sous forme de virions qui sont capables d'infecter d'autres cellules, soit sous forme de prophages, c'est-à-dire que leur génome est intégré à celui de leur hôte<sup>2</sup>. Les bactériophages sont l'entité biologique la plus abondante sur terre et sont trouvés dans tous les biotopes où des bactéries sont présentes<sup>2</sup>.

### 1.1.1 Historique : de 1915 à nos jours

#### 1.1.1.1 Découverte des bactériophages

En 1915, alors que Frédérick Twort tente de trouver un milieu de culture pour le virus de la vaccine, il observe des tâches vitreuses et transparentes sur les contaminants bactériens. Il émet trois hypothèses : pour lui il peut s'agir, d'un cycle de vie particulier des bactéries, d'un élément produit par les bactéries elles-mêmes, ou bien d'un « virus ultra microscopique »<sup>3</sup>. C'est la première découverte par sérendipité des bactériophages. Deux ans plus tard, Félix D'Hérelle, travaillant sur une épidémie de dysenterie, décrit ce phénomène de lyse bactérienne. Il observe que l'élément responsable persiste dans les solutions filtrées sur un filtre de Chamberland. Cet élément est capable de se transmettre entre les cultures et de s'amplifier. Félix D'Hérelle a l'intime conviction que cet élément est un organisme biologique capable de se reproduire en se « nourrissant » des bactéries, il le baptise « bactériophage ». Ayant observé que des bactériophages ciblant la bactérie pathogène apparaissent dans les selles de malades qui guérissent de la dysenterie, Félix D'Hérelle, conduisit, en 1919, une étude clinique sur des patients atteints de dysenterie, traités avec des bactériophages à l'hôpital des Enfants Malades de Paris : la phagothérapie était née<sup>4</sup>. Encouragé par les résultats positifs de cette étude, Félix D'Hérelle, part en Inde pour y traiter le choléra et la peste. Même si la nature exacte des bactériophages n'est pas connue à cette époque, le traitement par phagothérapie est bien accueilli. En 1923, Georges Eliava ouvrit un Institut consacré à la phagothérapie à Tbilisi en Géorgie, Félix D'Hérelle y travailla d'ailleurs quelques mois en 1934. Cet institut, ainsi que les laboratoires de Félix D'Hérelle à Paris, produisirent des préparations de phages commerciales contre différentes infections<sup>5</sup>. Cependant, après la deuxième guerre mondiale, cette pratique tombe petit à petit dans l'oubli pour deux raisons majeures :

- Les cocktails commerciaux de phages sont peu ou pas efficaces. À cette époque les phénomènes de lysogénie ne sont pas encore connus.
- La découverte des antibiotiques chimiques. En effet, ils peuvent être produits en grande quantité et sont plus simple d'utilisation. Ils sont également moins spécifiques que les phages et un même antibiotique peut être utilisé contre un grand nombre d'infections bactériennes.

La phagothérapie est cependant encore pratiquée dans l'ancien bloc de l'Est, et l'Institut Eliava est encore actif aujourd'hui. En outre, les phages ont été utilisés comme modèle d'étude et sont à la base de grandes découvertes sur la génétique.

### 1.1.1.2 Bactériophages et biologie moléculaire

Max Delbrück est un physicien de la fin des années 30. Il s'intéresse aux mécanismes de la vie, et veut utiliser les outils de la physique pour comprendre la biologie. Dans ce but, il cherche un organisme simple à étudier. Il rencontre alors, Emory Elis, qui étudie la biologie virale fondamentale à l'aide des bactériophages. C'est le point de départ de l'utilisation des phages comme modèle pour l'étude de ce qui deviendra la biologie moléculaire. Il se forme alors un groupe de chercheurs, d'origines diverses, comprenant des scientifiques comme : Salvador Luria, Alfred Hershey, François Jacob, Seymour Benzer, ... Ce groupe surnommé le « Groupe Phage », se réunit régulièrement au Cold Spring Harbor Laboratory pour des écoles d'été. Le maître mot de ce groupe est le partage des connaissances plutôt que la compétition, l'échange des protocoles et la définition de souches de bactéries et de phages communs pour que les manipulations soient reproductibles d'un laboratoire à l'autre. Les phages lytiques de la série T sont utilisés pour l'étude et le fonctionnement moléculaire de l'héritage. La régulation des gènes est étudiée à l'aide des phages tempérés comme le phage lambda<sup>6</sup>. Il s'ensuit 20 ans de recherche et de découvertes sur les mécanismes moléculaires de la vie, décrits dans la (Figure 1-1).

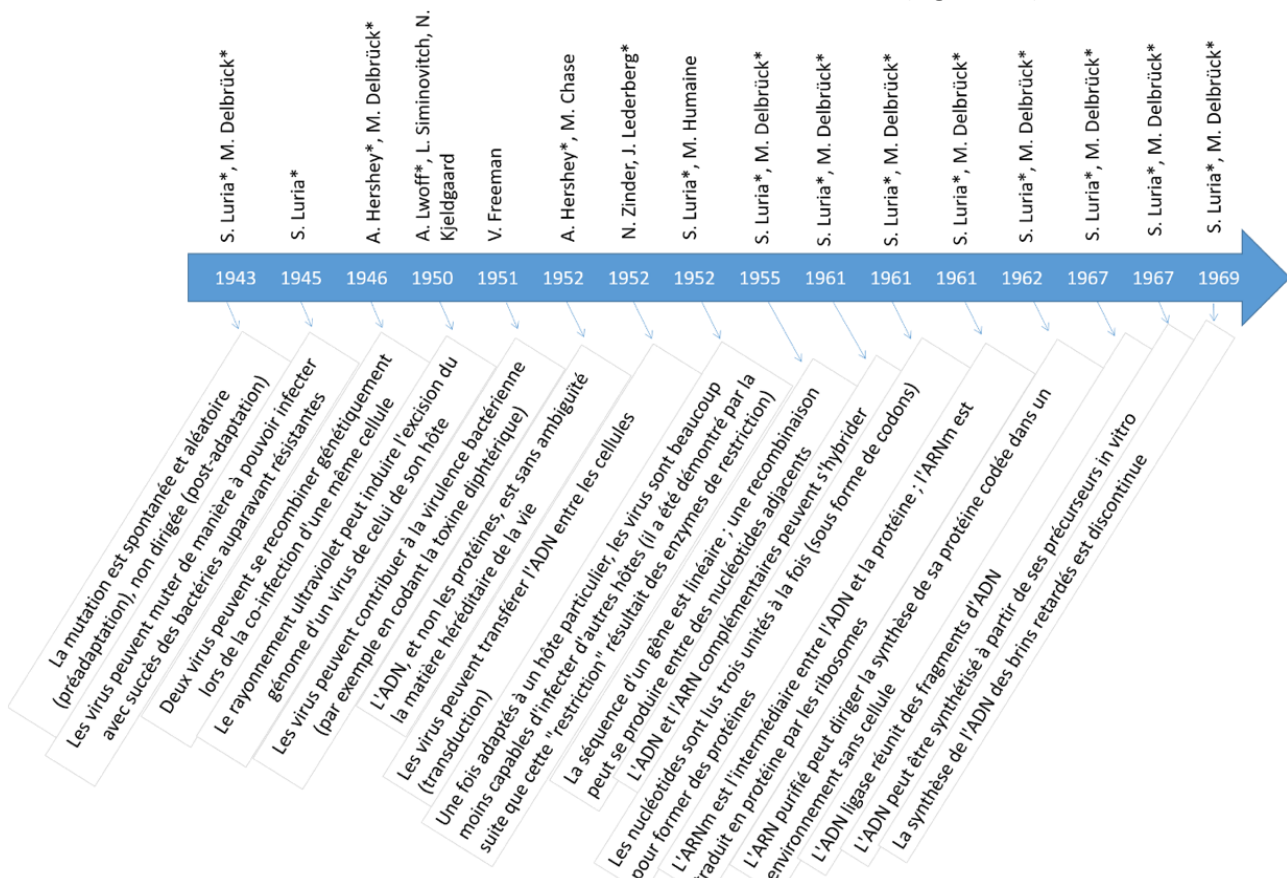


Figure 1-1 : Chronologie des découvertes des mécanismes moléculaires de la vie.

La fresque montre les principales découvertes de biologie moléculaire entre 1943 et 1969. Cette figure est réalisée d'après 6.

En plus de ces grandes découvertes, le « Groupe Phage » a façonné la biologie comme une discipline scientifique rigoureuse. En effet, la biologie était à l'époque l'héritière de ce qu'on appelle aujourd'hui l'histoire naturelle, définie en tant que discipline scientifique seulement en 1802 par Jean-Baptiste Lamarck : « Tout ce qui est généralement commun aux végétaux et aux animaux [...], doit constituer l'unique et vaste objet d'une science particulière qui n'est pas encore fondée, qui n'a même pas de nom, et à laquelle je donnerai le nom de Biologie »<sup>7</sup>. Le Groupe Phage, riche de ses scientifiques multidisciplinaires, a appliqué et communiqué à la communauté des biologistes un ensemble de normes rigoureuses et cohérentes pour la conception et l'interprétation des expériences. La Biologie s'imprègne alors du raisonnement déductif, et de l'évaluation des données obtenues, il en résulte la biologie telle qu'on la connaît aujourd'hui<sup>6</sup>.

De plus, les études sur les phages, ont donné à la biologie moderne des outils puissants pour le génie génétique et la production de protéines pour leur *étude in vitro*. On peut citer par exemple les enzymes de restriction, mais aussi les promoteurs forts, comme le promoteur T7 pour la surexpression des protéines. Plus récemment, la découverte des systèmes de défense des bactéries contre les bactériophages, a permis la conception d'un outil d'édition génétique universel, CRISPR-Cas<sup>8</sup> (Clustered Regularly Interspaced Short Palindromic Repeats - CRISPR Associated).

#### 1.1.1.3 L'étude des bactériophages aujourd'hui

Comme nous l'avons vu dans le paragraphe précédent, l'étude des bactériophages a permis l'essor de la biologie moléculaire au cours du XXe siècle. Le XXIe siècle marque un tournant dans l'étude des phages. En effet, on se rend compte aujourd'hui de l'importance d'étudier les phages d'un point de vue global.

D'un point de vue écologique par exemple, les études passées sur les bactériophages ont été réalisées dans les conditions de laboratoire bien contrôlées. Il apparaît maintenant clairement qu'il faut également étudier les phages dans la complexité du monde réel<sup>9</sup>. Il paraît indispensable de mieux comprendre la relation phages-hôtes dans des milieux complexes pour mieux évaluer leur impact écologique. En outre, on estime qu'il y a  $10^{31}$  phages sur terre. La majorité des phages reste inconnue, ils sont donc une réserve importante de nouveaux gènes et de protéines à découvrir. Étudier les bactériophages est donc essentiel pour comprendre pleinement leur rôle dans les cycles géochimiques de la matière, la régulation des populations bactériennes dans les océans et leur impact sur le climat etc. <sup>9, 10 et 11</sup>. De plus, comme l'histoire l'a montré, ces entités ont un fort potentiel pour les applications biotechnologiques.

D'un point de vue des applications technologiques, l'étude des phages permet de mettre au point des nouvelles technologies. Les méthodes d'édition génétique CRISPR-CAS mentionnées dans le

paragraphe précédent, ont été développées à partir de 2012. D'autres technologies sont en cours de mise au point, certaines seront développées dans le paragraphe 1.3 Applications.

L'études de la relation bactériophage / bactérie est également importante dans l'industrie. En effet plusieurs types d'industries utilisent les bactéries dans leurs procédures de fabrication de produit (pharmaceutique, agroalimentaire, industrie chimique, etc.). Dans ce cas les phages sont vus comme des ennemis car ils affectent la production issue des bactéries<sup>12</sup>. Cependant les bactéries ne sont pas sans défense face aux phages (voir section 1.1.4.2). Le contrôle des systèmes de défense bactérien peut être utilisé pour protéger des bactéries utilisées dans l'industrie.

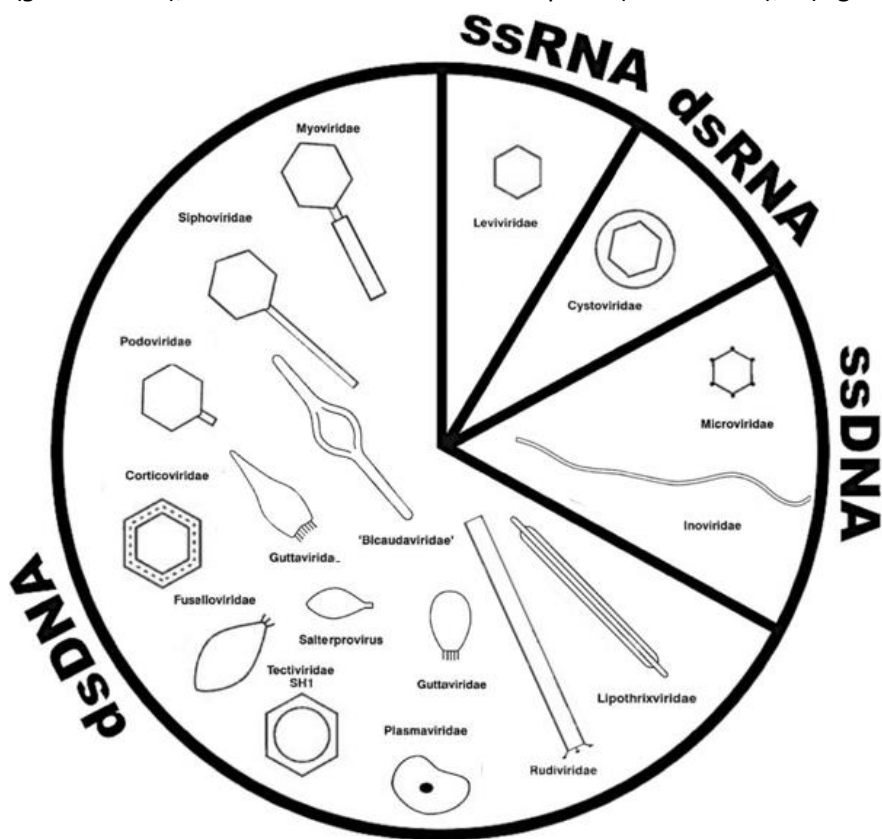
Un autre axe majeur de la recherche sur les bactériophages concerne la phagothérapie. En effet, l'émergence et l'amplification des bactéries résistantes aux antibiotiques, poussent les chercheurs à trouver des alternatives pour combattre les infections bactériennes. Les enjeux et difficultés des thérapies utilisant des phages ou leurs composés seront discutés dans le chapitre 1.3.1 Applications Médicales.

#### 1.1.2 Classification, morphologie et structure

Le comité international de la taxonomie (ICTV) a en charge la classification des virus. La première classification de virus proposée est celle de David Baltimore, qui répartit les virus en six groupes en fonction de leur type d'acide nucléique, mais aussi de leur mode d'expression de l'ARN messager viral. L'ICTV propose une classification complémentaire, en regroupant les virus par ordre, famille, sous famille, genres et espèces, comme cela se fait pour les autres règnes biologiques<sup>13</sup>. En 2019, 6590 espèces sont répertoriées et réparties dans 168 familles et 55 ordres<sup>14</sup>.

La classification des virus est un processus complexe. Le regroupement des espèces en famille, sous famille, ordre... se base sur la comparaison des génomes et leur morphologie observée au microscope électronique. De ce fait il dépend des outils bio-informatiques utilisés pour comparer les génomes comme Pairwise Sequence Comparison. De plus les bactériophages n'ont pas leurs propre ARN ribosomiaux sur lesquels se basent les autres études phylogénétiques. À ce jour, les génomes de bactériophages déposés dans GenBank® n'ont pas révélé de gènes communs aux bactériophages pouvant servir de marqueurs pour l'identification<sup>14</sup>. De plus, l'observation microscopique des phages dont le génome est isolé n'est pas toujours possible, il manque donc de l'information morphologique pour certains bactériophages. La classification des virus est aussi un processus dynamique. Par exemple, entre 2018 et 2019, le sous-comité des virus de procaryotes de l'ICTV, a proposé la création d'un nouvel ordre (*Tubulavirales*) mais aussi de 10 nouvelles familles, 22 sous famille, 424 genres et 964 nouvelles espèces<sup>13</sup>. De plus, chaque année entre 400 et 600 nouveaux génomes de phages sont déposés dans la base de données GenBank®, et doivent donc être classés par le comité de l'ICTV.

Les bactériophages sont des virus procaryotes. Ils sont donc inclus dans la classification des virus. Ils sont répartis en quatre groupes selon la nature de leur acide nucléique : ADN simple brin (sbADN), ADN double brin (dbADN), ANR simple brin (sbARN), ou ARN double brin dbARN). En 2007, une étude par microscopie électronique sur plus de 5500 bactériophages, montre que la majorité d'entre eux, 96%, appartient à l'ordre des *Caudovirales*, c'est-à-dire qui possèdent une capsid et une queue<sup>15,16</sup>. L'ordre des *Caudovirales*, comprend trois familles qui se distinguent par la morphologie de leur queue : 25% sont des *Myoviridae* (longues queues contractiles), 60% des *Siphoviridae* (longues queues non-contractiles) et 15% des *Podoviridae* (petites queues). Cependant, les phages présentent également des formes morphologiques variées : filamenteux (*inoviridae*), en forme de citron (*Fuselloviridae*), en forme de gouttes (*guttaviridae*), ou encore avec une double capsid (*Tectiviridae*),... (Figure 1-2).



**Figure 1-2 : Classification des bactériophages.**

Les bactériophages sont classés en 4 grands groupes : sbADN, sbADN, dbARN, sbARN. La majorité des bactériophages sont des particules à ADN double brins. On constate également une diversité importante de morphologies possibles pour les bactériophages. Cette figure est réalisée d'après<sup>15 & 16</sup>.

Au-delà de leur variété de formes, les bactériophages ont des génomes très variés. Il existe des bactériophages avec un tout petit génome de 3.5 kpb<sup>17</sup>, mais aussi des bactériophages avec des génomes allant jusqu'à 735 kpb<sup>18</sup>. L'ADN génomique peut également comporter des bases atypiques telle que la 5-hydréméthylcytosine qui permet de protéger le génome des enzymes de restriction de l'hôte (e.g. bactériophage T4)<sup>14</sup>. De même, la composition en bases nucléiques (A, T, G, C) varie beaucoup d'un phage à l'autre<sup>19</sup>.

### 1.1.3 Les applications biotechnologiques liées aux phages

#### 1.1.3.1 Phagothérapie médicale & agro-alimentaire

##### 1.1.3.1.1 L'antibiorésistance

L'antibiorésistance des bactéries est un phénomène naturel qui existait avant la découverte des antibiotiques par l'homme. En effet, les molécules antibiotiques sont sécrétées par des microorganismes tels que les levures. D'ailleurs, la découverte des antibiotiques est aussi un exemple de découverte par sérendipité. Ils ont été découverts en 1928 par Alexandre Fleming à cause d'une contamination de ses cultures bactériennes par le champignon *penicillium*. L'apparition de phénomène de résistance est multiple, mutation spontanée, transferts de gènes horizontaux (échange de gènes entre deux bactéries) ou intégration d'éléments géniques mobiles comme les plasmides<sup>20</sup>. Le taux de ces événements est modulé par le nombre de gènes de résistance dans les populations bactériennes et de la pression de sélection que représente les antibiotiques. Le microbiote de l'environnement, c'est-à-dire l'ensemble des micro-organismes non pathogènes et commensaux de leurs environnements spécifiques, est une source importante d'éléments génétiques de résistance. La cohabitation microbiote humain et animal avec le microbiote environnemental permet des échanges de gènes de résistance aux antibiotiques, favorisé par diverses pollutions de l'environnement par des antibiotiques. Les pollutions de l'environnement par les antibiotiques sont multiples : utilisation des antibiotiques en santé humaine, dans l'agroalimentaire (santé animale et agriculture), l'industrie de fabrication des antibiotiques, les eaux usées non traitées<sup>20</sup>. L'augmentation des échanges à l'origine des résistances aux antibiotiques pose un problème de santé mondiale, car les bactéries pathogènes acquièrent également des résistances aux antibiotiques utilisés en santé humaine. Dans certain cas de pathogène multi-résistant, les infections bactériennes humaines et animales ne peuvent plus être traitées avec les antibiotiques existants. Les problèmes de résistance aux antibiotiques en santé humaine et agroalimentaire sont souvent présentés séparément, cependant c'est un problème global qui implique ces deux domaines d'utilisation, mais également la dimension environnementale<sup>20</sup>.

##### 1.1.3.1.2 Les solutions contre l'antibiorésistance

Il n'existe pas une solution unique pour lutter contre l'antibiorésistance des bactéries. Larsson et al. 2021<sup>20</sup> proposent par exemple la surveillance de l'apparition des résistances dans les eaux usées, un meilleur traitement des déchets de l'industrie pharmaceutique, ou encore un meilleur traitement des eaux usées<sup>20</sup>. L'utilisation des bactériophages pour la phagothérapie humaine et dans l'agroalimentaire fait aussi partie de la solution.

L'agroalimentaire est un milieu très dépendant des antibiotiques. En effet l'augmentation de la population mondiale ainsi que la volonté de sûreté alimentaire, pousse les acteurs de ce domaine à utiliser des antibiotiques tout au long du processus afin d'augmenter et d'assurer une sûreté alimentaire suffisante<sup>21</sup>. Dans la perspective de limiter la résistance des bactéries, de nouvelles études ont été



menées, pour démontrer l'efficacité des bactériophages dans le domaine de l'agroalimentaire. Par exemple l'étude de Carvalho et al. (2010)<sup>22</sup>, montre que des poussins infectés artificiellement par *Campylobacter coli* et *Campylobacter jejuni* et traités par un cocktail de trois bactériophages, présentent une réduction des pathogènes de  $10^2$  tout au long de l'expérience<sup>22</sup>. Le même type d'expérience a été mené sur d'autres pathogènes de la santé animale, mais également sur les bactéries infectant les plantes<sup>21</sup>. Les bactériophages peuvent également être utilisés sur les produits issus de l'abattage ou des récoltes, ce qui a l'avantage de limiter l'apparition de résistances chez les bactéries<sup>22</sup>.

La santé humaine peut également profiter de la phagothérapie. En France, la phagothérapie est interdite aujourd'hui, mais son utilisation compassionnelle a été autorisée à nouveau en 2016 par l'ANSM (agence nationale de sécurité du médicament et des produits de santé). Depuis lors, le projet Phag-ONE à Lyon a traité 26 patients en impasse thérapeutique. Le projet Phag-ONE s'est concentré sur les infections ostéo-articulaires complexes par la bactérie *Pseudomonas aeruginosa* car des phages produits dans des bonnes pratiques de laboratoire sont disponibles grâce à l'entreprise PherecydesPharma. Cette entreprise a également été impliquée dans le projet PhagoBurn qui représente le premier essai clinique randomisé sur les phages dans le cadre de soin des grands brûlés et dont les résultats sont détaillés dans Jault, et al 2016<sup>23</sup>. Aujourd'hui PherecydesPharma développe des préparations de phages contre les infections à *Staphylococcus* qui devraient bientôt être testées dans le cadre clinique. Le retour d'expérience de ces cas montre des succès et mais aussi des échecs, et montre que la phagothérapie est prometteuse dans les infections ostéo-articulaire. Ce type de traitement peut aussi être envisagé pour les infections urinaires chroniques et les infections pulmonaires. En revanche l'utilisation des bactériophages en médecine doit encore être étudiée par des études scientifiques pour bien maîtriser les bénéfices et les risques<sup>24</sup>. En effet, il peut exister plusieurs problèmes dans l'utilisation des bactériophages, comme une immunisation contre les bactériophages conduisant à une moindre disponibilité de ceux-ci, ou des résistances de la bactérie au phage (voir ci-après). Néanmoins les outils biotechnologiques à notre disposition permettent de modifier génétiquement les bactériophages pour qu'ils soient plus efficaces, notamment en les rendant moins sensibles au système immunitaire humain<sup>25</sup>. Il est cependant à noter que la méthode actuelle de la phagothérapie (qui consiste à trouver pour chaque infection des phages dans la banque de phage ou dans l'environnement) est une méthode empirique. Il est donc important de mettre en place une procédure qui permette une utilisation sécuritaire et efficace des bactériophages et de faire valider ce processus par les autorités sanitaires. Un enjeu majeur de ces procédures réside en la production de bactériophages en évitant les contaminations bactériennes tel que les endo ou exo toxines bactériennes par exemple. En revanche il a déjà été montré une très bonne tolérance de l'utilisation des bactériophages en santé<sup>25</sup>.

Bien entendu, l'utilisation des bactériophages, peut également créer des résistances chez les bactéries. Cependant les phages co-évolue avec les bactéries permettant ainsi de contourner les

résistances des bactéries. De plus la résistance des bactéries face au phage peut s'accompagner d'une perte de virulence et parfois d'une re-sensibilisation aux antibiotiques. Il sera cependant important de surveiller l'apparition des résistances en cas d'utilisation importante des phages pour la thérapie pour ne pas reproduire les erreurs du passé<sup>26</sup>.

#### 1.1.3.2 Applications technologiques & médicales :

Comme décrit dans le paragraphe précédent, la phagothérapie est une application importante des bactériophages dans la lutte contre les infections. Cependant il existe aussi d'autres applications médicales.

Les lysozymes de la famille des peptidoglycanes hydrolases, découverts par Alexandre Fleming, en 1922, sont utilisés depuis une dizaine d'années comme traitement prophylactique<sup>25</sup>. Ils permettent notamment de traiter les contaminations par les bactéries Gram positive car leur peptidoglycane est directement accessible. Le lysozyme est largement utilisé dans l'industrie de la fermentation pour maîtriser les contaminations par les bactéries à Gram positive<sup>25</sup>.

Plus récemment les endolysines, qui sont produites par les bactériophages en fin d'infection afin de digérer le peptidoglycane et faciliter la libération des virions, ont été testées pour des traitements curatifs chez les mammifères chez lesquels elles ont fait régresser les infections<sup>25</sup>. Le même type d'expérience a été mené chez les plantes et a conduit aux mêmes résultats<sup>25</sup>. Comme mentionné dans le chapitre précédent, la phagothérapie peut entraîner des résistances bactériennes, cependant aucune résistance aux endolysines n'a été observée en laboratoire jusqu'à présent<sup>25</sup>. Les bactériophages produisent également des enzymes de digestion du peptidoglycane à l'extrémité de leur queue. Les hydrolases murines sont utilisées par les phages pour faciliter l'injection de l'ADN à l'intérieur de la bactérie. L'équipe de Paul et al. (2011), a testé, avec succès, une hydrolase fusionnée au domaine de liaison de la paroi de staphylocoque afin de traiter une infection.

D'un point de vue prophylactique, les bactériophages peuvent être utiles dans l'élaboration de nouveaux vaccins. Deux approches différentes sont explorées, le « phage DNA vaccin » et le « phage display vaccin ». La première approche consiste à utiliser des capsides de bactériophages comme vecteur. La capside contient l'ADN codant pour l'épitope à produire sous le contrôle d'un promoteur eucaryote fort. La capside sera reconnue par les macrophages qui intégreront la particule. Son ADN sera libéré dans le cytoplasme d'une cellule humaine qui produiront alors l'épitope à leur surface pour le présenter au système immunitaire<sup>25</sup>. La deuxième approche consiste à se servir de la capside des bactériophages comme présentateur d'épitope tout en activant le système immunitaire<sup>25</sup>. Il est à noter que dans les deux approches, les protéines de surface des bactériophages peuvent également inclure des facteurs d'attachement pour cibler spécifiquement les cellules dendritiques<sup>25</sup>.

Outre les applications médicales, les bactériophages ont conduit à des progrès technologiques. Comme décrit ci-dessus, ils ont fourni des outils efficaces pour la biologie moléculaire. D'autres technologies continuent à être mises au point grâce aux phages. Prenons l'exemple du phage display, développé par George Smith en 1998. Les applications sont multiples comme la sélection des anticorps monoclonaux ou la cartographie des épitopes<sup>27</sup>.

#### 1.1.4 Relations phage - bactérie et cycle de vie des phages

##### 1.1.4.1 Les cycles de vie des bactériophages

L'interaction entre le bactériophage et la bactérie est un point clé de l'infection et est largement étudiée pour les phages modèles. Classiquement, on présente trois types de cycle de vie.

Le cycle lytique dans lequel le bactériophage se lie irréversiblement à son récepteur spécifique (e.g LPS (Lipo-polysaccharide), protéine ou acides téichoïques) à l'aide d'une ou plusieurs copies de ses RBP (Receptor Binding Protein). Le matériel génétique du phage (sbADN, dbADN, sbARN ou dbARN) est alors injecté dans la bactérie. Les bactériophages protègent leur génome par différentes stratégies et le plus souvent le génome de la bactérie est détruit. Le phage détourne alors la machinerie de la bactérie pour répliquer son génome, produire ses protéines de structure et produire des protéines pour la lyse bactérienne comme des endolysines ou des holines. Les virions sont assemblés dans le cytoplasme bactérien et la lyse de la bactérie permet la libération des virions qui peuvent aller infecter une nouvelle cellule (figure 3). Ce type de cycle est généralement relativement court (quelques dizaines de minutes) et les bactéries sont rarement protégées contre une sur-infection par le phage<sup>28</sup>.

Le cycle lysogénique commence aussi par la reconnaissance de l'hôte et l'injection du génome viral dans le cytoplasme bactérien. Dans ce cycle, le génome viral est alors intégré au génome bactérien sous forme de prophage. La réplication du prophage est réalisée en même temps que la réplication du chromosome bactérien lors de la division cellulaire. Un répresseur viral maintient le prophage dans le génome bactérien en se fixant sur l'opérateur viral impliqué dans le cycle lytique. En condition de stress, le prophage est capable de s'exciser du génome bactérien pour repartir dans un cycle lytique. Il est fréquent que le prophage apporte des gènes actifs, conférant un avantage évolutif à l'hôte. C'est le cas par exemple du prophage CTX $\phi$  qui apporte à la bactérie *Vibrio Cholerae* la toxine du choléra, qui cause des diarrhées brutales chez l'Homme, favorisant la dissémination de la bactérie. Les bactériophages lysogéniques bloquent également souvent leurs récepteurs d'entrée pour protéger l'hôte d'une surinfection par un phage proche (voir paragraphe 1.1.6). Toutes les cellules portant le prophage sont alors protégées de la surinfection (Figure 3)<sup>28</sup>.

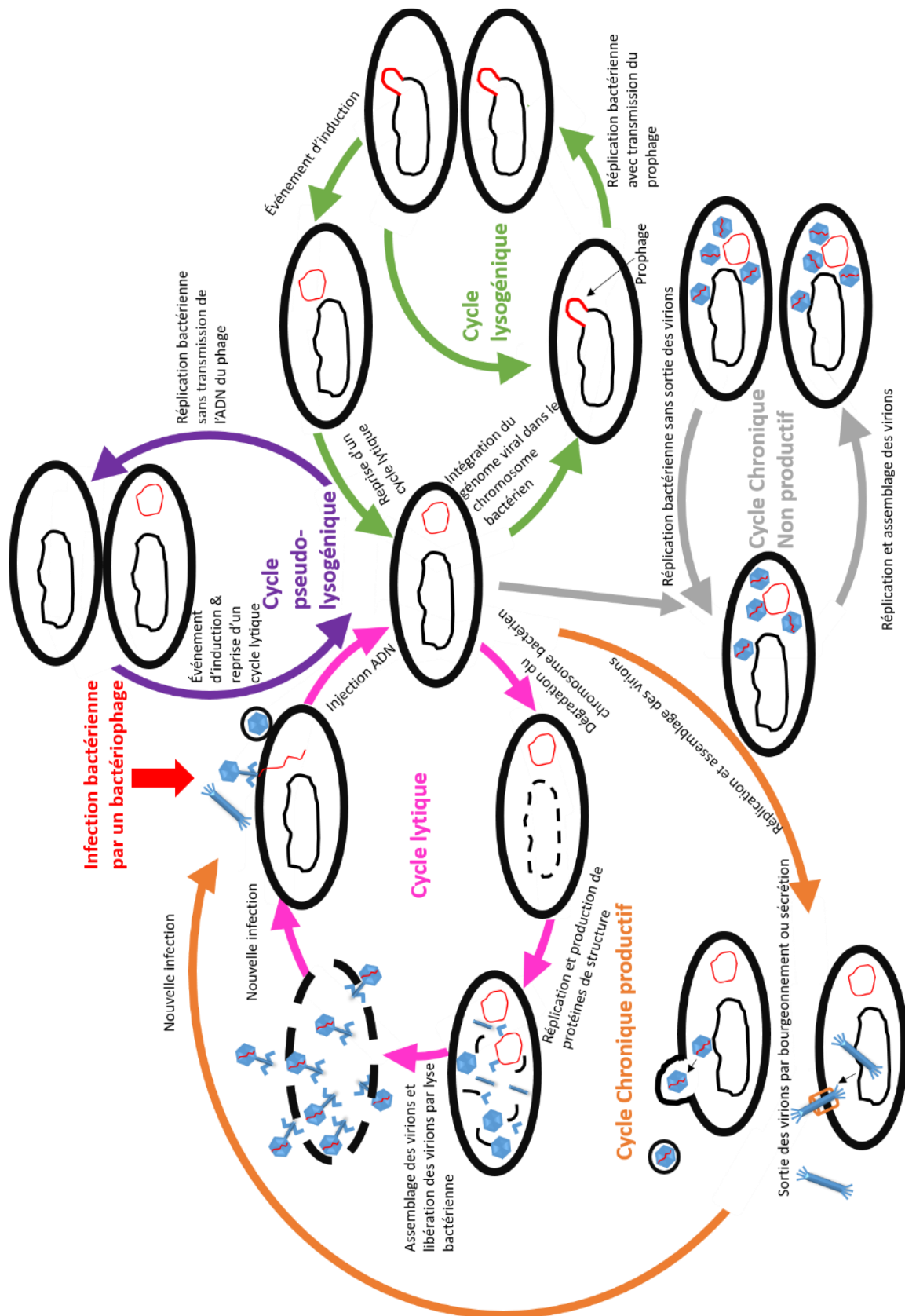
Le cycle chronique productif pour lequel le bactériophage est produit de manière continue sans lyse bactérienne. Le génome viral est injecté dans l'hôte et peut soit être intégré dans le génome bactérien, soit rester dans le cytoplasme. Les virus se répliquent grâce à la polymérase de l'hôte et les

nouveaux virions sont assemblés par bourgeonnement ou par sécrétion au travers de la membrane. C'est le cas par exemple pour les phages filamenteux M13, ff et f1 infectant *Escherichia coli*<sup>28</sup>.

Récemment des études ont montré qu'en plus des cycles classiques, les phages pouvaient adopter des cycles alternatifs en fonction de la bactérie infectée, de l'état physiologique de l'hôte et / ou des conditions environnementales. On distingue donc 2 autres cycles de vie pour les bactériophages que les cycles lytique et lysogénique.

Le cycle chronique non productif a été décrit une fois pour le bactériophage phi6 infectant *Pseudomonas*. Les variants de *Pseudomonas* résistant à phi6 sont obtenus facilement et certains montrent une grande quantité de virus dans le cytoplasme. L'hôte est protégé d'une surinfection et est capable de se multiplier. Ce mode d'infection de phi6 ressemble au cycle d'un virus ARN double brins fongique n'ayant pas de cycle de vie extracellulaire et où les virions sont transmis de manière intracellulaire (Figure 1-3)<sup>28</sup>.

Dans le cycle pseudo-lysogénique, le génome du phage est injecté dans le cytoplasme bactérien mais est peu actif (ni réplication, ni production de virion, ni intégration de l'ADN virale dans le chromosome bactérien). L'hôte est capable de se répliquer mais le génome du phage est transmis de façon asymétrique aux cellules filles. S'il existe un système d'exclusion de surinfection, seules les bactéries contenant le génome viral sont protégées. Il existe donc au sein de la descendance de la cellule infectée des bactéries sensibles au phage et d'autres insensibles. Pour le moment il n'est pas clair si ce type de cycle est un style de vie du virus ou un temps d'attente pour entrer dans un cycle lytique ou lysogénique (Figure 1-3)<sup>28</sup>.

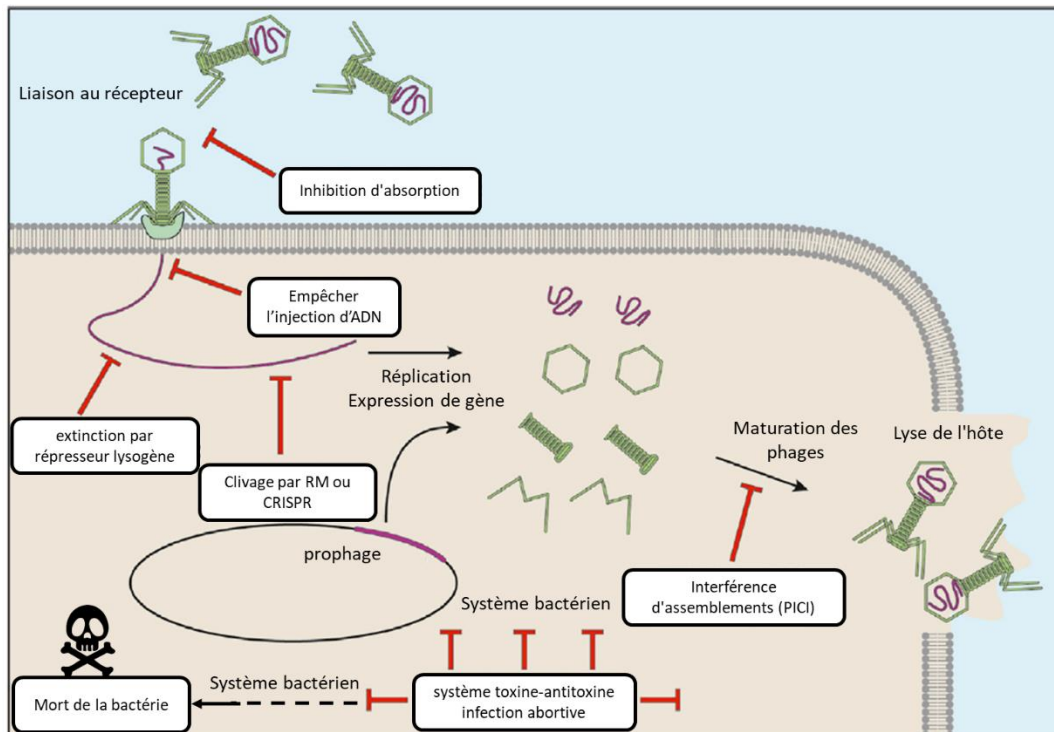


**Figure 1-3 : Cycle d'infection possible pour les bactériophages.**

Les cycles des bactériophages commencent toujours par une infection (flèche rouge) puis l'introduction du génome viral dans la cellule bactérienne (flèche rose injection ADN). Ensuite, 4 cycles différents sont décrits : cycle lytique (rose), cycle chronique productif (orange), cycle lysogénique (vert), plus rarement cycle pseudo-lysogénique, exceptionnellement cycle chronique non productif (gris) inspirée de 28.

### 1.1.4.2 Relation phage - bactérie, la course à l'armement

Les bactériophages et les bactéries cohabitent dans tous les habitats naturels. On estime que les bactériophages tuent environ 20 % de la biomasse bactérienne par jour<sup>29</sup> et jouent donc un rôle important pour la régulation des populations bactériennes. Cependant les bactéries ne sont pas sans défense contre les infections des phages. Les bactéries peuvent se protéger des phages en agissant sur des étapes clé de l'infection et empêcher celle-ci (Figure 1-4). Dans certains cas, la résistance ne permet pas l'absorption, c'est le cas des résistances liées au récepteur. Dans d'autres cas l'absorption est possible mais n'aboutit pas au relargage de virions, c'est le cas de tous les autres systèmes de défense des bactéries. La bactérie est aussi capable de se sacrifier pour protéger la population bactérienne générale avec des systèmes de toxine-antitoxine qui permettent d'abrégé le cycle viral mais provoquent aussi la mort de l'hôte (Figure 1-4). Un des systèmes de défense des bactéries et le système CRISPER-Cas9. C'est un système immunitaire adaptatif permettant de détecter un ADN viral et de le couper spécifiquement après une première infection. Je ne développerai pas tous les systèmes de résistance des bactéries qui sont détaillés dans plusieurs revues<sup>30-32</sup>. La relation phage – hôte est un système complexe qui doit être étudié plus intensément dans le contexte d'enjeux de la phagothérapie. En outre une bonne connaissance de ces systèmes peut apporter des solutions pour protéger les bactéries pour les processus industriels (pharmaceutique, agroalimentaire, chimique, ect.).

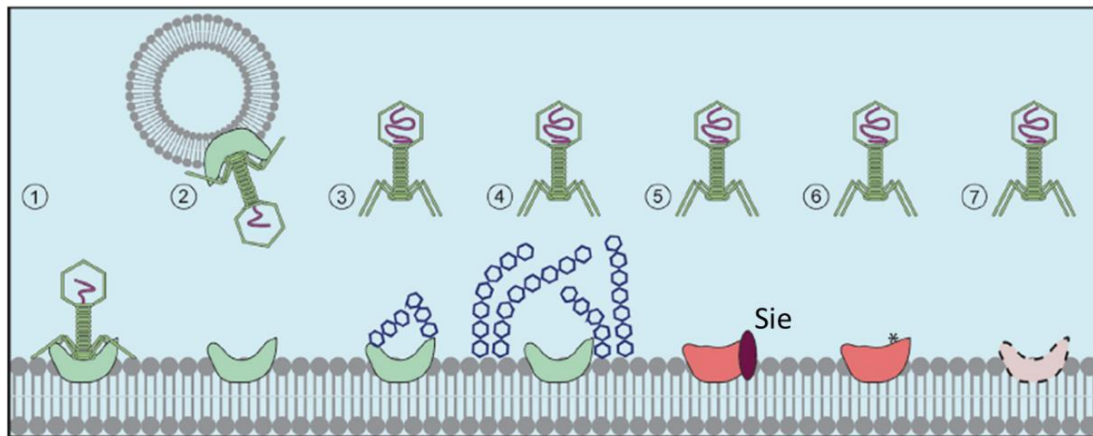


**Figure 1-4 : Cycle d'infection des bactériophages et mécanismes de résistance des bactéries.**

Pour une infection réussie, un bactériophage réalise un cycle productif (flèche noire) qui nécessite d'interagir avec son récepteur, faire entrer son ADN dans l'hôte, répliquer son ADN et exprimer les gènes de structure. Les virions sont ensuite assemblés. L'accumulation de virions et la production d'endolysine permettent la sortie du phage par lyse de l'hôte. Les processus de défense des bactéries peuvent cibler chacune de ces étapes (flèche d'inhibition rouge). Figure adaptée de 31

Le moyen le plus efficace pour protéger la cellule de l'infection par un bactériophage est d'éviter l'interaction entre le phage et son récepteur. On peut décrire 5 mécanismes différents. La bactérie peut sécréter des vésicules de membranes externes (OMVs) contenant le récepteur du phage servant de leurre (Figure 1-5 (2)). Les bactériophages injectent alors leur ADN dans une vésicule vide ne permettant pas la réalisation du cycle viral. Le récepteur peut être inaccessible, soit en masquant la surface d'interaction par des capsules de glycane à la surface bactérienne, soit par son absence (régulation de l'expression du récepteur) (Figure 1-5 (4 et 7)). La reconnaissance du récepteur par la RBP du phage peut être affectée soit par des modifications post traductionnelles comme des glycosylations soit par des mutations ponctuelles du récepteur (Figure 1-5 (3 et 6)).

Parfois les phages eux-mêmes protègent la cellule hôte d'une surinfection en modifiant le récepteur d'entrée par la liaison d'une protéine d'exclusion de surinfection (Figure 1-5 (5)). Je détaillerai plus longuement cette résistance dans le paragraphe 1.1.5 phénomène d'exclusion de surinfection.



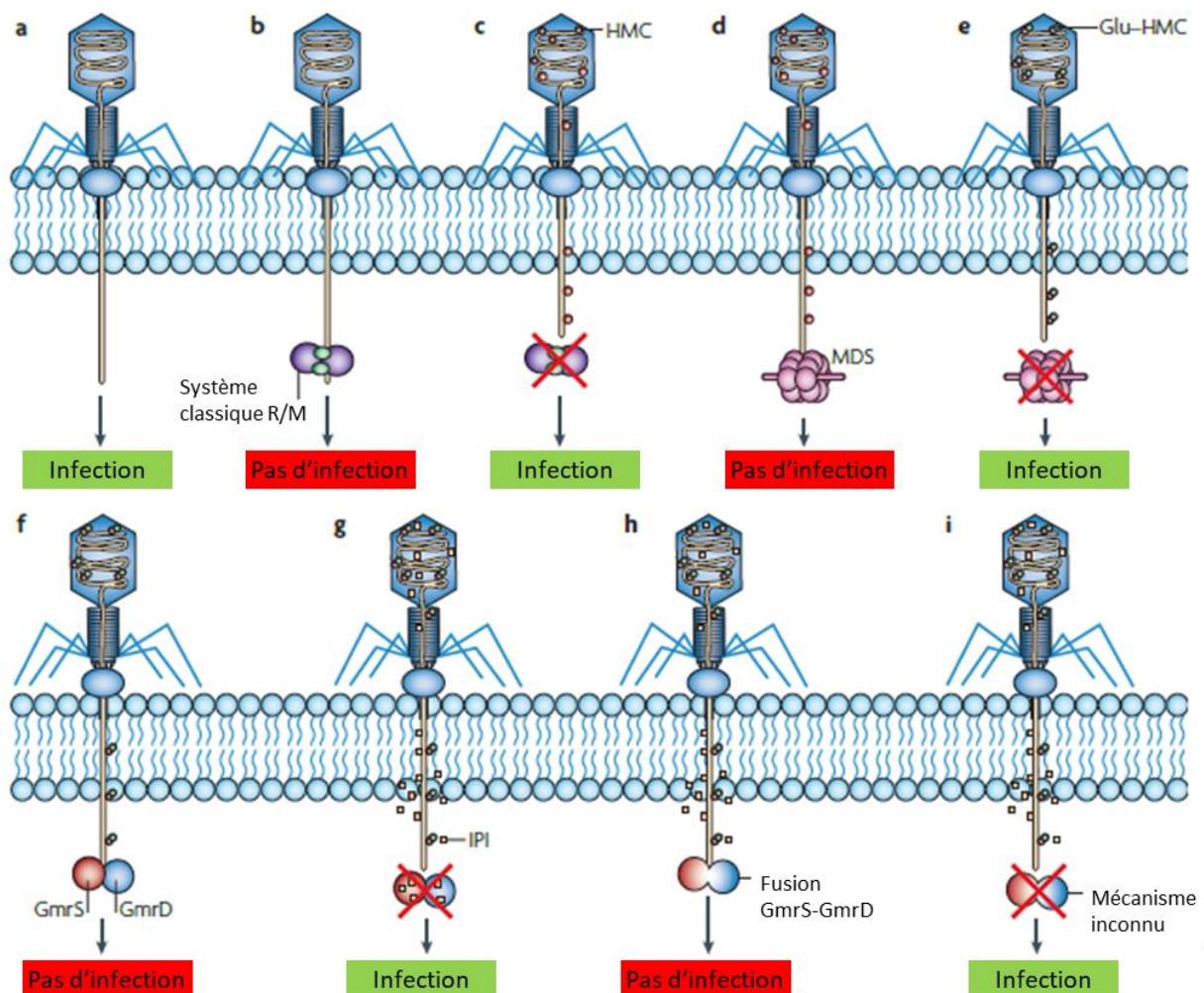
**Figure 1-5 : différents mécanismes de prévention d'absorption de phage sur leur récepteur.**

1. Reconnaissance du récepteur par la RBP du phage, infection possible. 2. Production de vésicules de membrane externe pour séquestrer le phage par une interaction avec les récepteurs contenus dans les vésicules. 3. Modifications post traductionnelles du récepteur empêchant sa reconnaissance par les RBPs. 4. Modification de la surface bactérienne pour occluser le récepteur et empêcher l'accès des RBPs. 5. Modification du récepteur par liaison d'une protéine d'exclusion de surinfection produite par les phages. 6. Modification du récepteur par des mutations ponctuelles empêchant sa reconnaissance par les RBPs. 7. Régulation de l'expression du récepteur le rendant inaccessible par son absence. Figure issue de 31.

Les mécanismes de défenses des bactéries sont multiples, et surtout ne sont pas figés. En effet les populations de phages et de bactéries sont impliquées dans des cycles continus de coévolution. Il existe plusieurs exemples documentés de coévolution phage-bactérie. Par exemple le système d'immunité CRISPR-Cas9 peut être inhibé par des protéines de la famille Acrs. La protéine AcrIIA6 est capable d'inhiber le système CRISPR-Cas9 en se liant à la protéine Cas9 provoquant sa dimérisation<sup>24</sup>. Cette même étude décrit des variants naturels de Cas9 résistants à l'inactivation par AcrIIA6.



Un autre exemple de coévolution est l'échappement du bactériophage T4 au système de restriction modification classique (Figure 1-6 (a,b))<sup>32</sup>. Pour échapper aux enzymes de restriction le phage a intégré dans son génome des bases particulières d'hydroxyméthylcytosine (HMC) (Figure 1-6 (c)). En réponse, une sous population bactérienne a acquis un système de clivage spécifique des bases modifiées (MDS) (Figure 1-6(d)). Le bactériophage a alors acquis des glycosylations (glu-HMC) pour préserver son génome des coupures (Figure 1-6 (e)). Encore une fois la population bactérienne a répliqué en produisant les protéines GmrS, GmrD permettant de cliver les bases glu-HMC (Figure 1-6 (f)). Une résistance à ce système apparait chez le phage T4 avec une protéine interne (IPI) qui entrave le système Gmr et permet à ces phages d'infecter avec succès les souches *E. coli* contenant ce système (Figure 1-6 (g)). Et ainsi de suite... Cet exemple montre parfaitement la coévolution imposée par les deux parties, d'un côté par le phage pour pouvoir infecter la population bactérienne, de l'autre par les bactéries pour empêcher l'infection.



**Figure 1-6 : La course aux armements entre les bactéries et les phages.**

Exemple de stratégies du phage T4 pour échapper aux systèmes de restriction-modification (R-M) et aux systèmes dépendants des modifications (MDS) qui sont codés par *E. coli* pour résister à l'infection par le phage. (a) à (g) voir le texte. h. Protéine de fusion GmrS-GmrD ne peut plus être inhibée par la protéine IPI. i. le phage a contourné la fusion GmrS-GmrD par un système inconnu. Figure issue de 32.



### 1.1.5 Phénomène d'exclusion de surinfection

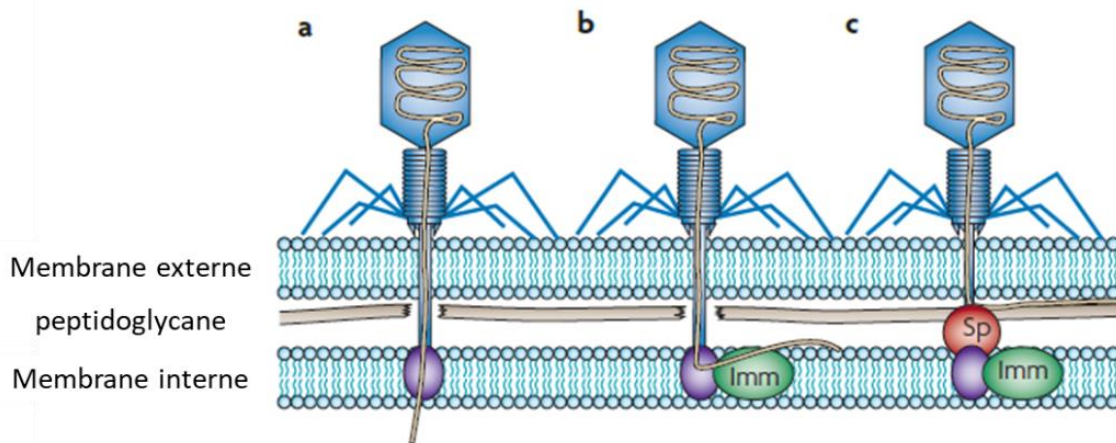
Le phénomène d'exclusion de surinfection (Sie) est la capacité pour un virus à bloquer ou interférer avec une seconde infection virale par le même virus ou un virus génétiquement proche. Des phénomènes de Sie ont été décrits pour tous les systèmes virus hôte (chez l'homme, les animaux, les plante et les virus procaryotes)<sup>33,34</sup> (Tableau 1-1).

Family	Genus	Virus	Superinfection exclusion protein	strategy
<i>Baculoviridae</i>	<i>Alphabaculovirus</i>	AcMNPV	undocumented	undocumented
<i>Closteroviridae</i>	<i>Closterovirus</i>	Citrus tristeza virus	p33	undocumented
<i>Flaviridae</i>	<i>Hepacivirus</i>	Hepatitis C	Undocumented	post entry
<i>Siphoviridae</i>	<i>Lambdaliikevirus</i>	Enterobacteria phage HK97	gp15	Undocumented
<i>Siphoviridae</i>	<i>T5likevirus</i>	Enterobacteria phage T5	Lipoprotein Lip	Blocks T5 entry receptor FhuA
<i>Myoviridae</i>	<i>T4likevirus</i>	Enterobacteria phage T4	Immunity protein Imm Protein spackle Sp	Blocks DNA translocation into host cytoplasm Inhibits the lysozyme activity
<i>Siphoviridae</i>	unclassified	Bacteriophage TP-J34	Lipoprotein LTP	Blocks DNA ejection into host cytoplasm
<i>Siphoviridae</i>	unclassified	Lactococcus phage Tuc2009	ORF2	Blocks DNA ejection into host cytoplasm
<i>Podoviridae</i>	<i>P22likevirus</i>	Enterobacteria phage P22	SieA	Blocks DNA translocation into host cytoplasm
<i>Myoviridae</i>	<i>Punaliikevirus</i>	Enterobacteria phage P1	Sim	Probably blocks DNA translocation into host cytoplasm

Tableau 1-1 : Exemple de virus possédant un système d'exclusion de surinfection.  
Consulté le 10/06/2022 sur <https://viralzone.expasy.org/3971>

Chez les bactériophages, plusieurs systèmes de Sie ont été décrits chez des prophages suggérant qu'ils sont importants pour les interactions phage-phage plutôt que pour les interactions phage-hôte. C'est le cas du phage HK97, dont le récepteur est LamB, et qui est capable de protéger son hôte contre l'infection par les phages HK97 et HK75<sup>35</sup>. La protéine Sie identifier est la protéine gp15 dont les prédictions de structure secondaire donne une hélice transmembranaire. Les auteurs ont alors testé l'entrée de l'ADN dans la cellule puisque qu'elle est corrélée à un export de potassium depuis la cellule vers le milieu extracellulaire. Dans cette étude il est démontré que la protection de l'hôte par le système Sie est dû au blocage de l'injection d'ADN dans le cytoplasme. En revanche certain phage utilisant le récepteur LamB sont insensible à la présence de gp15 démontrant que le récepteur n'est pas bloqué. Les auteurs proposent que le blocage du transfert d'ADN se fait en interagissant avec des composants non identifiés du système de transfert de l'ADN du phage.

Les systèmes Sie sont aussi retrouvés dans les bactériophages strictement lytiques comme le coliphage T4. Ce phage possède deux systèmes Sie portés par les protéine Imm et Sp qui protègent contre la surinfection des bactériophages de type T ciblant le récepteur OmpC. Ces deux systèmes agissent sur l'injection de l'ADN dans le cytoplasme de l'hôte au niveau de la membrane interne (Imm) ou du peptidoglycane (Sp) (Figure 1-7). Dans ce cas l'absorption des phages n'est pas affectée.



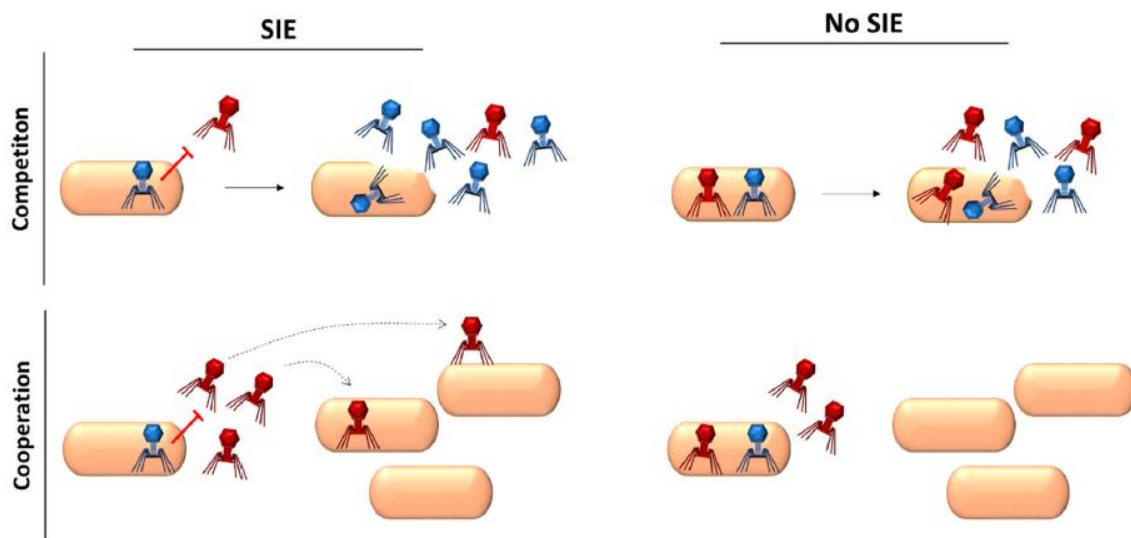
**Figure 1-7 : Blocage de l'injection de l'ADN dans la bactérie par le système Sie du coliphage T4.**

*a. Au cours de l'infection le peptidoglycane est digéré et l'ADN est injecté dans le cytoplasme de l'hôte en passant par les membranes externe et interne. B. La protéine Imm bloque le passage de l'ADN dans la membrane interne ne permettant pas l'infection. c. La protéine Sp agit en synergie avec la protéine Imm et bloque la digestion du peptidoglycane empêchant l'ADN viral d'être injecté dans le cytoplasme de l'hôte. Figure issue de 32.*

Certaines Sie agissent directement sur le récepteur du phage probablement en en modifiant la structure. Dans ce cas, l'absorption irréversible du phage n'est plus possible. C'est le cas des protéines Cor codées par les phages mEp167, mEp213, HK022, N15 et  $\phi 80$ . Un article récent Arguijo-hermandez et al. 2018<sup>36</sup> montre que les protéines Cor<sub>mEp197</sub> sont des lipoprotéines adressées à la membrane externe. Ces protéines protègent l'hôte contre les phages utilisant FhuA comme récepteur, mais cette protection n'est pas assurée dans les cellules ne possédant pas les protéines OMP(s) (outer membrane protein). Les auteurs supposent donc que les protéine OMP(s) ont un rôle dans la résistance, mais ne proposent pas de mécanisme précis. La structure des protéines Cor n'est pas connue<sup>36</sup>.

La protéine Llp du phage T5 est aussi une lipoprotéine d'exclusion de surinfection qui partage des caractéristiques des protéines Cor (voir 1.2.4. La protéine Llp).

Le rôle évolutif des Sie peut être interprété soit par une utilité dans la compétition entre phages soit par une utilité de coopération. Dans le cas de la compétition la présence de Sie permet à un phage d'utiliser l'ensemble des ressources de la bactérie pour se reproduire et donc augmenter la capacité d'adaptation du virus. Dans le cas de la coopération, le coût supplémentaire d'assurer la protection de la cellule, permettrait d'aider les autres membres de la population à trouver les cellules susceptibles qui sont disponibles pour produire de nouveau virion (Figure 1-8)<sup>37</sup>. En outre dans, M. hunter et al <sup>38</sup>, les auteurs montrent que la surinfection permet une meilleure diversité génétique des virus, en revanche l'exclusion de surinfection apporte un avantage à court terme. En effet en cas de mélange de sous populations virales permettant ou non la surinfection, c'est toujours celle qui porte un système d'exclusion qui l'emporte. Ceci indique que l'exclusion de surinfection augmente la sélection à court terme expliquant l'abondance de ce phénomène dans la nature<sup>38</sup>.



**Figure 1-8 : Avantages et inconvénients liés au Sie en fonction des mécanismes évolutifs de compétition ou de coopération.** En compétition le système Sie permet de garder les ressources cellulaires pour un type de phage unique. Dans la coopération le système Sie permet au phage de mieux se diffuser dans la population bactérienne et utiliser toutes les ressources à disposition pour une population de virus. Issu de 37.

### 1.1.6 Le phage T5

Le bactériophage T5 appartient à la famille des *Demereviridae*, possédant une capside et une longue queue flexible non contractile. C'est un phage strictement lytique, qui infecte *Escherichia coli*. À l'extrémité distale de la queue du bactériophage T5, la protéine de liaison au récepteur, pb5 reconnaît comme récepteur la protéine FhuA également reconnue par les phages T1,  $\Phi$ 80 et UC-1. La fonction physiologique de FhuA est de transporter le fer ferrichrome mais FhuA est également capable de transporter certains antibiotiques comme l'alomycine, la rifamycine, la colicine M et la microcine J25<sup>39</sup>.

#### 1.1.6.1 Génome

Le génome du bactériophage T5 (AY587007) contient 121 752 bp dont 111 613 bp sont des séquences non répétées. Les régions répétées sont comprises entre les séquences de 0 à 10 129 pb au début du génome, puis à la fin du génome ces mêmes 10 129pb sont répétées. L'ADN du phage T5 contient en moyenne 39.3% de CG, et 168 Open Reading Frame (ORF) (Figure 1-9). Le bactériophage T5 a une densité de gènes (ORFs ou cadres ouverts de lecture) de 83,1% ce qui est considéré comme faible pour un procaryote, mais qui s'explique par la présence d'une longue région répétée aux deux extrémités du génome et de plusieurs régions regroupant de nombreux d'ARN de transfert<sup>40</sup>. L'ADN de T5 est interrompu par l'absence de liaison phosphodiester entre deux nucléotides, en moyenne en 8 endroits, sur le brin sens uniquement. Toutes ses interruptions sont précédées d'un site de reconnaissance d'une endonucléase. Seulement 4 positions sont quasi systématiquement clivées (dans 80-90% des particules virales). Pour le reste des séquences de clivage, la distribution n'est pas reproductible. A ce jour, on ne sait pas à quoi servent ces coupures, mais on sait qu'elles ne sont pas essentielles. En effet le mutant st0 possède une délétion sur un site de coupure majeur et reste infectieux<sup>40</sup> et il existe un mutant T5amHA911 dit « nickless », chez lequel le gène codant pour l'endonucléase responsable de la formation de ces interruptions simple brin ou « Nicks » qui affecte le gène codant pour l'endonucléase responsable de la formation de ces interruptions simple brin et qui également infectieux<sup>41</sup>.

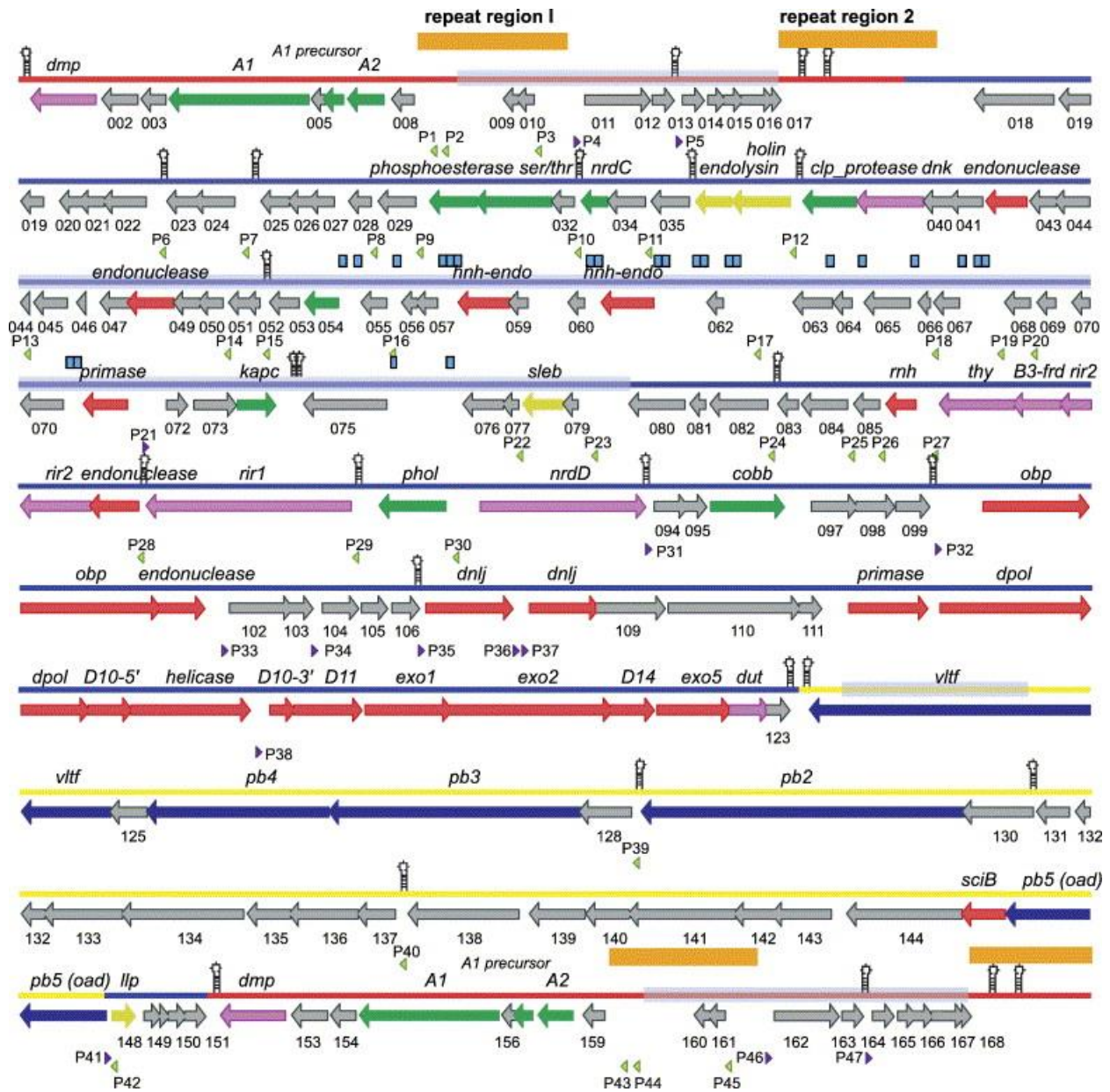
Le génome se compose de trois régions fonctionnelles<sup>40</sup>.

Une région très précoce (« pre-early » en anglais) qui est composée de gènes codant pour des protéines permettant d'hydrolyser l'ADN bactérien, d'inactiver la transcription et la traduction de l'hôte, bloquant ainsi toute répllication du génome bactérien. Cette région correspond aux 8 premiers % de l'ADN qui sont injectés dans la cellule hôte. La reprise d'éjection de l'ADN est conditionnée par l'expression de ces gènes.

Une grande région dite précoce où 111 ORF ont été identifiés. Les protéines codées par cette région sont exprimées entre 5 et 10 min après l'infection environ. Elles contiennent des gènes codant pour des protéines impliquées dans la répllication, la recombinaison, la transcription, le métabolisme des nucléotides, ou encore dans la fonction lytique du phage comme les endolysines ou les holines<sup>40</sup>. On note que toutes ces séquences sont situées en amont des gènes tardifs à l'exception du précurseur de

la lipoprotéine de conversion lytique (Llp) qui se situe juste avant les séquences très précoces répétées en fin de génome (Figure 1-9). En effet ce gène se situe juste après le gène *oad* (o-antigène dépendant), gène tardif codant pour pb5, protéine de liaison au récepteur, qui cible également FhuA. Cette position particulière de ces deux gènes à cheval entre les régions précoce et tardive permet de limiter les recombinaisons génétiques qui pourraient séparer ces deux gènes qui agissent en synergie autour de FhuA. Les ORF 148 à 151 qui suivent *llp* ont été proposés pour être des protéines chaperonnes aidant au repliement de Llp, mais cette hypothèse n'a pas été investiguée<sup>40,42</sup>.

La dernière région dite tardive, exprimée à partir de 10 min d'infection et jusqu'à la fin de l'infection (c'est-à-dire environ 30 min), contient principalement des gènes de structure du phage T5, notamment pb5. On note quand même que l'ORF 145 code pour la protéine sciB, une endonucléase. L'organisation des gènes de structure sera détaillée et comparée à d'autres phages dans la section 1.3.2 Aspect structuraux.



**Figure 1-9 : Structure du génome du bactériophage T5.**

Les régions très précoce, précoce et tardive sont indiquées respectivement par des lignes de base rouges, bleues et jaunes. Les gènes et leurs directions de transcription sont indiqués par des flèches colorées indiquant les directions de la transcription. Les gènes sont colorés en fonction de leur catégorie de fonction : Réplication et réparation de l'ADN, saumon ; métabolisme des nucléotides, prune ; interaction avec l'hôte, jaune ; autres enzymes, vert ; protéines structurales, bleu ; fonction inconnue, gris ; les promoteurs sont représentés par des triangles violets (gènes codés par le brin plus) ou vert (gènes codés par le brin moins). Les terminateurs transcriptionnels Rho-indépendants sont indiqués par un symbole en forme d'épingle à cheveux. Les régions riches en répétitions sont encadrées en orange au-dessus des points de repère. Cette figure est reprise sans modification de <sup>40</sup>

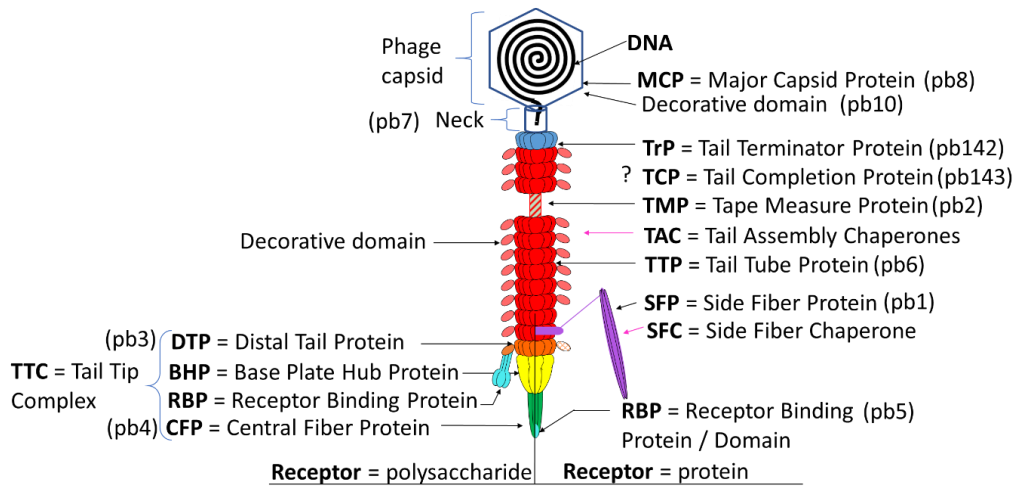


### 1.1.6.2 Aspect structuraux

Le phage T5 est composé d'une capsidie icosaédrique d'un diamètre moyen de 900 Å<sup>43</sup> et d'une queue d'un diamètre de 150 Å pour une longueur totale de 1900 Å<sup>44</sup>.

La capsidie est composée de 775 copies de la protéine majeure de capsidie pb8, organisée en hexamères sur les faces de l'icosaèdre et en pentamères sur 11 des sommets. Sa paroi a une épaisseur de 40 Å. Cette capsidie porte 120 copies de la protéine de décoration appelée pb10, fixée au centre de chacun des hexamères<sup>45</sup>. Elle contient un ADN compacté ; en microscopie électronique, l'ADN est visible sous forme de couches concentriques de densités séparées de 24,4 Å<sup>43</sup> qui montre que l'ADN est hautement compacté dans la capsidie. Quand la capsidie est vide, le diamètre de celle-ci augmente d'environ 10 Å<sup>43</sup>. La capsidie est fermée par la protéine portale pb7, un dodécamère localisé à un des vertex de l'icosaèdre<sup>46</sup>. La capsidie du bactériophage T5 est reliée à la queue par la protéine p144 protéine qui ferme la capsidie et qui est connectée à la protéine p142 protéine de terminaison de la queue lors de l'assemblage tête-queue (TrP)<sup>41</sup> (Figure 1-10). Contrairement à ce qui avait été proposé en raison de sa position dans le génome, la protéine p143 nommée protéine de complétion de queue (TCP), n'est pas présente entre la capsidie et la queue, mais au niveau distal de la queue<sup>47</sup>. La fonction de cette protéine n'est pas connue (Figure 1-10 et 1-11).

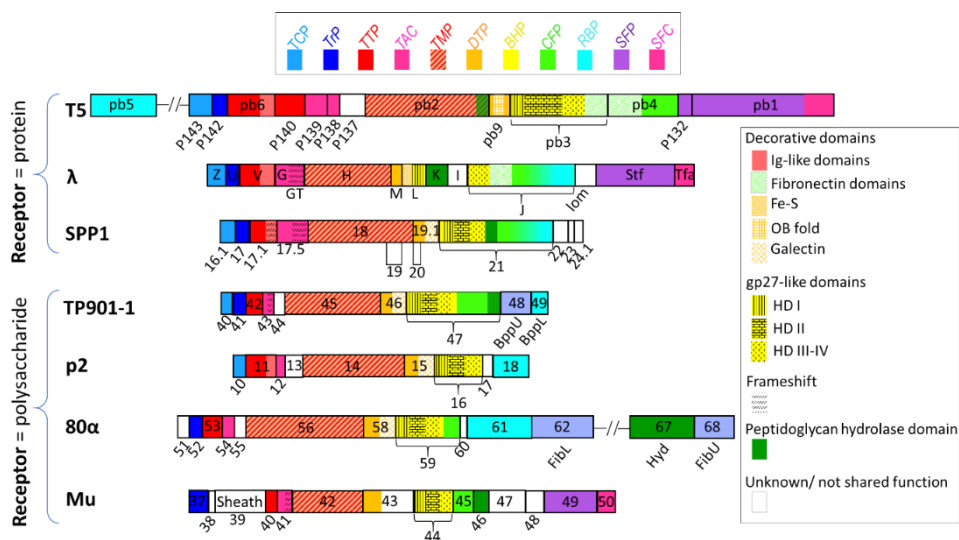
La partie proximale de la queue forme un tube de 190 nm de long et la partie distale forme le complexe de l'extrémité de queue (TTC) directement impliqué dans la reconnaissance de l'hôte. Le long tube de la queue est formé par l'oligomérisation de la protéine de tube pb6 (TTP) autour de la protéine vernier pb2 (TMP) associés en trimère. La protéine pb2 possède un site de coupure séparant une partie C terminale avec une activité peptidoglycane hydrolase et de la partie formant la protéine vernier. Comme la plupart des protéines de tube, pb6 possèdent un domaine de décoration, dans le cas de T5 il s'agit d'un domaine Ig-like. Contrairement à la queue de la plupart des phages dont la protéine de tube a une symétrie d'ordre 6, pb6, résultant d'une duplication/fusion de gène, adopte une symétrie d'ordre 3<sup>48</sup>. Le complexe de l'extrémité de la queue est composé de la protéine distale de queue (DTP) pb9 formant un hexamère, d'une protéine de plaque basale pb3 (BHP) qui est trimérique, de la protéine de fibre centrale pb4 (CFP) trimérique qui forme un réseau allongé de domaines fibronectine (FNIII) et une hélice bêta<sup>47</sup>. pb3 est la protéine trimérique qui permet de fermer le tube, elle est composée d'une partie proximale continuant le tube, fermée par un bouchon et d'une partie distale composée de 2 domaines fibronectine, qui forment le début de la fibre droite<sup>47</sup>. À l'extrémité distale de la queue on retrouve pb5. La queue du phage T5 contient également des fibres en L composées de la protéine pb1<sup>41</sup>(Figure 1-10).



**Figure 1-10 : Représentation schématique des siphophages.**

Architecture type de la queue des siphophage avec les protéines de structure conservées. Les flèches roses indiquent les chaperonnes. Les protéines TTP et DTP ne présentent pas toujours des domaines de décoration. Lorsque le récepteur est un polysaccharide, les RBPs sont présentes en plusieurs exemplaires et sont liées directement ou indirectement à la DTP, et la CFP peut être absente (à gauche). Lorsque le phage cible une protéine comme récepteur principal, la RBP est située à l'extrémité de la CFP, comme une protéine individuelle ou un domaine de la CFP (à droite). Les noms des protéines spécifiques du bactériophage T5 sont entre parenthèses. Les fibres latérales ne sont pas toujours présentes. Figure tirée de 48.

La structure des protéines composant la queue des siphophages est conservée, mais également leur organisation dans le génome <sup>48</sup>(Figure 1-11). Le phage T5 présente une exception pour pb5, la protéine de liaison au récepteur (RBP) le gène ne se situe pas à proximité des gènes de la fibre droite comme dans les autres siphophages mais du côté des gènes de la capsid (voir ci-dessus, paragraphe 1.6.1).



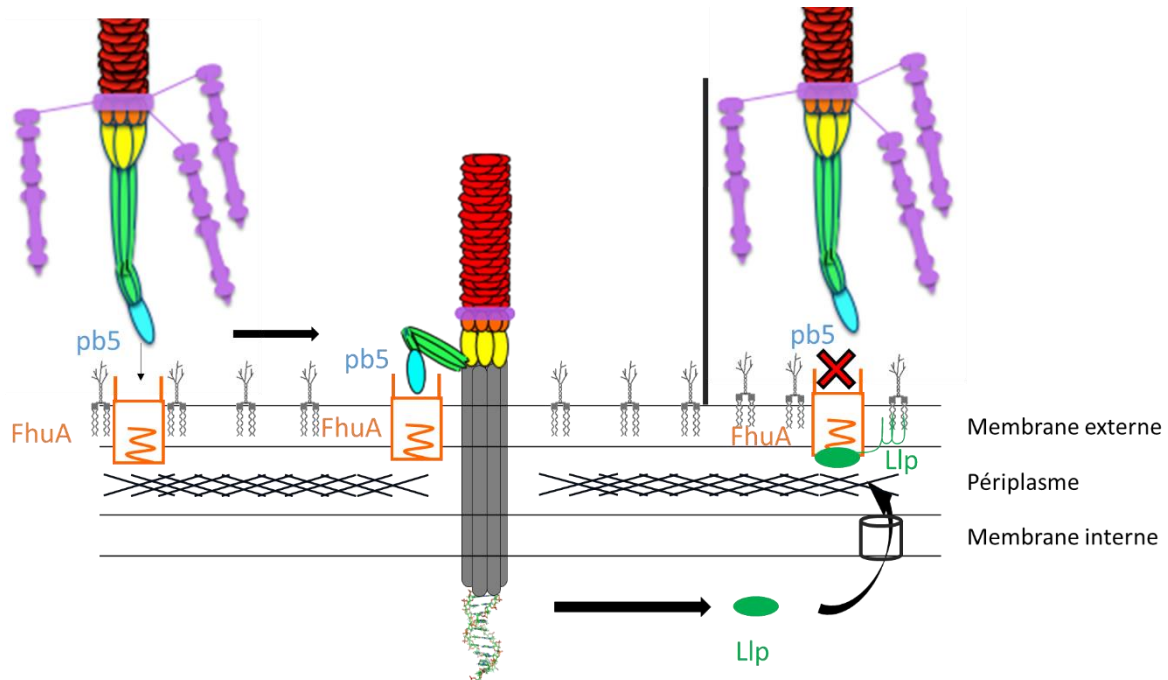
**Figure 1-11 : Disposition générale des gènes de queue de siphophages.**

Les régions codant pour les protéines de queue sont représentées pour les siphophages  $\lambda$  et T5 de *E. coli*, SPP1 de *B. subtilis*, TP901-1 et p2 de *L. lactis*, 80 $\alpha$  de *S. aureus*, ainsi que pour le myophage Mu de *E. coli*. Les gènes codant pour des protéines de fonction/structure homologues sont représentés dans la même couleur. Les abréviations et le code couleur sont les mêmes qu'en figure 9. Les homologies ont été établies par des données de structure et de séquence par HHpred. Figure de 48.



### 1.1.6.3 Cycle d'infection

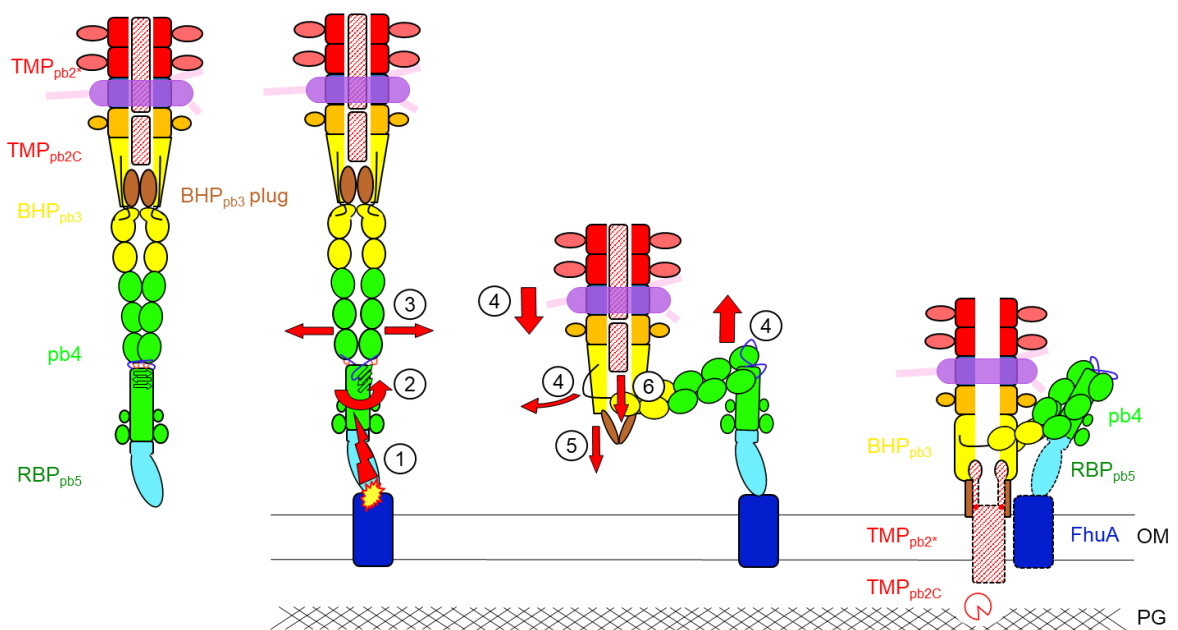
L'infection de l'hôte se déroule en 4 étapes : le phage se lie d'abord de manière réversible aux lipopolysaccharides à la surface cellulaire grâce à pb1 qui compose les fibres en forme de L. Cette interaction réversible permet au phage de sonder la surface bactérienne et facilite la liaison de pb5 à FhuA<sup>49</sup>. L'interaction entre pb5 et FhuA est nécessaire et suffisante pour déclencher la perforation de la paroi bactérienne et l'injection de l'ADN dans le cytoplasme<sup>44</sup> (Figure 1-12). Le phage T5 injecte son ADN en deux étapes : environ 8% de l'ADN correspondant au FST (pour « first step transfer ») sont injectés. Après une pose de quelques minutes, où des gènes contrôlant les fonctions cellulaires de l'hôte inhibent les systèmes de défense de la bactérie, le reste de l'ADN est injecté<sup>50</sup>. Cette deuxième étape de l'infection est suivie par la réplication du génome, la synthèse des protéines structurales, l'assemblage de nouveaux virions puis leur libération par la lyse de la bactérie. Pendant la réplication et l'assemblage, T5 protège la nouvelle usine virale d'une surinfection par la production de la protéine Llp (Sie)<sup>42</sup>. Llp est produite dans le cytoplasme puis transportée dans la membrane externe d'*E. coli*, puis se lie à FhuA. Cette interaction empêche toute nouvelle liaison de pb5 mais aussi le transport de tous les ligands de FhuA<sup>42,51</sup> (figure 1-12). A la fin du cycle viral qui dure environ 25 à 30 min, Llp est toujours présente et permet d'empêcher l'inactivation des virions par FhuA présente dans les débris des cellules lysées, augmentant ainsi leurs chances d'infecter un nouvel hôte<sup>51</sup> (Figure 1-12).



**Figure 1-12 : Différentes étapes de l'infection par le phage T5.**

Les fibres en L (violet) permettent de sonder la surface bactérienne en interagissant avec le LPS. La liaison de pb5 (bleu) à FhuA (rouge) permet l'injection d'ADN dans le cytoplasme de la bactérie. Llp (vert) codé par le phage T5, est produit et transporté par le système bactérien. Llp interagit avec FhuA prévenant toute nouvelle liaison.

Le mécanisme permettant la perforation de la paroi cellulaire a été élucidé récemment<sup>47</sup>. La détermination de la structure du tube de la queue T5 par cryo-microscopie électronique (Cryo-EM) suggère que le tube de la queue (pb6) ne subit pas de changement de conformation lors de la liaison de l'extrémité de la queue (pb5) à son récepteur FhuA<sup>52</sup>. En revanche, la liaison de pb5 à FhuA produit un changement de conformation de pb5 jusque-là inconnu (Figure 12 (1)) qui induit une torsion de l'hélice bêta de pb4 (Figure 1-13), tirant sur la liaison entre l'hélice bêta et les domaine FNIII de pb4. Ceci à son tour déstabilise le réseau de FNIII de la fibre centrale (3). La réorganisation de la chaîne FNIII autour de l'hélice bêta de pb4 induit la flexion de pb4 (2), rapprochant le tube de la membrane et désengageant le lien de pb3 ce qui libère le bouchon (former par la partie C-terminal de pb3) permettant d'ouvrir le tube (4,5). Ce dernier événement déstabilise la partie C-terminale de pb2, qui est expulsée (6). La partie C-terminale de pb2, serait transloquée au périplasm où elle digèrerait localement le peptidoglycane. L'autre partie de pb2 pourrait former un canal transmembranaire en intégrant ses segments hydrophobes dans la membrane externe<sup>47</sup>. Ce canal protégerait l'ADN phagique contre la dégradation par les nucléases périplasmique s'il se prolonge dans le périplasm. Le mécanisme d'infection du phage T5 est à présent mieux compris, mais les changements de conformation de pb5 qui induise cette cascade d'évènements ne sont pas encore connus.



**Figure 1-13 : Mécanismes proposés pour le déclenchement de l'infection, l'ancrage à la membrane et la formation de canaux par le phage T5.**

Le code couleur est le même que dans la figure 10. Mécanisme d'ouverture de la queue du phage permettant le transfert de l'ADN à travers la paroi bactérienne. Etapes 1 à 6 expliquées dans le texte. Figure issue de 47.

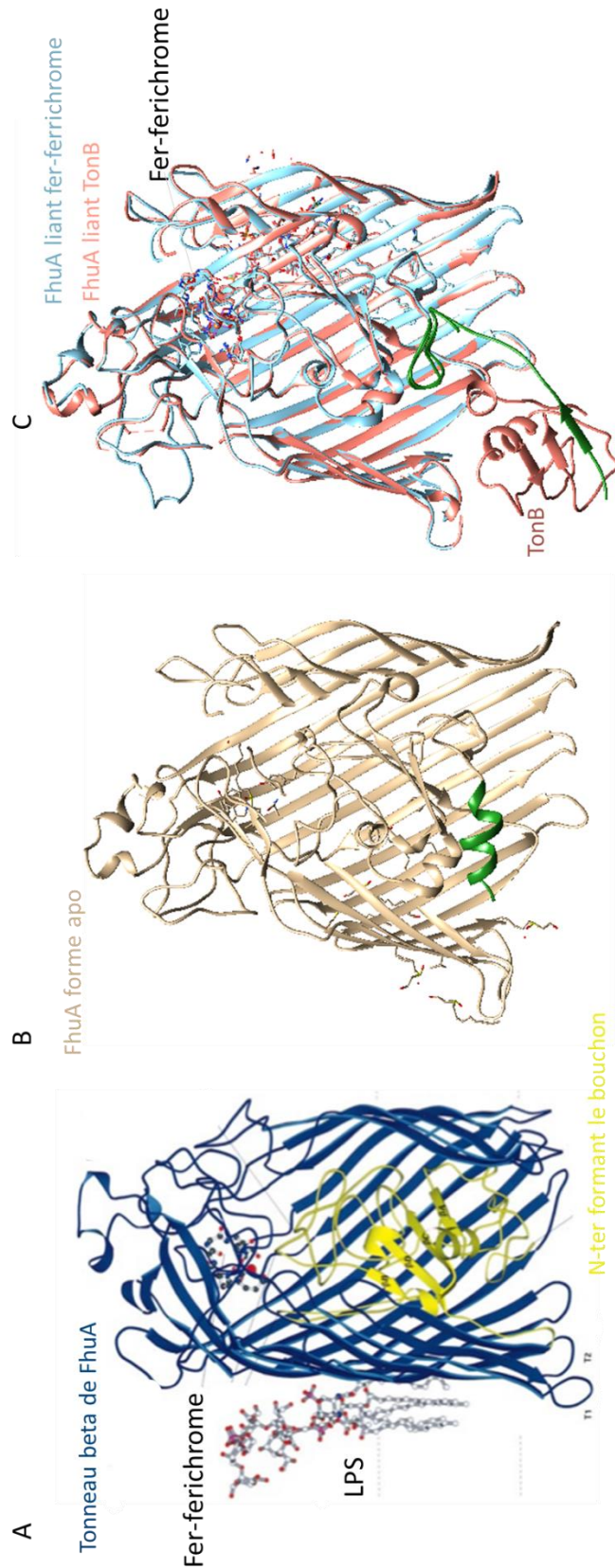
## 1.2 Présentation des protéines

### 1.2.1 La protéine FhuA

La Protéine FhuA a d'abord été appelée TonA car elle est découverte comme étant le récepteur du phage T1 en 1943. TonA a été la deuxième protéine pure isolée de l'enveloppe cellulaire d' *E. coli*<sup>39</sup>. Plus tard TonA a été renommée FhuA pour Ferric hydroxamate uptake protein A car son rôle physiologique est le transport du ferrichrome, un sidérophore complexant le fer du type hydroxamate. FhuA fait partie des transporteurs actif TonB dépendants (TBDTs) qui peuvent transporter des sidérophores, des vitamines, des complexes de nickel ou des sucres<sup>53</sup>. Mais FhuA est aussi capable de transporter la colicine M, la microcine J25, l'albomycine et la rifamycine qui sont des antibiotiques<sup>39</sup>. En plus d'être le récepteur des phages T1 et T5, FhuA l'est également pour les phages  $\phi$ 80, et UC 1. Pour l'ensemble des fonctions de FhuA (sauf l'infection par T5) la protéine FhuA a besoin d'énergie, apportée par le gradient électrochimique de la membrane interne par le complexe TonB-ExbB-ExbD<sup>54</sup>.

La structure de FhuA a été résolue en 1998<sup>55,56</sup>. FhuA est un monomère et se compose d'un tonneau de 22 brins  $\beta$  antiparallèles (Figure 13 en bleu), et contrairement aux porines, il contient un bouchon (en jaune) qui remplit la lumière du tonneau et permet le transport actif du ferrichrome. Le bouchon est formé par le domaine N-terminal globulaire qui pénètre dans le tonneau  $\beta$  du côté périplasmique et ferme hermétiquement le pore de la protéine. Les brins  $\beta$  sont reliés par de courtes boucles périplasmique et les boucles extracellulaires plus grandes. Comme beaucoup de protéines membranaires, FhuA contient deux ceintures de résidus aromatiques marquant les limites de la zone apolaire du tonneau beta<sup>54</sup>. Malgré la délipidation réalisée au cours de la purification une molécule de LPS reste associée à FhuA lors de la cristallisation et permet de déterminer la position de la membrane. Ceci montre que les boucles externes s'étendent bien au-dessus de la bicouche lipidique. La structure de FhuA a été résolue sans ligand ou avec différent ligand<sup>53</sup>. L'Ensemble des 12 structures de FhuA se superpose remarquablement bien avec un RMSD (Root-Mean-Square Deviation) de 0,5 sur 675 résidus (Figure 1-14). Les principales différences entre les structures se situent dans la TonB box qui est désordonnée dans toutes les structures, sauf quand le ferrichrome ou TonB sont présents.

FhuA est une protéine de 80 kDa et le tonneau et le bouchon se comportent comme 2 domaines indépendants présentant une température de dénaturation ( $T_m$ ) à 75°C et 65°C respectivement. Cependant le bouchon stabilise aussi le tonneau puisque dans FhuA $\Delta$ plug déletée pour le bouchon la dénaturation du tonneau arrive à 61.6°C<sup>57</sup>.



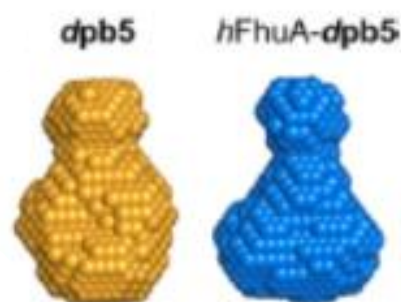
**Figure 1-14 : Structure de FhuA en complexe avec le fer ferrichrome.**

A. structure de FhuA vue de profil avec le tonneau en bleu et le bouchon (résidu 1-160) en jaune. Les résidus 621 à 723 ont été retirés pour permettre de visualiser le tonneau. Le LPS et le fer ferrichrome sont représentés en bâtonnés avec le fer représenté par une sphère rouge. Les deux ceintures aromatiques de FhuA sont indiquées par les pointillés. B. FhuA en forme Apo (sans ligant). C. superposition de FhuA liant au fer ferrichrome et au FhuA liant en plus TonB. Figure adaptée de 55.

### 1.2.2 La protéine pb5

La protéine pb5 a été identifiée comme une protéine RBP en 1984 grâce à un mutant du gène *oad* de T5 (Heller et al., 1984). Le gène de pb5 a été nommé *oad* (O-Antigen dependent phenotype) car la liaison d'un phage T5 muté dans ce gène est dépendante de l'antigène O du LPS. Dans Krauel and Heller 1991<sup>58</sup>, les auteurs ont utilisé le gène *oad* à la place du gène *hrs* (host receptor specificity) du phage BF23, ce qui a eu pour conséquence de changer la spécificité du récepteur passant de ButB à FhuA. Cette expérience confirme que la spécificité d'attachement à FhuA est due à pb5 dans le phage T5. Un peu plus tard la mutation qui a permis d'identifier pb5 comme la RBP a été identifiée comme une mutation ponctuelle G166T<sup>59</sup>. Dans Mondigler et al. 1996<sup>59</sup>, les auteurs ont cherché à trouver quelle zone est en interaction avec FhuA. Ainsi ils ont montré que la délétion des 153 acides aminés du côté C-terminal n'est pas indispensable à l'interaction puisque pb5 peut encore se lier à FhuA. De plus ils ont étudié les similarités de séquences entre pb5 du phage T5 et pb5 du phage BF23 et ont montré des zones conservées du côté N-Terminal, probablement en interaction avec le phage, et des zones non conservées probablement en interaction avec leur récepteur respectif (acide aminé 41-66, 89-305, 361-412, 461-545, et 562-628).

La protéine pb5 est une protéine de 640 acides aminés avec un pI basique (8.02) et peu soluble (0.5mg/ml) avec un Tm de 46°C. Pb5 migre en AUC (Analytical UltraCentrifugation) comme une protéine globulaire de 60 kDa (poids théorique 70 kDa) suggérant une protéine plutôt allongée. En revanche le profil d'élution de pb5 sur une chromatographie d'exclusion de taille migre comme une protéine de 40 kDa ce qui suggère que la protéine a des interactions aspécifiques avec la colonne<sup>60</sup>. Les expériences de dichroïsme circulaire montre que pb5 est composé de 51 % de feuillet, 6% d'hélice, 22 % de coude et 22 % de région non structurée<sup>60</sup>. Des expériences de SANS (Small-Angle Neutron Scattering) montrent une protéine allongée, qui ne subit pas de changements conformationnels importants lors de la liaison à FhuA (Figure 1-15)<sup>61</sup>.

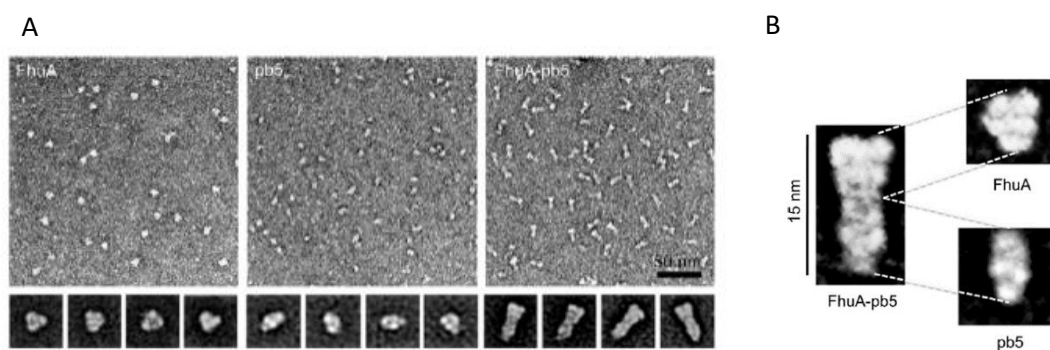


**Figure 1-15 : Structure des enveloppes de pb5 à partir des données SANS.**  
pb5 seul (orange) et pb5 en complexe FhuA (bleu). Figure issu de 61.

### 1.2.3 Le complexe FhuA-pb5

La fixation de pb5 à FhuA déclenche l'infection du phage T5. Obtenir la structure de ce complexe permettrait d'avoir la première étape du mécanisme de la perforation de la paroi de *E. coli* par T5 (voir paragraphe 1.1.6.3). De plus il est important de savoir où se lie pb5 pour comprendre l'inhibition de la fixation de pb5 à FhuA par Llp. Le complexe FhuA-pb5 peut être formé *in vitro* par un mélange équimolaire de FhuA et pb5 déjà purifiées. Ce complexe est extrêmement stable. En effet, il n'est pas dénaturé par le SDS (sodium dodecyl sulfate) sur gel en présence de  $\beta$ -mercaptoéthanol. Pour séparer le complexe sur gel SDS, il faut chauffer le complexe au moins 3 min à 70°C<sup>60</sup>. Des expériences de SPR (résonance plasmique de surface) ont été réalisées sur le complexe et le  $K_d$  (constante de dissociation) est trop faible pour être mesuré par cette méthode (il est inférieur à 100 pM) (Flayhan et al. 2012). De plus, pb5 et FhuA sont toutes deux stabilisées lorsqu'elles interagissent : la température de dénaturation ( $T_m$ ) du complexe est de 89°C alors que la  $T_m$  de pb5 est de 43°C et que FhuA présente deux  $T_m$  65°C (pour le bouchon) et 74°C (pour le tonneau béta). Une analyse par dichroïsme circulaire a été réalisée. Cette méthode permet de sonder l'abondance des différentes structures secondaires dans la protéine. Le résultat montre une différence entre la somme des spectres de pb5 et de FhuA issus des protéines seules et les spectres expérimentaux du complexe FhuA-pb5. Le complexe FhuA-pb5 contient plus de feuillets béta<sup>62</sup>. Pour explorer ces changements de conformation, les chercheurs ont obtenu la forme globale du complexe FhuA-pb5 grâce à des expériences de microscopie électronique en coloration négative et de SANS. Malgré les images recueillies, la forme globale ne donne pas d'information sur le changement de conformation<sup>61</sup> (Figure 1-16).

Le fer ferrichrome n'empêche pas la formation du complexe FhuA-pb5, bien qu'il entre en compétition avec la formation du complexe puisque qu'à 50  $\mu$ M de fer ferrichrome il y a un peu de FhuA et de pb5 libre<sup>60</sup>. Le  $K_d$  du fer ferrichrome étant de 200 nM<sup>63</sup>, cela implique que le  $K_d$  de pb5 pour FhuA est plus petit. La fixation de pb5 sur FhuA a été testée pour des FhuA délétés pour chacune des boucles. À part pour la boucle 8 les délétions n'empêchent pas la fixation<sup>64</sup>. En outre pb5 ne peut pas se fixer sur un FhuA $\Delta$ plug<sup>65</sup>.



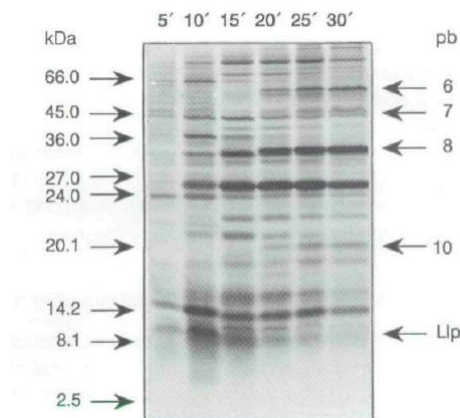
**Figure 1-16 : structure basse résolution de FhuA, pb5 et FhuA-pb5 en coloration négative.**

A. image de FhuA, pb5 où le complexe FhuA-pb5 en microscopie a une coloration négative. B. comparaison de la forme globale de FhuA isolé, pb5 isolé ou FhuA-pb5 en complexe. Figure adaptée de 61.



#### 1.2.4 La protéine Llp

La protéine Llp pour « Late lipoprotein » a été décrite par Decker et al. en 1994<sup>42</sup>. Llp est codée par le gène *llp* situé juste à côté du gène *oad* de pb5 (voir ci-dessus). Les études montrent que le promoteur de Llp est un promoteur précoce qui chevauche le promoteur *oad* tardif qui lui est orienté de sens opposé. Une expérience de « pulse chase », où l'expression des protéines est suivie par l'insertion d'acides aminés radioactifs, montre que Llp est produite dès 5 min après l'infection et que sa production est maximum à 10 min après l'infection. Pour comparaison, la production des protéines de structure comme pb6 ou pb10 ne commence qu'à 20 min après l'infection<sup>42</sup> (Figure 1-17). Les bactéries portant le plasmide contenant Llp sont résistantes à T5 et il n'y a pas d'absorption suggérant qu'elle agit au niveau de la surface bactérienne. Les auteurs font donc l'hypothèse que Llp est une protéine Sie et la renomme « lytic conversion protéine »<sup>42</sup>.

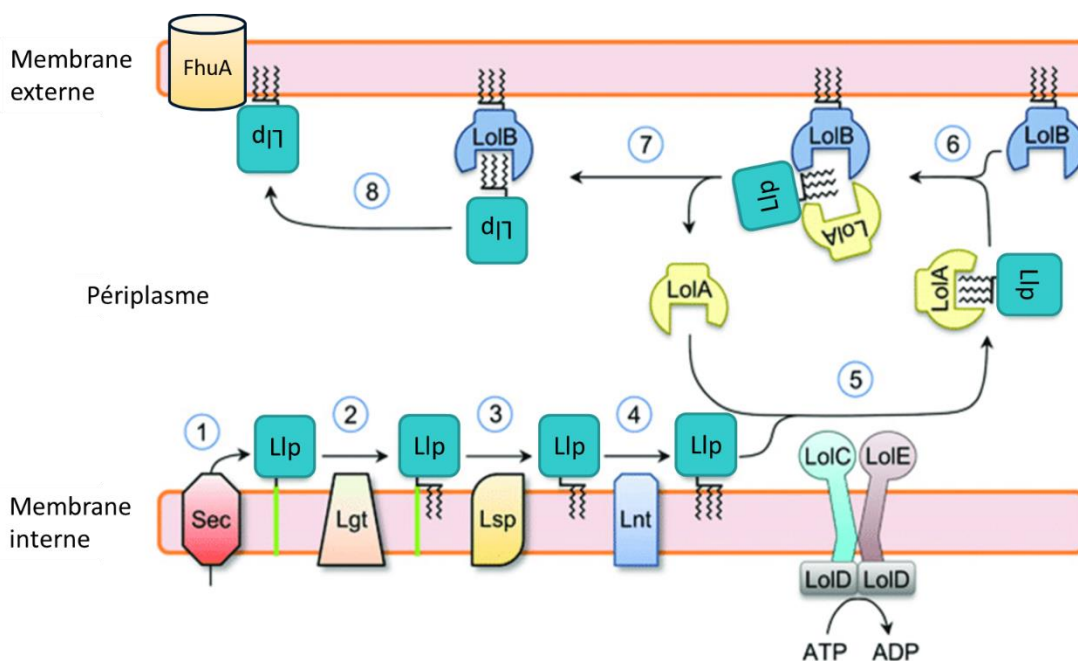


**Figure 1-17 : Expression des protéines du phage T5 pendant l'infection.**

Des échantillons prélevés à 5, 10, 15, 20, 25 et 30 min après l'infection de cellules d'*E. coli* radiomarquées et infectées par T5 ont été analysées sur gel d'acrylamide à 15 % et détectées par autoradiographie. Dans la marge de droite sont indiquées les positions de certaines protéines de structure (indiquées comme des bandes protéiques (pb + numérotation)) et de Llp. L'échelle de marqueurs de masse moléculaire des protéines est indiquée à gauche en kDa. Figure issue de 41.

La même année Braun et al.<sup>66</sup> identifient Llp comme une lipoprotéine qui interagit avec FhuA. Ils confirment que Llp empêche la liaison de pb5, mais inhibe également le fer ferrichrome et le transport de la colicine M. En revanche la sensibilité à la colicine B reste inchangée montrant que Llp n'est pas capable d'interagir avec FepA<sup>66</sup>. Dans la même étude les auteurs étudient l'interaction de Llp avec différents mutants de FhuA. Certains mutants du côté périplasmique mais aussi extracellulaire ne sont plus inhibés par Llp, suggérant des changements de conformation à longue distance. En revanche, toutes les boucles périplasmiques ne sont pas nécessaires pour que FhuA soit inactivé par Llp<sup>66</sup>. En 1998, Pedruzzi et al.<sup>51</sup> ont produit la protéine Llp sous une forme acylée et soluble. Ils ont montré, par dichroïsme circulaire, que la protéine de 7,8 kDa est formée principalement de feuilletés bêta par dichroïsme circulaire et qu'elle est capable d'inhiber FhuA in vitro<sup>51</sup>.

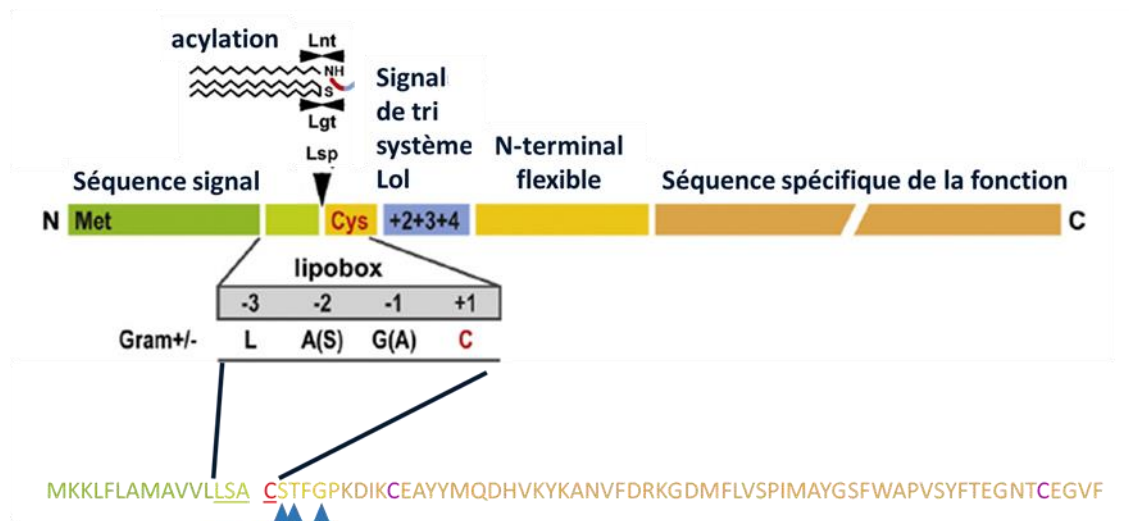
Braun et al. ont suggéré que Llp est une lipoprotéine adressée à la membrane externe pour interagir avec FhuA. La protéine Llp est produite sous forme d'un précurseur : elle possède une séquence signal d'adressage au périplasma et une lipobox, séquence reconnue par la machinerie de l'hôte permettant d'acyler la Cystéine N-terminale <sup>67</sup>(Figure 1-18). La séquence signal de Llp est reconnue par le système de transport Sec (Figure 1-18 (1)) pour être transportée dans le périplasma et prise en charge par 3 enzymes successives. La phosphatidylglycerol-prolipoprotein diacylglyceryl transferase (Lgt) catalyse l'ajout d'un diacylglycérol sur le thiol de la cystéine +1 de Llp (2), puis la lipoprotein signal peptidase (Lsp) coupe la séquence signal (3), et enfin la protéine apolipoprotein n-acyltransferase (Lnt) catalyse l'acylation de la cystéine +1 sur l'amine terminal de la cystéine +1 (4). La protéine Llp est adressée à la membrane externe par une séquence signal reconnue par le système Lol ABCDE. Le complexe Lol CDE est un transporteur ABC et apporte l'énergie pour transférer Llp de la membrane interne à la chaperonne périplasmique Lol A (5) ce qui permet à Llp de traverser le périplasma avant d'être accepté par Lol B (6-7). Enfin, Llp est insérée dans le feuillet interne de la membrane externe où elle va pouvoir interagir avec FhuA (8)<sup>68</sup>.



**Figure 1-18 : Mécanisme de transport de Llp par la bactérie.**  
*Numérotation de chaque étape détaillée dans le texte. Schéma adapté de 67.*



L'adressage à la membrane interne ou externe des lipoprotéines est contrôlé par la nature de l'acide aminé +2 de la protéine<sup>69</sup>. Une serine en position +2 permet d'adresser la protéine à la membrane externe alors qu'un acide aminé aromatique, une proline ou un aspartate permet la rétention dans la membrane interne. Llp possède une serine en position +2, ainsi, elle est transportée à la membrane externe<sup>70</sup> (Figure 1-19). Llp a été utilisée pour étudier le mécanisme du tri d'adressage à la membrane interne ou externe. Dans Robichon et al, les auteurs prouvent que Llp possède un pont disulfure essentiel à sa fonction et qui influe sur la rétention ou non de Llp dans la membrane interne<sup>71</sup>.



**Figure 1-19 : Structure générale des domaines des lipoprotéines (haut) et séquence du précurseur de la protéine Llp (bas).** Séquence signal N-terminale en vert, lipobox : signal pour une di-acylation en vert clair et cystéine acylée en rouge, système de tri pour le système Lol ABCDE en bleu, partie N-terminale flexible en jaune et séquence spécifique de la protéine en orange. Dans la séquence de Llp les 2 autres cystéines formant un pont disulfure sont colorées en violet. Figure adaptée de 69.

## 2 Matériels et méthodes

## 2.1 Biologie moléculaire

### 2.1.1 Clonage des différentes formes de Llp

Les plasmides pET20b (Ac-Llp) et pMAL-p2E (Sol-Llp) ont été utilisés pour l'expression de Llp. Les deux plasmides utilisés portent la résistance à l'ampicilline pour la sélection des clones. Le gène de la protéine sous forme acylée, a également été introduit dans des vecteurs pASK7+ (expression de souche dépourvue de polymérase T7 avec une résistance à amplicilline) et pASKIAB3C (expression de souche dépourvue de polymérase T7 avec une résistance au chloramphénicol). Enfin des clonages pour produire des protéines de fusion FhuA-Llp sans lien entre FhuA et Llp, ou avec un lien de 5 ou 10 acides aminés ont été produits avec les séquences GGGGSF et GGGSGGGGSF respectivement pour chacun des liens. Il a été choisi d'insérer Llp à la suite de FhuA dans son plasmide d'origine qui est donc le vecteur d'expression de la protéine de fusion<sup>72 et 73</sup>.

Les vecteurs d'expression de Llp acylée ou non et des fusions FhuA-Llp sont réalisés grâce à une méthode de clonage par double PCR. La première PCR permet, à partir de l'ADN génomique de T5, d'amplifier un fragment linéaire d'ADN qui va être inséré dans le plasmide de destination par une deuxième PCR (voir Tableau 2-3 pour les amorces utilisées). C'est aussi à cette étape que l'on ajoute, au niveau des amorces, les séquences additionnelles comme les étiquettes ou les sites de clivage.

La première PCR est réalisée avec environ  $6.10^9$  particules de phage, ou 100 ng de plasmide pCHAP6507 (CD-Llp)<sup>71</sup>, 1X tampon  $MgSO_4$ , 250 ng de chaque couple d'amorces, 10 nM de dNTP et 2 unités de polymérase haute-fidélité (F530L). Cycle de PCR suivant :

Température	Temps	Répétition
95°C	3 min	1
95°C	45 s	25
57 ou 62 °C	1 min	
72 °C	40s	
71 °C	10 min	1

*Tableau 2-1 : Cycles de PCR1 utilisés pour amplifier les séquences du gène llp.*

L'insertion du fragment amplifié dans le vecteur de choix est réalisée à l'aide du QuickChange Lightning Multi Site-directed Mutagenesis Kit (Agilent Technologies), utilisé selon les recommandations du constructeur. Les amorces utilisées à cette étape sont les produits de PCR de la première étape. Les produits de PCR sont purifiés par migration sur gel d'agarose 1,2 %, TAE 1X 90V 50 min, puis une extraction à l'aide du kit NucleoSpin Gel and PCR Clean-up (REF 740609.50) selon les recommandations du constructeur. Le cycle de PCR utilisé pour l'insertion de la séquence dans le vecteur est le suivant :

Température	Temps	Répétition	étape
95°C	2 min	1	Dénaturation initiale
95°C	1 min	22	Dénaturation
55 °C	1 min		Hybridation
68 °C	30 sec/kb		Élongation
68 °C	10 min	1	Élongation finale

construction	Temps d'élongation
Ac-Llp	2 min
MalE-Llp	4 min
FhuA-Lx-Llp	4 min

*Tableau 2-2 : Cycle de PCR2 utilisé pour insérer les gènes dans le plasmide de sur expression.*

La digestion des plasmides parentaux par la Dnpi, 5 min à 37°C, permet de transformer les bactéries uniquement avec les plasmides contenant l'insert issu de la deuxième PCR. La présence de l'insert des plasmides obtenus est contrôlée par digestion enzymatique (BamHI), puis par séquençage (GENEWIZ).

Les étapes de PCR et la digestion des plasmides sont contrôlées par migration sur gel d'agarose 0,8 %, migration dans du tampon Tirs Acetate EDTA TAE 1X 90V 50 min.

Les constructions sont ensuite amplifiées dans une souche 10XGold et purifiées par midi Prep en utilisant le kit NucleoBBind Xtra Midi/Maxi. Les plasmides sont stockés à -20°C.

Les différentes souches utilisées dans cette étude ont toutes été transformées par la même méthode. Entre 100 et 200 ng de plasmide sont incubés 30 minutes sur glace avec les souches à transformer (*BL21DE3*, *C43*, *AW740* ou *Salmonella enterica indica*). Puis un choc thermique de 40 secondes à 42 °C est effectué. Les bactéries sont ensuite incubées 2 minutes sur glace puis 100 µL de SOC (Dextrose : 3,603 g, KCL : 0,186 g, MgSO4 : 4,8 g, Tryptone : 20 g, Extrait de levure : 5g, KCL : 0,186g /L) est ajouté. Les bactéries sont incubées 1 h à 37°C, 180 rpm, puis isolées sur une boîte LB agar avec l'antibiotique approprié, ampicilline 125 µg/ml pour pET20b-Ac-Llp, pMAL-p2E (Sol-Llp), pASK7+-Ac-Llp et pH405X-FhuA-Llp, ou Chloramphénicol 125 µg/ml pour pASKIAB3C-Ac-Llp.

Nom de la construction	Source d'ADN	Amorces utilisées
Ac-Llp	Phage T5 sauvage	<b>Amorce directe :</b> TTTAAGAAGGAGATATACATATGAAAAAATTATTTTAGCTATGG <b>Amorce indirecte :</b> CCAGCAGCAGCGGTCCGGCTCAGTGATGGTGATGGTGATGGCCGAAAAC CCCTCGCATGTATTC
Sol-Llp	Phage T5 sauvage	<b>Amorce directe :</b> TTTAAGAAGGAGATATACATATGTCTACTTTTGGACCTAAAG <b>Amorce indirecte :</b> CAGCAGCAGCGGTCCGGCTCAGTGATGGTGATGGTGATGGCCGAAAAC CCTCGCATGTATTC
MalE-Llp	Phage T5 sauvage	<b>Amorce directe :</b> CAATAACAACAACCTCGGGGAAAACCTGTATTTTCAGGGCTCTACTTTTG GACCTAAAGAG <b>Amorce indirecte :</b> GATTTAATCTGTATCAGGCTGTTAGAAAACCTCCCTCGCATGTATTC
MalE-his-Llp	Plasmide MalE-Llp	<b>Amorce directe :</b> CAATAACAACAACCTCGGGCATCACCATCATCACCATCACCACGAAAACC TGATTTTTCAGGG <b>Amorce indirecte :</b> CCCTGAAAATACAGGTTTTCTGGTGATGGTGATGGTGATGCCGA GGTTGTTGTTATTG
Opéron Llp	Phage T5 sauvage	Étape 1 récupérée la séquence de l'opéron <b>Amorce directe :</b> TTTAAGAAGGAGATATACATATGAAAAAATTATTTTAGCTATGG <b>Amorce indirecte :</b>

		<p>CAGCAGCAGCGGTCGGCTTAGCGCATAGAATGGTCATC</p> <p>Étape 2 insérés l'étiquette histidine au niveau de Llp</p> <p><b>Amorce directe :</b> GAAATACATGCGAGGGAGTTTTTCGGCCATCACCATCACCATCACTAAGTA AAAACCACCTTAGTAACC</p> <p><b>Amorce indirecte :</b> GGTTACTAAGGTGGTTTTTACTTAGTGATGGTGGTGGTGGCCGAAA ACTCCCTCGCATGTATTC</p>
PuIA-Ac-Llp	pCHAP6507	<p>Étape 1 récupérée la séquence de Llp pour l'inséré dans pET20b</p> <p><b>Amorce directe :</b> TTTAAGAAGGAGATATACATATGCTCAGATATACCCGTAATG</p> <p># Amorce 1 : Tm = 62°C et NT = 22 / Amorce 2 : Tm = 58°C et NT = 23</p> <p><b>Amorce indirecte :</b> CCAGCAGCAGCGGTCGGCTCAGTGATGGTGGTGGTGGCCGAAA CCCTCGCATGTATTC</p> <p>Étape 2 mutations dans la séquence d'adressage de D en S</p> <p><b>Amorce directe :</b> GGTTTTACTTAGCGGCTGCAGCAACAGCGCCGAATTCTCTAC</p> <p><b>Amorce indirecte :</b> GTAGAGAATTCGGCGCTGTTGCTGCAGCCGCTAAGTAAAACC</p>
Ac-Llp-pASK7	Ac-Llp	<p><b>Amorce directe :</b> CTTAAGAAGGAGATATACAAATGAAAAATTATTTTTAGCTATGG</p> <p><b>Amorce indirecte :</b> CTTCACAGGTCAAGCTTAGTTAGTGATGGTGGTGGTGGC</p>
Ac-Llp-pASKIAB3C	Ac-Llp	<p><b>Amorce directe :</b> TAGATAACGAGGGCAAAAAATGAAAAATTATTTTTAGCTATGG</p> <p><b>Amorce indirecte :</b> GTGGCTCCAAGCGCTGAGGAAAACCCCTCGCATGTATTC</p>
FhuA-L0-Llp	Ac-Llp & pWSK29-FhuA	<p><b>Amorce directe :</b> CAACCGCAACCTTCCGTTTCTTCTACTTTTGGACCTAAAGATAT</p> <p><b>Amorce indirecte :</b> CGTGCCCCAAAAGAGAAATTATTTTTCGAACTGCGGGTGGC</p>
FhuA-L5-Llp	Ac-Llp & pHX405 (FhuA)	<p><b>Amorce directe :</b> CAACCGCAACCTTCCGTTTGGCGGTGGTGGCAGCTTTTCTACTTTTGGC CTAAAGATAT</p> <p><b>Amorce indirecte :</b> CGTGCCCCAAAAGAGAAATTATTTTTCGAACTGCGGGTGGC</p>
FhuA-L10-Llp	Ac-Llp & pHX405 (FhuA)	<p><b>Amorce directe :</b> CAACCGCAACCTTCCGTTTGGCGGTGGTGGCAGCGGTGGCGCGGTAG CTTTTCTACTTTTGGACCTAAAGATAT</p> <p><b>Amorce indirecte :</b> CGTGCCCCAAAAGAGAAATTATTTTTCGAACTGCGGGTGGC</p>

**Tableau 2-3 : Amorges utilisées pour le clonage des différentes constructions.**

Rouge : codons Start et Stop, vert : séquences s'hybridant sur llp, bleu foncé : étiquette histidine, jaune : glycine additionnelle, rose : séquences s'hybridant sur MalE, violet : séquences s'hybridant sur Orf 4 de l'opéron de llp, bleu clair : mutation unique, marron : séquence de l'étiquette streptavidine, orange : acides aminés de liaison, noir : séquences s'hybridant sur les plasmides de destination.

### 2.1.2 Mutagenèse dirigée :

Dans cette étude, des mutants de FhuA et de Ac-Llp ont été produits pour analyser les acides aminés essentiels à la protection de la bactérie au bactériophage T5 du côté de FhuA ou de Llp. La mutagenèse dirigée a été réalisée à partir des plasmides décrits dans la section précédente et à l'aide du QuickChange Lightning Multi Site-directed Mutagenesis Kit (Agilent Technologies), utilisé selon les recommandations du constructeur. Les amorces utilisées pour réaliser ces mutants sont décrites dans le tableau 2-5 pour FhuA et 2-6 pour Ac-Llp. La digestion des plasmides parentaux est effectuée par la Dnpi, 5 min à 37°C. La présence de la mutation est contrôlée par séquençage (GENEWIZ).

Le cycle de PCR utilisé pour l'insertion des codons mutés dans le vecteur est le suivant :

Température	Temps	Répétition	étape		
95°C	2 min	1	Dénaturation initiale		
95°C	20 sec	22	Dénaturation	construction	Temps d'élongation
60 °C	10 sec		Hybridation		
68 °C	30 sec/kb		Élongation		
68 °C	10 min	1	Élongation finale	FhuA-Lx-Llp	4 min

*Tableau 2-4 : Cycle de PCR utilisé pour insérer les codons mutants.*

Nom de la construction	Source d'ADN	Amorce utilisée
FhuA-APAΔ333-335	pHX405 plasmide	<b>Amorce directe :</b> GCAAACAGTGTGCGGCATTAXGATAAAGGCCATTATCTGGCA <b>Amorce indirecte :</b> TGCCAGATAATGGCCTTTATCXTAATGCCGCACACTGTTTGC
FhuA-D292A		<b>Amorce directe :</b> GCTTCGATCACGAATTTAACGCGACCTTTACTGTGCGTCAGAAC <b>Amorce indirecte :</b> GTTCTGACGCACAGTAAAGGTCGCGTTAAATTCGTGATCGAAGC
FhuA-D369D371-AA		<b>Amorce directe :</b> GAGCAAGTTTGCCACTGGCGCGATTGCCACACCCTGCTGACCGG <b>Amorce indirecte :</b> CCGGTCAGCAGGGTGTGTCAATATCGCCAGTGGCAAACCTTGCTC
FhuA- D395Y		<b>Amorce directe :</b> CGCCTGGTTTGGTTACGACTATTCTGTGCCACTGCTCAATCTG <b>Amorce indirecte :</b> CAGATTGAGCAGTGGCACAGAATAGTCGTAACCAAACCAGGCG
FhuA-DDK224-226-NNQ		<b>Amorce directe :</b> GCGTTCACCTGGCGTCCGAAACAACCAGACCAATTTTACCTTCCTTTCTTAC <b>Amorce indirecte :</b> GTAAGAAAGGAAGGTAATAATTGGTCTGGTTGTTCCGGACGCCAGGTGAACGC
FhuA-I9P		<b>Amorce directe :</b> GTTGAACCGAAAGAAGACACTCCGACCGTTACCGCTGCACCTG <b>Amorce indirecte :</b> CAGGTGCAGCGGTAACCGGTCGGAGTGTCTTCTTTCCGGTTCAAC
FhuA-L106P		<b>Amorce directe :</b> GCCAAAGCCAGAATAACTATCCGAATGGCCTGAAGTTGCAG <b>Amorce indirecte :</b> CTGCAACTTCAGGCCATTCCGGATAGTTATTCTGGCTTTGGC
FhuA-PANΔ321-323		<b>Amorce directe :</b> GTTACGGCGTCTGCTCCGATXGCTTACAGCAAACAGTGTGC <b>Amorce indirecte :</b> GCACACTGTTTGCTGTAAGCXATCGGAGCAGACGCCGTAAC
FhuA-T12P		<b>Amorce directe :</b> GAAAGAAGACACATCCACCGTTCCGGCTGCACCTGCGCCGC <b>Amorce indirecte :</b> GCGGCGCAGGTGCAGCCGGAACGGTGGATGTGTCTTCTTTC

**Tableau 2-5 : Amorces utilisées pour les mutations de FhuA.**

En rouge sont notés les codons mutés. Les délétions sont notées par un x rouge, et en noir les séquences s'hybridant sur les plasmides d'origine.

Nom de la construction	Source d'ADN	Amorce utilisée
Ac-Llp_D28K	Ac-Llp-pET20b	<b>Amorce directe :</b> GTGAAATATAAGGCTAATGTTTTTAAAAGGAAAGGGGATATGTTTTTG <b>Amorce indirecte :</b> CAAAAACATATCCCCTTTCTTTAAAAACATTAGCCTTATATTTTAC
Ac-Llp_KD7-8AR		<b>Amorce directe :</b> GGCTTGCTCTACTTTTGGACCTGCACGCATTAATGTGAGGCATACTATA TG <b>Amorce indirecte :</b> CATATAGTATGCCTCACATTTAATGCGTGCAGGTCCAAAAGTAGAGCAAG CC
Ac-Llp_RK28-30SS		<b>Amorce directe :</b> GAAATATAAGGCTAATGTTTTTGATAGCAGCGGGGATATGTTTTTGGTGT C <b>Amorce indirecte :</b> GACACCAAAAACATATCCCCGCTGCTATCAAAAACATTAGCCTTATATTTT
Ac-Llp_YK22-23AA		<b>Amorce directe :</b> CTATATGCAGGATCATGTGAAA GCGAGCGCTAATGTTTTTGATAGGAAA G <b>Amorce indirecte :</b> CTTTCCTATCAAAAACATTAGCGCTCGCTTTTACATGATCCTGCATATAG
Ac-Llp_YYHQD14-18AFK		<b>Amorce directe :</b> CTAAAGATATTAATGTGAGGCAGCATTTCGAGCAAAACATGTGAAATA TAAGGCTAATG <b>Amorce indirecte :</b> CATATAGTATGCCTCACATTTAATTCTTGCAGGTCCAAAAGTAGAGCAAG C
Ac-Llp_D18K21K23NNQ		<b>Amorce directe :</b> CTAAAGATATTAATGTGAGGCAGCATTTCGAGCAAAACATGTGAAATA TAAGGCTAATG <b>Amorce indirecte :</b> CATATAGTATGCCTCACATTTAATTCTTGCAGGTCCAAAAGTAGAGCAAG CC
Ac-Llp_FTE52-54YKK		<b>Amorce directe :</b> GGGCCCCGTAAGCTACTATAAAAAAGGAAATACATGCGAGGGAG <b>Amorce indirecte :</b> CTCCCTCGCATGTATTTCTTTTTATAGTAGCTTACGGGGGCC
Ac-Llp_W46S		<b>Amorce directe :</b> CAATTATGGCCTATGGATCTTTCAGCGCCCCGTAAGCTACTTTAC <b>Amorce indirecte :</b> GTAAAGTAGCTTACGGGGGCGCTGAAAGATCCATAGGCCATAATTG
Ac-Llp_YGSFW42-46STRAS		<b>Amorce directe :</b> GGTGTACCAATTATGGCCAGCACCCGCGCCAGCGCCCCGTAAGCTACT TTAC <b>Amorce indirecte :</b> GTAAAGTAGCTTACGGGGGCGCTGGCGGGGTGCTGGCCATAATTGGT GACACC

**Tableau 2-6 : Amorces utilisées pour les mutations de FhuA.**

En rouge sont notés les codons mutés. Les délétions sont notées par un x rouge, et en noir les séquences s'hybridant sur les plasmides d'origine.



### 2.1.3 Cellule compétente pour la biologie moléculaire :

Les souches *AW740* et *Salmonella enterica indica* sont rendus chimiocompétentes comme décrits dans Sambrook et al 1989. Les souches d'amplification Top10 et XL10Gold sont des souches chimiocompétentes commerciales. Les souches d'expressions BL21(DE3) et C43(DE3) sont des souches chimiocompétentes commerciales.

## 2.2 Tests de surexpression :

Des tests d'expression ont été menés sur 3 constructions de Llp dans différentes souches et conditions de culture, ainsi que sur une construction de la fusion FhuA-L10-Llp (voir Tableau 2-7).

Nom de la construction	Souche testée	Conditions de culture et induction
Ac-Llp	BL21(DE3)	Milieu : LB, TB Température : 20°C, 30°C et 37°C induction (1 mM IPTG) ou non
	C43(DE3)	
Llp Soluble	BL21(DE3)	
	Shuffle(DE3)	
MalE-Llp MalE-his-Llp	BL21(DE3)	
FhuA-L10-Llp	AW740	Milieu : LB Température : 37°C induction Bipyridyl

Tableau 2-7 : Conditions d'expression testées.

La souche BL21(DE3) est utilisée dans les tests d'expression, car il s'agit de la souche usuelle utilisée au laboratoire. Pour la construction Ac-Llp, la souche C43(DE3) a été testée, car cette souche permet la production de protéine toxique<sup>74</sup>. Pour la construction Llp Soluble sans séquence signal, la souche Shuffle a été testée, car elle permet la production de pont disulfure dans le cytoplasme. Pour la protéine de fusion FhuA-L10-Llp, la souche AW740 déplétée des principales porines de la membrane externe et pour FhuA endogène est utilisée. Pour contrôler le niveau d'expression, des gels d'acrylamide 15 % (Sol-Llp, Ac-Llp) ou 10% (FhuA-L10-Llp) colorés au bleu de Coomassie ou immunorévélés ont été réalisés. Les anticorps anti histidine ou anti MBP (Maltose-binding protein) ont été utilisés pour les immuno-révélation s (voir paragraphe 2.4.2).

## 2.3 Production et purification des protéines

L'ensemble des purifications présenté si après sont effectuées à l'aide du dispositif BioRad BioLogic DuoFlow Medium-Pressure Chromatography Systems équipé d'un détecteur multi longueur d'onde BioLogic QuadTec™ Detector UV-Visible et d'un collecteur de fraction BioFrac™ Fraction Collector.

### 2.3.1 Llp acylée

#### 2.3.1.1 Surexpression

Llp acylée est surexprimée dans une souche d'*E. coli* C43 transformée avec le plasmide Ac-Llp. Une préculture est réalisée en LB et en présence d'ampicilline 125 µg/ml, incubés à 37°C, 180 rpm, sur une nuit. Les surexpressions sont réalisées dans 1 L de LB, en présence d'ampicilline 125 µg/ml, inoculés

à 0,1 de densité optique à 600 nm avec la préculture et incubés à 20°C, 180 rpm, pendant 24 heures. En fin de culture, les bactéries sont récupérées par une centrifugation de 20 minutes, 3060 g, à 4 °C. Le culot est congelé dans l'azote liquide et conservé à -80 °C avant son utilisation.

#### 2.3.1.2 Solubilisation

Les cellules bactériennes sont décongelées dans du tampon de lyse à température ambiante (50 mM Tris pH 8 ; 150 mM NaCl ; 2 mM MgCl<sub>2</sub> ; DNase ; Clapa 1X (inhibiteurs de protéase)) avec un ratio de 10 g de cellules pour 30 ml de tampon. Les bactéries sont lysées à l'aide de 10 passages à 14000 PSI dans un microfluidizer. Les cellules non brisées et les gros débris cellulaires sont éliminés par une centrifugation de 15 minutes à 6900 g, 4 °C. Les membranes sont récupérées par une ultra centrifugation de 20 minutes à 100000 g, 4°C. Le culot est repris dans un tampon de stockage (25 mM Tris pH 8,5, 150 mM NaCl) à une concentration de protéine totale déterminée par BCA (BicinChoninic Acid ) de 20 mg/ml. Les membranes totales sont diluées 10 fois et incubées, soit 30 minutes à 37 °C avec agitation, soit toute une nuit à 4°C avec agitation, avec du tampon de solubilisation des membranes (25 mM Tris pH 8,5 ; 150 mM NaCl ;163 mM OPOE). Le matériel non solubilisé est éliminé à l'aide d'une ultracentrifugation dans les mêmes conditions que précédemment.

#### 2.3.1.3 Purification

Le surnageant contenant Llp solubilisée est chargé sur une colonne de 5 ml HisTrap Chelating Nickel column (GE Healthcare), pré-équilibrée avec un tampon à 25 mM Tris pH 8,5 ; 150 mM NaCl ; 12,2 mM OPOE (octyl polyoxyethylene). Un échange de détergent est réalisé par lavage avec un tampon contenant 25 mM Tris pH 8,5 ; 150 mM NaCl ; 19,6 mM DDM (Dodécyl-β-D-maltoside). Un lavage est ensuite effectué avec un tampon contenant 25 mM Tris pH 8,5 ; 150 mM NaCl ; 50 mM d'imidazole ; 0,4 mM DDM. Enfin la protéine est éluée avec un tampon d'éluion 25 mM Tris pH 8,5 ; 150 mM NaCl ; 0,4 mM DDM et 200 ou 400 mM imidazole ou 50 mM EDTA (Éthylènediaminetétraacétique). Les fractions contenant la protéine sont regroupées et diluées trois fois avec du tampon d'équilibration de la colonne suivante, et déposer sur celle-ci. La colonne échangeuse d'anion (HisTrap Q HP 1ml, GE Healthcare) est pré-équilibrée avec 25 mM Tris pH 8,5 ; 0,4 mM DDM. La protéine est éluée grâce à un gradient linéaire de 0 à 1 M NaCl pour un volume d'éluion total de 25 ml par mélange avec le tampon d'équilibration. Une étape alternative à la colonne échangeuse d'anion été utilisée, un passage sur une colonne de tamis moléculaire (Superdex 200 Increase 10/300 GL, GE Healthcare) a été réalisé avec un tampon 25 mM Tris pH 8,5 ; 150mM NaCl ; 0,4 mM DDM. Une version de Ac-Llp avec un site de coupure à la TEV a été réalisée ; dans ce cas la coupure est effectuée sur une nuit avec 1 mg de TEV pour 20 mg de Ac-Llp sous dialyse contre un tampon 25 mM Tris pH 8,5 ; 150mM NaCl ; 0,2 mM DDM. La protéine est repassée sur la colonne NiNTA (Nickel NiTriloacetic Acid), ce qui passe à travers est récupéré, purifié et reconcentré par la colonne échangeuse d'anion.

## 2.3.2 Llp fusionnée a la MBP

### 2.3.2.1 Surexpression

Llp fusionnée à la MBP est surexprimée dans une souche d'*E. coli* BL21(DE3) transformée par le plasmide MalE-his-Llp. Les précultures sont réalisées en LB et en présence amplicilline 125 µg/ml, inoculé avec un stock glycérol et incubé à 37°, 180 rpm, sur une nuit. Les surexpressions sont réalisées dans 1 L de LB, en présence d'ampicilline 125 µg/ml, inoculés à 0,1 de densité optique à 600 nm avec la préculture. La croissance à 37°C est suivie par densité optique jusqu'à une DO de 0,6 puis les cultures sont transférées à 30 °C pendant 1 h. L'expression est induite avec 1 mM d'IPTG (isopropyl β-D-1-thiogalactopyranoside) sur la nuit à 30°C et agité à 180 rpm. À la fin de la culture, les bactéries sont traitées comme celle surexprimant Ac-Llp.

### 2.3.2.2 Marquage isotopique

En vue d'un marquage <sup>13</sup>C et <sup>15</sup>N de Sol-Llp pour étudier la protéine en RMN, des tests de croissance bactérienne ont été réalisés en milieu M9, avec ou sans vitamines, à 37°C, 30°C et 20 °C.

Le protocole décrit précédemment a été utilisé avec quelques changements. La préculture est centrifugée et l'ensemble du culot est utilisé pour ensemercer un litre de culture. Le milieu de culture est un milieu M9 complété en vitamines (50mM Na<sub>2</sub>HPO<sub>4</sub>; 20mM K<sub>2</sub>HPO<sub>4</sub>; 10mM NaCl ; 20mM <sup>15</sup>NH<sub>4</sub>Cl ; 2mM MgSO<sub>4</sub>; 100 µM CaCl<sub>2</sub>; 0,1 mM MnCl<sub>2</sub>; 50 µM ZnSO<sub>4</sub>; 50 µM FeCl<sub>3</sub>; vitamine : pyridoxine ; biotine ; panthothenate hemi calcium ; acide folique ; chlorure de choline ; niacine amide ; riboflavine ; thiamine ; et 4 g/L<sup>-1</sup> glucose <sup>13</sup>C). Les cultures sont transférées à 20°C avant l'ajout d'IPTG. La condition retenue pour la production de Sol-Llp a donc été : M9 supplémenté en vitamine à 20°C expression sur la nuit.

### 2.3.2.3 Purification

Deux méthodes de lyse ont été utilisées n'ayant pas d'incidence sur la présence des ponts disulfure à la fin de la purification :

- Application d'un choc osmotique : 10 g de cellule pour 12,5 ml de tampon (25 mM Hepes pH 7,4 ; 150 mM NaCl ; 0,5 M saccharose) incubé 1 h à 4°C. Centrifugation 13000 g / 15 min / 4°C. Le culot est resuspendu dans 25 ml H<sub>2</sub>O bien froide : incubation 10 min sur glace. Une deuxième centrifugation est appliquée comme précédemment. Le deuxième surnageant est utilisé pour la purification.
- Application d'un choc mécanique : 10 g de cellule pour 30 ml de tampon de lyse (50 mM Tris pH 8, 150 mM NaCl, 2 mM MgCl<sub>2</sub>, DNase, Clapa 1X (inhibiteurs de protéase)). 10 passages à 14000 PSI dans un microfluidizer. Une ultracentrifugation est appliquée pour éliminer les débris cellulaires et les membranes, 30 min / 100000 g / 4°C.

Toutes les étapes de purification sont effectuées à l'aide du dispositif BioRad BioLogic DuoFlow à température ambiante ou sur le système ÄKTA pure à 4°C. Le surnageant contenant MBP-his-Llp est

chargé sur 2 colonnes en série de 5 ml MBPTrap™ HP Prepacked Columns (GE Healthcare), équilibrée avec un tampon d'équilibration (25 mM Tris pH 8, 150 mM NaCl, 1 mM EDTA). Après lavage avec le même tampon, la protéine est éluée avec un tampon d'éluion (25 mM Tris pH 8, 150 mM NaCl, 1 mM EDTA, 10 mM maltose). Les fractions contenant MBP-Llp sont regroupées. La concentration de MBP-Llp est calculée à partir de spectre UV-Visible (voire paragraphe 3.4.1) 1 mg de TEV (Tobacco Etch Virus Protease) pour 20 mg de MBP – Llp est ajouté, ainsi que 3 mM de glutathion réduit et 0,3 mM de glutathion oxydé. Le mélange est dialysé toute la nuit à 4°C (Spectra/Por™, cut off 3 kDa) contre un tampon de dialyse (25 mM Tris pH 8, 150 mM NaCl, 3 mM glutathion réduit, 0,3 mM glutathion oxydé). Le produit de dialyse est chargé sur une colonne de 5 ml HisTrap Chelating Nickel column (GE Healthcare) pour éliminer la TEV et la plus grande partie de la MBP, toutes les deux portant une étiquette polyHis, pré équilibré avec un tampon 25 mM Tris pH 8,5 / 150 mM NaCl. Ce qui passe à travers la colonne est récupéré, puis concentré jusqu'à un volume de 500 µL à l'aide d'un Amicon® Ultra Centrifugal Filters (cut off 3 kDa) et de centrifugations successives de 15 min, 2280 g, 4°C. La dernière étape est un tamis moléculaire, Superdex 75 équilibré avec le tampon final :

- 25 mM Tris pH 6,5 / 150 mM NaCl pour la RMN
- 25 mM Tris pH 8,5 / 150 mM NaF pour le dichroïsme circulaire
- 25 mM Tris pH 8,5 / 150 mM NaCl pour la SPR.

### 2.3.3 FhuA

FhuA sauvage et tous les mutants de FhuA sont traités de la même façon selon le protocole de Ali Flayhan et al. 2012<sup>62</sup>. FhuA sauvage est purifié pour étudier le complexe FhuA-pb5 et FhuA-Llp. Les mutants de FhuA sont produits pour explorer la surface d'interaction entre FhuA et Llp ; une partie des mutants sont repris de Braun et al 1994<sup>66</sup>.

#### 2.3.3.1 Surexpression

FhuA est surexprimée dans une souche d'*E. coli* AW740 ( $\Delta ompF$  *zcb:Tn10*  $\Delta ompC$  *fhuA*)<sup>75</sup> dépourvue des principales porine de la membrane externe, résistante à la tétracycline, et transformée par le plasmide pHX405<sup>73</sup> qui porte une résistance à l'ampicilline et code pour FhuA avec une étiquette histidine dans la quatrième boucle du tonneau  $\beta$ . Les précultures sont réalisées en LB et en présence d'ampicilline 125 µg/ml et de tétracycline 10 µg/ml et incubé à 37°, 180 rpm, sur une nuit. Les surexpressions sont réalisées dans 1 L de LB ou de TB, en présence de 125 µg/ml d'ampicilline ; 10 µg/ml tétracycline; et 100 µM bipyridyl, inoculé à 0,1 de densité optique à 600 nm avec la préculture est incubée à 37°C, 180 rpm, pendant 22 heures. À la fin de la culture, les bactéries sont traitées comme celles surexprimant Ac-Llp.

#### 2.3.3.2 Solubilisation différentielle des membranes

La lyse bactérienne et l'élimination des déchets cellulaire sont effectuées de la même façon et avec les mêmes tampons que pour Ac-Llp. Les membranes totales sont récupérées avec une ultracentrifugation de 20 minutes à 100000 g, 4°C. Les membranes totales sont incubées, soit 30 minutes à 37 °C avec agitation, soit toute une nuit à 4°C avec agitation, avec du tampon de solubilisation des membranes internes (50 mM Tris pH8 ; 81,5 mM OPOE). Les membranes non solubilisées, contenant FhuA, sont récupérées par ultracentrifugation dans les mêmes conditions que précédemment. Le culot contenant les membranes externes est remis en suspension et incubé, soit 1 heure à 37°C sur une nuit à 4 °C, avec le tampon de solubilisation des membranes externes (50 mM Tris pH 8 ; EDTA 1mM ; 43,6 mM Lauryldiméthylamine oxyde (LDAO)). Une dernière ultracentrifugation dans les conditions précédentes est effectuée pour éliminer le matériel non solubilisé.

#### 2.3.3.3 Purification de FhuA

Le surnageant contenant FhuA solubilisé est chargé sur une colonne de 5 ml HisTrap Chelating Nickel column (GE Healthcare), équilibrée avec un tampon d'équilibration (20 mM Tris pH 8 ; 5 mM imidazole ; 4,4 mM LDAO). Après lavage avec le même tampon, un tampon de délipidation (20 mM Tris pH 8 ; 5 mM imidazole ; 43,6mM LDAO) est passé sur la colonne à 0,2 ml/min, permettant d'éliminer le LPS lié à FhuA dans le cas où FhuA est utilisé pour la cristallographie, cette étape est omise pour les tests fonctionnels. Enfin la protéine est éluée avec un tampon d'éluion (20 mM Tris pH 8 ; 200 mM imidazole ; 4,4 mM LDAO). Les fractions contenant FhuA sont regroupées et diluées trois fois avec du tampon d'équilibration de la colonne suivante, et déposées sur celle-ci. La colonne échangeuse d'anion (HisTrap Q HP 1ml, GE Healthcare) est pré-équilibrée avec 20 mM Tris pH 8 ; 2,2 mM LDAO. La protéine est éluée grâce à un gradient linéaire de 0 à 1 M NaCl s'ajoutant au tampon d'équilibration.

#### 2.3.3.4 Échange de détergent

En fin de purification, FhuA est solubilisée dans 2,2 mM LDAO, cependant pour certaines études un échange de détergent est nécessaire. Dans ce cas FhuA est diluée 20 fois avec de l'eau afin de la faire s'agréger. L'agrégation de la protéine est vérifiée par spectrométrie UV-visible (spectre diffusant). Une ultracentrifugation de 60 min à 271000 g, à 4°C permet de récupérer FhuA agrégée. Le culot est rincé avec de l'eau, puis la protéine est reprise dans un tampon de composition adapté aux expériences contenant le nouveau détergent. La quantité du nouveau détergent est calculée en multipliant le nombre de mol de FhuA par 100 puis en ajoutant une CMC (concentration micellaire critique) puis 1 mM.

#### 2.3.3.5 FhuA-L10-Llp

Le protocole est le même que pour FhuA seul (voir section précédente) avec le plasmide FhuA-L10-Llp (cette étude) qui portent une résistance à l'ampicilline et code pour FhuA avec une étiquette

histidine dans la quatrième boucle du tonneau  $\beta$  un lien de 10 acides aminés puis Llp finissant par une étiquette Streptavidine.

#### 2.3.4 Production de protéines utilitaires

##### 2.3.4.1 TEV

Le plasmide et le protocole de la TEV nous ont été transmis au sein du laboratoire par une autre équipe. La protéine TEV (*Tabacco etch virus*) est une protéase qui permet de couper le site ENLYFQ|S(G,A) entre la MBP et la Llp soluble. Le plasmide pcDNA3.1 TEV contenant la protéine TEV est fraîchement transformé dans une souche *E. coli* Rosetta(DE3). L'expression se fait dans 1 L de TB en présence de 50  $\mu\text{g/ml}$  d'ampicilline et de 34  $\mu\text{g/ml}$  de chloramphénicol. Les bactéries sont incubées à 37°C jusqu'à 0,4 de DO puis la température est réduite à 20°C. L'expression de la protéine est induite avec 1 mM d'IPTG à une DO de 0,8 puis incubé sur la nuit à 20°C. Les bactéries sont lysées dans un tampon (25 mM tris pH 8 ; 500 mM NaCl ; 15 mM imidazole et 10 % glycérol), par 6 passages au Microfluidizer 15000 PSI. Les débris cellulaires sont éliminés par une centrifugation : 45 min / 70000 g / 4°C. Le surnageant de lyse est purifié sur une colonne de 5 ml HisTrap Chelating Nickel column (GE Healthcare), équilibrée avec un tampon d'équilibration (25 mM Tris pH 8, 300 mM NaCl, 30 mM imidazole, 10% glycérol). La protéine est éluée avec le même tampon et 300 mM d'imidazole. La TEV est ensuite dialysée contre un tampon 50 mM Tris pH 8, 150 mM NaCl, 10% glycérol, 1 mM DTT. Après la dialyse la protéine est diluée à 1 mg/ml avec une concentration finale de glycérol de 40%, congelée dans l'azote liquide et conservée à -80°C.

##### 2.3.4.2 MSP

Les nanodisques, permettant de stabiliser les protéines membranaires<sup>76</sup>, ont été formés grâce à un mélange de lipides et à la MSP (Membrane Scaffold Protein). Le plasmide contenant la MSP1E3D1<sup>77</sup> est transformé dans des BL21(DE3) avant chaque purification pour augmenter le rendement. L'expression se fait dans 1 L de milieu d'auto-induction (10g Tryptone ; 5g extrait de levures ; 2,5 g  $\text{MgSO}_4$  ; 0,5 g  $(\text{NH}_4)_2\text{SO}_4$  ; 2,7 g  $\text{KH}_2\text{PO}_4$  ; 2,8 g  $\text{Na}_2\text{HPO}_4$  ; 12,5 g Glycérol ; 1,25 g glucose ; 5 g  $\alpha$ -lactose pour 1 L de milieu) en présence de 50  $\mu\text{g/ml}$  de kanamycine. Les bactéries sont incubées à 27°C sur la nuit, récupérées par centrifugation et resuspendues dans un tampon de lyse (50 mM tris pH 8 ; 300 mM NaCl). Les bactéries sont lysées par deux passages au Microfluidizer 15000 PSI. Les débris cellulaires sont éliminés par une centrifugation (45 min / 70000 g / 4°C). Le surnageant de lyse est purifié sur une colonne de 5 ml HisTrap Chelating Nickel column (GE Healthcare), équilibrée avec un tampon d'équilibration (50 mM Tris pH 8, 300 mM NaCl). Trois lavages sont appliqués sur la colonne. Premier lavage : 50 mM Tris pH 8 ; 300 mM NaCl ; 20 mM imidazole ; 1% triton. Deuxième lavage : 50 mM Tris pH 8 ; 300 mM NaCl ; 20 mM imidazole ; 50 mM Na cholate. Troisième lavage : 50 mM Tris pH 8 ; 300 mM NaCl ; 40 mM imidazole. La protéine est éluée dans le troisième tampon contenant 400 mM d'imidazole. La protéine est ensuite digérée par la TEV, sous dialyse sur la nuit (50 mM Tris, 150 mM

NaCl). La TEV et les protéines non clivées sont enlevées par une colonne NiNTA et la protéine clivée, qui passe dans le non-retenu, est congelée dans l'azote liquide et conservée à - 80°C.

## 2.4 Caractérisations biochimiques

### 2.4.1 Détermination de la concentration de protéine:

La détermination de la concentration des protéines est primordiale pour ajuster les conditions de mesure en biologie structurale. La justesse de cette concentration est particulièrement importante pour les études d'interaction des protéines afin de déterminer un  $K_d$ .

La concentration en protéines est déterminée par mesure spectrophotométrique sur un volume de 100  $\mu$ l dans une microcuvette en quartz de 1 cm ou 1 mm de trajet optique ( $\epsilon_{Ac-Llp} = 1,64 \text{ (l.g}^{-1}\text{).cm}^{-1}$  /  $\epsilon_{MBP-HIS-Llp} = 1,54 \text{ (l.g}^{-1}\text{).cm}^{-1}$  avant coupures /  $\epsilon_{sol-Llp} = 1,85 \text{ (l.g}^{-1}\text{).cm}^{-1}$  après coupure,  $\epsilon_{FhuA} = 1,29 \text{ (l.g}^{-1}\text{).cm}^{-1}$ ).

### 2.4.2 SDS-PAGE

La qualité et la pureté des protéines sont vérifiées par gel d'acrylamide (10 ou 15 % acrylamide ; 0,371 M Tris pH 8,8 ; 0,1% SDS ; APS et TEMED pour la polymérisation). Les échantillons sont préparés dans un tampon de charge (0,1 M Tris pH 6,8 ; 4% SDS ; 20% glycérol ; 0,05% bleu de bromophénol) avec ou sans  $\beta$  mercaptoethanol (0,2 à 20%) pour être déposés dans le puits du gel d'acrylamide. Pour FhuA et FhuA-L10-Llp les échantillons sont chauffés ou non à 80°C pendant 2 min. La migration est réalisée à l'aide du système d'électrophorèse vertical 1-D Mini-PROTEAN Tetra Cell (BIO-RAD) selon les recommandations du constructeur. Pour augmenter la résolution du gel 15%, des tampons Tris-Tricine sont utilisés (tampon de la cathode : 0,1 M Tris ; 0,1M Tricine ; 0,1% SDS tampon de l'anode : 0,2 M Tris-Hcl pH 8,9) et les gels sont mis à migrer avec un courant de 180 mV pendant 1h.

Les gels sont colorés par une incubation de 30 min à 1 h dans un bain de bleu de Coomassie (2% éthanol ; 3% acide acétique ; 1,82 mM bleu brillant R), puis décolorés, par un passage de 10 min au four à micro-ondes dans de l'eau milliQ™.

Pour les immunotransferts, le transfert sur une membrane de nitrocellulose 0,2 $\mu$ m est réalisé avec le système Trans-Blot® Turbo™ (Bio-Rad) en utilisant le programme "Bio-Rad Mixed" (7 min / 1,3 A / 25 V), et le tampon commercial associé (25 mM Tris ; 192 mM glycine pH 8,3 ; 20% MeOH). La membrane de nitrocellulose est ensuite saturée 30 minutes à température ambiante (lait écrémé 5% poids/volume (Regilait) ; 0,1% TBS-tween 20). Pour Ac-Llp, FhuA et FhuA-L10-Llp les anticorps utilisés sont des anti-histidine (Sigma-Aldrich) couplés à la HRP (Horse Radish Peroxydase) dilués au 1/5000<sup>ème</sup> dans le même tampon, incubé pendant 1 heures à température ambiante puis laver 15 min (0,1% TBS-tween 20) et révélés par précipitation colorimétrique en utilisant les tablettes SIGMAFAST™ 3,3'-Diaminobenzidine. Pour MBP-HIS-Llp les anticorps utilisés sont des anticorps anti-MBP de souris diluer au 1/1000<sup>ème</sup> (M1321-200UL, Sigma-Aldrich) (lait écrémé 5% poids ; volume (Regilait) ; 0,1% TBS-tween 20) incubés pendant 1 heure à température ambiante. La membrane est ensuite lavée 3 fois 5 min (0,1%

TBS-tween20) puis hybridée avec des anticorps anti-anticorps de souris couplé à la HRP (A9044, Sigma-Aldrich) diluer au 1/5000<sup>ème</sup> et révélée comme précédemment.

### 2.4.3 SEC-MALS

Le SEC-MALS (Size-Exclusion Chromatography - Multi-Angle Light Scattering) est une méthode qui couple une chromatographie d'exclusion stérique (SEC) avec un détecteur de diffusion de lumière multiangle (MALS) et un détecteur UV. L'analyse de l'ensemble des signaux UV et de diffusion de lumière permet de détermination la de masse. Cette technique permet de caractériser et d'identifier les solutions protéiques, contenant des monomères, oligomères, hétérocomplexes, ...

Les expériences de SEC-MALS ont été réalisées sur la plateforme de l'ISBG, par Aline Le Roy, à température ambiante. Des échantillons de 20-100 µL ont été injectés à 0,5 ml/min<sup>-1</sup> sur une colonne Superdex 75 10/300 GL ou Superdex 200 10/300 GL (GE Heathcare), équilibré avec 20 mM Tris-HCl pH8.5, 150 mM NaCl, et détergent si nécessaire (voir Tableau 2-8). De l'albumine de sérum bovin à 2 mg/ml<sup>-1</sup> dans un tampon PBS (Phosphate-Buffered Saline) a été injectée comme contrôle. L'analyse a été réalisée par Aline Le Roy (M&P, IBS) avec le logiciel ASTRA, v5.4.3.20 (Wyatt, Santa-Barbara, USA).

Ensemble des expériences menées :

Noms du projet	type de colonne	volume injecté en µL	Detergent en concentration en mM
Sol-Llp-MBP-TEV	SD 75	100	Non
Ac-Llp	SD 200	100	0,4 mM DDM
Ac-Llp DDM/LMNG/OG	SD 201	50-80	2,2 mM DDM / 21,9 mM OG / 2 mM LMNG
Ac-Llp DDM	SD 202	50	5,1 mM DDM
Ac-Llp DDM	SD 203	50	1 mM DDM
FhuA/Ac-Llp DDM	SD 204	20	0,4 mM DDM
FhuA/Ac-Llp DDM / LMNG	SD 205	40-60	3,9 mM DDM / 0,2 mM LMNG
Sol-Llp / MBP / TEV	SD 75	20	Non

Tableau 2-8 : Protéine et solvants testées en SEC-MALS.

### 2.4.4 Dichroïsme circulaire

Les structures secondaires des protéines absorbent différemment la lumière polarisée circulairement à gauche ou à droite. La mesure des longueurs d'ondes dans les UV lointains permet donc de déterminer la proportion d'hélices alpha, de feuillets bêta ou de boucles dans la protéine. Les structures secondaires et le repliement global des protéines sont mesurés par la différence d'absorption entre la lumière polarisée circulairement à gauche ou à droite.

Les mesures ont été réalisées sur un spectropolarimètre de dichroïsme circulaire (CD) JASCO J-810 équipé d'un régulateur de température (système Peltier). Les spectres ont été enregistrés sur les protéines Ac-Llp et Sol-Llp à 30 µM dans un tampon 25 mM Tris pH 8,5 / 150 mM NaF (+ 0,4 mM DDM pour Ac-Llp) en absence ou en présence d'EDTA. Les spectres ont été enregistrés sur 15 acquisitions successives dans une cuvette de 1 cm. Les valeurs sont valables pour une tension inférieure à 700 V soit dans notre cas entre 260 nm et 200 nm (UV lointains). Les valeurs ont été normalisées par soustraction du blanc et converti en ellipticité molaire moyenne (deg.cm<sup>2</sup>.dmol<sup>-1</sup>) à l'aide de la formule suivante :



$$\theta(\text{deg. cm}^2 \cdot \text{dmol}^{-1}) = \frac{\text{ellipticit  (mdeg)} \times 10^6}{\text{longueur de cuve (mm)} \times \text{concentration de prot ine (\mu M)} \times n}$$

Avec n le nombre de liaison peptidique dans la prot ine. La pr diction du type de structure secondaire a  t  r alis e   l'aide d'un outil en ligne K2D (<http://cbdm-01.zdv.uni-mainz.de/~andrade/k2d3/>).

#### 2.4.5 Temp rature de d naturation par la nano-DSF

Le nano-DSF est une m thode de « differential scanning fluorimetry » qui d termine la stabilit  de prot ine en suivant la fluorescence intrins que de tryptophane et des tyrosines en fonction de la d naturation thermique de la prot ine. La temp rature de fusion ( $T_m$ ) est mesur e gr ce   l'instrument Prometheus NT.48 (NanoTemper Technologies). Les  $T_m$  des deux formes de Llp ont  t  mesur s en absence ou en pr sence de  $\beta$  mercapto thanol et d'EDTA. 10  $\mu\text{L}$  de prot ine   0,2 mg/ml ont  t  charg s dans un capillaire en verre (Cat. No. PR-C002, nanoDSF Grade Standard Capillaries) et plac s sur le rack de lecture. Les courbes de d naturation sont obtenues sur une plage de temp rature comprise entre 20 et 95°C, avec une rampe de 1°C.min<sup>-1</sup>   une puissance d'excitation de 20%. Les  $T_m$  ont  t  calcul s avec le logiciel PR.ThermControl (NanoTemper Techonologies) en analysant la premi re courbe d riv e de l'auto fluorescence   350 nm trac e en fonction de la temp rature. Les interactions de Ac-Llp et Sol-Llp avec FhuA ont aussi  t  contr l es par cette technique. La concentration de FhuA a  t  fix e   10  $\mu\text{M}$ , et plusieurs ratios molaires ont  t  r alis s (de 1 /0,5   1/200 FhuA/Llp).

## 2.5  tudes fonctionnelles et interactions

### 2.5.1 Test de fonctionnalit 

#### 2.5.1.1 Test de sensibilit  *in vivo*

Le bact riophage T5 infecte les souches d'*E.coli* en formant des plages de lyse sur un tapis bact rien. Le test de sensibilit  consiste   mettre en pr sence une souche et le phage T5st0 pour v rifier la lyse bact rienne. Les diff rentes souches d'*E.coli* test es sont C43 et BL21(DE3) en absence ou en pr sence de Ac-Llp ; et la souche AW740 en pr sence ou en absence de FhuA et de Ac-Llp. Des souches pathog nes proches d'*E. coli* ont  galement  t  test es : *Shigella flexneri* (CIP 106171), *Shigella sonnei* (CIP 106204), *Salmonella enterica serovar Typhimurium* (CIP 104474), *Salmonella enterica indica* (CIP 102501T), *Salmonella bongori* (CIP 82.33T) and *E.coli EPEC* (CIP 52.170). Le titre de phage a  t  contr l  pour chaque exp rience sur la souche F.   partir d'une pr culture, une culture est r alis e   37°C jusqu'  une DO de 0,5 ou 0,8 ou 1. 0,3 ml de cette culture sont pr lev s et m lang s   1.5 ml de LB Agar (5 g/L) puis  tal s sur une boite de p tri. Une fois la g lose molle g lifi e, 5  $\mu\text{L}$    des concentrations de moins en moins concentr s de phage sont d pos s sur la boite. Les boites sont incub es toute une nuit   37°C. Le titre approximatif est calcul    partir du nombre de plages de lyse sur un d p t de 5 $\mu\text{L}$ . Pour acc der   un titre plus pr cis, la m me exp rience est r alis e, mais 100  $\mu\text{L}$  d'une concentration de phage choisi est m lang  directement avec les bact ries et le LB Agar. Les plages de lyse sont alors compt es sur l'ensemble de la boite.

### 2.5.1.2 Test d'absorption de phage

Le test d'absorption permet de vérifier si les phages sont absorbés (donc liés à leurs récepteurs) même dans le cas où ils ne parviennent pas à se multiplier dans la bactérie. L'absorption irréversible du bactériophage T5 est caractérisée pour l'ensemble des souches utilisées pour cette étude (*F*, C43, BL21(DE)3, AW740, *Shigella sonnei*, *Shigella flexneri*, *Salmonella enterica*, *Salmonella bengori*). À partir d'une préculture, une culture est inoculée à 0,01 et incubée à 37°C avec agitation jusqu'à une  $DO_{600nm}$  de 0,2. Ensuite, 5 mL de cette culture est inoculée avec le phage T5 à une MOI de 0,5 avec 0,01 M  $CaCl_2$  et 0,4 M  $MgCl_2$  et incubée à 37°C. 200  $\mu$ l de ce mélange sont prélevés à des temps réguliers (0 ; 2,5 ; 5 ; 8 ; 12 ; 18 ; 25 et 35 min), puis vortexés et centrifugés 10 minutes à 12000 g. Les surnageants sont dilués en série et déposés sur des tapis bactériens en spot de 10  $\mu$ l ou 100  $\mu$ l d'une dilution de surnageant sont mélangés directement avec les bactéries et le LB Agar pour être étalés sur une boîte complète. Le titre obtenu dans les surnageants correspond au nombre de phages non liés, la différence avec le titre initial donne le nombre de phages liés irréversiblement sur le récepteur (FhuA).

### 2.5.1.3 Test d'interaction fonctionnelle FhuA-Ac-Llp *in vitro*

L'incubation de FhuA même en détergent avec le bactériophage T5 provoque l'éjection d'ADN du phage dans le milieu réactionnel, ce qui provoque l'inactivation des phages. En revanche si FhuA est bloquée par Ac-Llp il n'est plus capable de lier pb5, les phages restent infectieux. Les bactériophages sont donc incubés avec FhuA, seule ou en présence de Ac-Llp et Sol-Llp ou avec FhuA-L10-Llp. Le ratio de molécules de FhuA/ T5 utilisé est de  $4,5 \times 10^7$  avec une concentration de FhuA fixée à 110 nM. Pour chaque expérience *in vitro*, le nombre de phages initial est contrôlé, ainsi que la capacité de FhuA à inactiver l'ensemble des bactériophages présents. L'intégrité du bactériophage en présence du complexe FhuA-Llp est réalisée comme décrit dans le paragraphe 2.5.1.2. Test d'absorption de phage sur la souche F.

## 2.5.2 Mesure de l'interaction entre FhuA et Llp

### 2.5.2.1 SPR et BLI

La SPR (surface plasmon resonance) est une technique qui permet d'explorer les interactions protéine – protéine ou protéines-molécule. Un des partenaires est fixé sur une surface de particule d'or et l'autre partenaire est passé en flux sur cette surface. Si le composé se fixe, il y a alors un changement d'indice de réfraction du milieu. Cette technique permet à la fois de mesurer le  $K_d$  (constante de dissociation) qui traduit la force d'interaction, mais aussi la cinétique d'interaction  $k_{on}$  et  $k_{off}$  (cinétique d'association et de dissociation) qui traduit la vitesse d'association des complexes. La réponse est mesurée en temps réel et donne un sensorgramme. L'intensité de la réponse dépend à la fois de la concentration et de la taille du partenaire qui se fixe avec (1000 RU  $\leftrightarrow$  1 ng/mm<sup>2</sup>  $\leftrightarrow$  10 mg/ml)<sup>78</sup>.

Les interactions ont été réalisées dans les deux sens (FhuA fixée Llp en flux ou Llp fixée et FhuA en flux) à 25°C sur un Biacore T200. Les premières expériences ont été réalisées avec des surfaces sur lesquelles FhuA était fixée, sur une puce de détection CM3. FhuA a été immobilisée sur des cellules d'écoulement en utilisant une méthode classique de couplage d'amine. La cellule d'écoulement 1 (Fc1) a été fonctionnalisée avec de la BSA et utilisée comme surface de référence. Les Fc1 à 4 (Flow Cell) ont été activées avec 50 µL d'un mélange EDC (N-ethyl-N'-(diméthylaminopropyl) carbodiimide) 0,2 M ; NHS (N-hydroxysuccinimide) 0,05 M ; 0,2 mM DDM a une vitesse de 5 µL/min et fonctionnalisés avec 20 µg/ml de BSA (Fc1 et 3) ou 100 µg /ml de FhuA (Fc2 et 4), les groupes activés restants de toutes les cellules ont été bloqués avec 30 µL d'éthanolamine 1 M pH 8,5. Les quatre Fc ont été traitées à 100 µL/min avec 5 µL de 10 mM HCl pour éliminer les protéines non spécifiquement liées et 5 µL de 50 mM NaOH ; 1M NaCl pour exposer la surface au protocole de régénération. Les surfaces ont été fonctionnalisées avec 2989, 1284, 1695 et 2300 RU de protéine respectivement pour les surfaces Fc 1 à 4.

Les expériences dans l'autre sens ont été réalisées comme décrit précédemment avec des surfaces sur lesquelles Ac-Llp et Sol-Llp ont été fixées. La protéine Ac-Llp a été fonctionnalisée sur la cellule d'écoulement 2 (Fc2), et Sol-Llp sur la 4 (Fc4). Les cellules 1 et 3 ont été fonctionnalisées avec de la BSA. La densité de surface était de 2258, 796, 2290, et 1192 RU de protéine respectivement pour les surfaces Fc 1 à 4.

Des gammes de Ac-Llp, Sol-Llp ou FhuA ont été réalisées dans le tampon d'écoulement 25 mM Tris pH 8,5 ; 150 mM NaCl ; détergent (0,2 mM DDM pour FhuA, Ac-Llp, FhuA-pb5 ou 0,1 mM LMNG (Lauryl Maltose Neopentyl Glycol) pour Ac-Llp). Entre chaque injection les puces ont été régénérées à l'aide d'un tampon de régénération 50 mM Glycine PH12 ; 25 mM EDTA ; 2,3 mM Triton ; 0,2 mM DMM ou 0,1 mM LMNG.

Protéine en flux	Surface	Gamme
Ac-Llp	FhuA	233 – 0,91 µM
Ac-Llp	FhuA	700 - 0,68 µM
Sol-Llp	FhuA	509 - 0,99 µM
FhuA-pb5	Ac-Llp & Sol-Llp	13 – 0,4 µM
FhuA	Ac-Llp & Sol-Llp	13 – 0,4 µM

*Tableau 2-9 : Gammes de concentration des protéines en flux pour la SPR et le BLI.*

Les sensorgrammes résultants ont été corrigés en fonction de la surface de référence (soustraction du signal Fc1 ou Fc3). L'affinité apparente des composés a été déterminée en ajustant le modèle d'affinité à l'état d'équilibre aux courbes des réponses de liaison en fonction de la concentration afin de pouvoir tenir compte de la correction du solvant sur le signal de SPR.

Comme pour la SPR le BLI (BioLayer Interferometry) permet d'explorer les interactions protéine – protéine ou protéines-molécule. Un des partenaires est fixé sur un capteur qui est incubé dans une

solutions de l'autre partenaire sous agitation. La détection de l'interaction est permise par la variation de l'interférence créée lors de la réflexion d'un faisceau lumineux à la surface du biocapteur qui dépend de l'épaisseur de biomolécule. Pour les expériences de BLI les protéines FhuA et Sol-Llp ont été biotinylées à l'aide du kit EZ-Link™ Sulfo-NHS-LC-LC-Biotin (Thermo Scientific™). Les protéines FhuA et Llp ont été transférées dans un tampon 25 mM Hepes pH 7,5 ; 150 mM NaCl ; (FhuA + 0,2 mM DDM) pour cette étape par échange de détergent (voir paragraphe 3.3.3.4) pour FhuA ou par une colonne d'échange de tampon (PD10columns Sephadex™ G25M, réf.17055101) pré équilibré avec le nouveau tampon pour Llp. Selon les recommandations, la biotine a été préparée juste avant la réaction, et le mélange avec la protéine à marquer est incubé 2h à température ambiante avec un ratio 15 µmol de protéine pour 330 µmol de biotine pour Sol-Llp et un ratio 0,01 µmol / 1500 µM pour FhuA. La biotine en excès est enlevée par une colonne de dessalage PD-10 (PD10columns Sephadex™ G25M, réf.17055101) avec le tampon Hepes. La biotinylation a été contrôlée par spectrométrie de masse.

Les protéines biotinylées FhuA ou Sol-Llp ont été immobilisées sur des capteurs Octet® SA Biosensors (Sartorius) en 600 s à une agitation de 1000 rpm. Les concentrations de protéine dans la phase liquide d'immobilisation testées sont 5 µg/ml, 10 µg/ml et 50 µg/ml pour FhuA et 1 µg/ml, 5µg/ml et 50µg/ml pour Sol-Llp. Pour les mesures, toutes les étapes ont été réalisées à 1000 rpm avec un temps d'association de 600 s et un temps de dissociation identique avec des gammes de concentration de protéine identiques à la SPR. Les capteurs ont été régénérés 20 s à 1000 rpm dans le même tampon de régénération que la SPR. Les sensorgrames ont été traités avec le logiciel constructeur dédié « Data Analysis HT 11.1 softwares ».

#### 2.5.2.2 MST (*MicroScale Thermophoresis*) :

La technique de MST permet de détecter l'interaction entre biomolécule (protéine – protéine ou protéine petit molécule). Le principe est de suivre le mouvement et l'état d'hydratation d'une protéine marquée par fluorescence dans un gradient de température induit par un laser infra-rouge, en présence de différentes concentrations de son partenaire. La protéine FhuA a été marquée par un fluorophore à l'aide du kit Monolith His-Tag Labeling Kit RED-tris-NTA 2nd Generation (NanoTemper Technologies) selon le protocole recommandé. La concentration de FhuA a été fixée à 500 nM et les protéines non marquées Ac-Llp et Sol-Llp ont été diluées en série dans le même tampon que FhuA (25 mM Tris pH8.5 ; 150 mM NaCl ; 0,2 mM DDM). Sol-Llp a été testé entre 48 nM et 1,6 mM, Ac-Llp entre 5,3 nM et 700 µM. La protéine marquée a été titrée en dilution 1:1 et insérée dans des capillaires Monolith NT.115. Les mesures ont été effectuées en utilisant une puissance de 30 % pour la LED (light-emitting diode) et de 20 % pour la MST pendant 20s. Le  $k_d$  a été déterminé en utilisant l'analyse MST du logiciel MO.Affinity Analysis 3.

### 2.5.2.3 AUC

L'AUC est une technique qui étudie la migration des macromolécules soumises à une centrifugation haute vitesse. La migration au cours de la centrifugation dépend de la masse des protéines, mais aussi de leurs densités, de leur forme globale. Dans le cas des protéines membranaires, le détergent peut avoir une grande influence sur la migration des protéines. La sédimentation est suivie à la fois en diffusion de lumière et en absorption à 280 nm, ce qui permet d'identifier les différents éléments d'un système complexe. Dans le cas d'un complexe, cette technique permet de déterminer son  $K_D$  et sa stœchiométrie.

Les expériences d'AUC ont été réalisées dans une ultracentrifugeuse analytique XLI (Beckman, Palo Alto, CA) de la plateforme de l'ISBG (Integrated Structural Biology Grenoble), par Aline Le Roy, utilisant un rotor ANTi-60 ou ANTi-50, avec des pièces centrales Ti à double canal (Nanolytics, Allemagne) de 12, 3 ou 1,5 mm de longueur de chemin optique équipées de fenêtres en saphir, le canal de référence étant généralement rempli de solvant sans détergent. L'acquisition a été faite à 20°C et à 130 000  $g$ , pendant la nuit, en utilisant l'absorbance (280 nm) et la détection d'interférence. Le traitement et l'analyse des données ont été effectués par Aline Le Roy (M&P, IBS) à l'aide du programme SEDFIT<sup>79</sup> de P. Schuck (NIH, USA), REDATE<sup>80</sup> et GUSI<sup>81</sup> de C. Brautigam (USA), et en utilisant les équations et protocoles standards décrits dans<sup>82,83,83</sup>. Les valeurs de sédimentation corrigées pour la densité et la viscosité du solvant,  $sw_{20}$ , sont calculées, pour les protéines membranaires, en considérant pour leur volume spécifique partiel, la valeur moyenne entre la protéine et le détergent.

projet	rotor	Taille de la cellule en mm
FhuA/Ac Llp	Anti 50	12; 3 et 1,5
FhuA/Ac Llp	Anti 50	3 et 1,5
FhuA mutants	Anti 60 et 50	3
FhuA mutants $\pm$ Ac Llp	Anti 50	3 et 1,5
FhuA/Ac Llp $\pm$ EDTA	Anti 60	3

Tableau 2-10 : Liste des échantillons testés en AUC.

### 2.5.2.4 Fidabio

La technologie FIDA (Flow Induced Dispersion Analysis) est l'analyse de la dispersion d'un profil d'écoulement hydrodynamique parabolique des macromolécules. Cette dernière dépend de facteurs comme la dimension du canal d'écoulement ou le débit, mais aussi du rayon hydrodynamique de la molécule (directement lié à la taille en kDa pour les protéines solubles globulaires). Dans le cas des protéines membranaires, le rayon hydrodynamique prend aussi en compte la bouée de détergent autour de la protéine. Le facteur d'écoulement peut être impacté par un changement de conformation et/ou une liaison d'une autre particule de taille suffisamment importante ce qui permet d'explorer le  $K_D$  entre deux molécules, ou encore différents stades d'oligomérisation par exemple<sup>84</sup>.

Les expériences FIDABIO ont été réalisées à l'aide d'un instrument Fida 1 utilisant une détection de fluorescence induite par une diode électroluminescente (DEL) avec une longueur d'onde d'excitation de 480 nm et une longueur d'onde d'émission > 515 nm (Fida Biosystems ApS, Copenhague, Danemark). Les protéines Ac-Llp ou FhuA ont été marquées alternativement avec de l'FITC (Isothiocyanate de fluorescéine) (ref : F7250). Dans le cas de FhuA, l'excès de marqueur fluorescent a été éliminé par une chromatographie d'exclusion SD200 alors que pour Ac-Llp, l'FITC a été enlevée avec le « ThermoFisher Dye Removal Kit » (ref : 22858). Des capillaires d'un diamètre intérieur de 75 µm, d'un diamètre extérieur de 375 µm, d'une longueur totale de 100 cm et d'une longueur jusqu'à la fenêtre de détection de 84 cm (Fida Biosystems) ont été utilisés pour toutes les expériences. 39 nL d'indicateur (protéine marquée et suivie) et 12 µL d'analyte (tampon ou protéine non marquée) ont été injectés par mesure. Quand FhuA a été marquée (indicateur), Llp (l'analyte) a été injectée sur une gamme de 0,1 à 50 µM. Quand Ac-Llp a été marquée FhuA a été titré de 0,68 à 78 µM.

## 2.6 Études structurales

### 2.6.1 RMN

#### 2.6.1.1 Tests préliminaires et attribution

Les expériences qui ont permis les tests de faisabilité et les attributions ont été acquises sur un spectromètre Bruker Avance III 600 MHz équipé d'une sonde à triple résonance ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) refroidie par cryogénie. Les expériences de faisabilité ont été acquises par Alicia Vallet de la plateforme RMN de l'ISBG (IBS, Grenoble). Les expériences d'attributions ont été acquises par Loïc Salmon et Emeline Mestdach au (CRMN, Lyon).

Les données ont été enregistrées sur un échantillon uniformément marqué au  $^{15}\text{N}$  ou au  $^{15}\text{N}$  et  $^{13}\text{C}$  de Sol-Llp dans un tampon 25mM Tris (pH 6,5), 75mM NaCl, 10 %  $\text{D}_2\text{O}$ . Les spectres HSQC (Heteronuclear Single Quantum Coherence)  $^1\text{H}$ ,  $^{15}\text{N}$  ont été enregistrés à 600 MHz sur une gamme de températures de 293 à 303 K, afin de déterminer la température la plus appropriée. D'une manière générale, les dimensions spectrales étaient autour de 1522 Hz (F1) et 9615 Hz (F2), correspondant à 25 x 16 ppm, pour des temps d'acquisition d'environ 84 ms et 106 ms respectivement. Les porteuses RF (radio Fréquence) étaient généralement à 4,70 ppm pour  $^1\text{H}$  et 118,5 ppm pour  $^{15}\text{N}$ , et le délai de recyclage est fixée à 1 s. Des spectres HSQC  $^1\text{H}$ ,  $^{15}\text{N}$  ont été acquis à 303 K à 600 MHz entre chaque spectre d'attribution pour vérifier l'état de la protéine. Les dimensions spectrales étaient identiques à celles indiquées précédemment.

L'attribution de résonance des noyaux des protéines ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) est la première étape vers une étude structurale et de dynamique par spectroscopie RMN. Dans les expériences de RMN, la magnétisation est transférée du proton excité aux noyaux des atomes voisins par les liaisons covalentes. Cette magnétisation est ensuite retransférée aux protons pour être détectée. Chaque noyau sondé donne un déplacement chimique dans une dimension dans le spectre final.

Habituellement, des expériences 2D et 3D sont acquises pour relier les groupes amine NH du squelette carboné aux deux premiers carbones des chaînes latérales. Les résonances  $^1\text{H}$ ,  $^{15}\text{N}$  et  $^{13}\text{C}$  sondés peuvent appartenir soit au même résidu (i) et/ou au résidu précédent (i-1), ce qui permet de relier les acides aminés entre eux.

Par exemple, dans l'expérience HNCA, l'aimantation est transmise d'un proton ( $^1\text{H}$ ) à l'azote directement lié ( $^{15}\text{N}$ ) puis au  $\text{C}\alpha$  par couplage J. Cependant, l'azote a un couplage assez élevé avec le  $\text{C}\alpha$  du résidu i et le  $\text{C}\alpha$  du résidu i-1. Les déplacements chimiques des deux carbones alpha sont visibles sur le spectre pour un même couple  $^1\text{H}$ ,  $^{15}\text{N}$ . Le couplage intra résidu est toujours plus élevé que les couplages des résidus successifs, les pics du résidu i seront plus intenses que ceux du résidu i-1. Dans les expériences HNCACB, les deux déplacements chimiques du  $\text{C}\alpha$  et du  $\text{C}\beta$  des chaînes latérales évoluent en même temps puisque leur fréquence se situe dans la même gamme, et leurs résonances sont donc toutes visibles dans les spectres  $\text{C}\beta$ . Il est important de noter que les déplacements chimiques du  $\text{C}\beta$  sont plus sensibles à l'environnement chimique de la chaîne latérale et sont donc des indicateurs utiles du type de résidu. De même que pour l'expérience HNCA les résidus i-1 sont visibles. Pour faciliter l'identification du pic du résidu i et i-1, la magnétisation peut être orientée vers le résidu précédent uniquement en faisant passer la magnétisation par le CO.

L'ensemble des spectres qui ont permis d'attribuer les déplacements chimiques des NH,  $\text{C}\alpha$  et  $\text{C}\beta$  sont : HNCO, HNcaCO, HNCA, HNcoCA, HNcaCB, HncocaCB. Les dimensions spectrales pour  $^1\text{H}$  (F3) et  $^{15}\text{N}$  (F2) étaient identiques à celles indiquées précédemment, la dimension  $^{13}\text{C}$  (F1) était 4228 Hz (28 ppm) pour HNCA et HNcoCA; 2114 Hz (14 ppm) pour HNCO ; 3322 Hz (22 ppm) pour HNcaCO ; 11326 Hz (75 ppm) pour HNcaCB et HncocaCB. Les porteuses RF étaient placées à 4,70 ppm pour  $^1\text{H}$  et 118,5 ppm pour  $^{15}\text{N}$ , et pour  $^{13}\text{C}$  à : 54,5 ppm dans HNCA et HNcoCA ; 173 ppm dans HNCO; 176 ppm dans HNcaCO; 39 ppm dans HNcaCB et HncocaCB. Pour l'attribution des chaînes latérales les protons HA et HB ont été attribués avec des spectres HAHBNH et HAHBcoNH. En outre en vue d'une attribution automatique, des spectres APSY ont également été enregistrés : CBCACONH, HACACONH et HACANH.

Toutes les données RMN ont été traitées à l'aide du programme NMRPipe. Le programme CCPNMR version 3<sup>85</sup> a été utilisée pour afficher et attribuer les spectres RMN 2D et 3D. Une attribution automatique a été réalisée en parallèle de l'attribution manuelle. L'entrée pour l'attribution automatisée de la résonance du squelette à l'aide du programme UNIO-MATCH version 2.9.5<sup>86</sup> était constituée de la séquence de la protéine et de listes de positions de pics générées manuellement dans les six spectres à triple résonance. Les approches manuelles et automatisées ont produit des résultats identiques pour l'attribution des fréquences de résonance des noyaux du squelette polypeptidique.

### 2.6.1.2 Information structurale 3D.

Les expériences de type NOESY (Nuclear Overhauser Effect Spectroscopy) ont été enregistrées à l'aide d'un spectromètre Bruker Avance III 1 GHz équipés de sondes à triple résonance ( $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ) refroidies par cryogénie, par Emelie Mestdach et Loïc Salmon (CRMN, Lyon).

En plus de l'attribution des déplacements chimiques du squelette carboné, la détermination de la structure tertiaire d'une protéine a besoin d'information spatiale. Les corrélations dans l'espace sont mesurées en utilisant les séquences NOESY qui permettent de déterminer si un proton est à proximité d'un autre, en fonction du temps de mélange de l'expérience. Les distances entre protons sont ainsi mesurées sur une gamme allant de 2.5 à 5 Å. Les contraintes de distances ainsi que des contraintes d'angles dièdres (obtenues à partir de l'analyse des déplacements chimiques) sont ensuite utilisées pour calculer un ensemble de structures tridimensionnelles de la protéine qui respectent au mieux l'ensemble des contraintes expérimentales.

Les calculs de structure ont été effectués à l'aide de l'ensemble des spectres suivants :  $^1\text{H},^{15}\text{N}$  NOESY-HSQC ;  $^1\text{H},^{13}\text{C}$  NOESY-HSQC ;  $^1\text{H}$ -aromatique, $^{13}\text{C}$  NOESY-HSQC ; et HCCH-TOCSY. Les spectres ont été enregistrés à 303K, à 1 GHz. Les dimensions spectrales étaient de 13889 Hz (F1 et F3) et 2534 Hz (F2), correspondant à 13,9 x 25 ppm dans la  $^1\text{H},^{15}\text{N}$  NOESY-HSQC ; 13889 Hz (F1 et F3) et 20123 Hz (F2), correspondant à 13,9 x 80 ppm dans l'analyse des  $^1\text{H},^{15}\text{N}$ -NOESYHSQC,  $^1\text{H},^{13}\text{C}$ -NOESY-HSQC, et  $^1\text{H},^{13}\text{C}$ -NOESY-HSQC (aromatique) ; 13889 Hz (F3) et 20123 Hz (F1 et F2), correspondant à 13,9 x 80 ppm dans la HCCH-TOCSY. Toutes les expériences NOESY-HSQC ont été enregistrées avec un temps de mélange de 120 ms. Le schéma pour le découplage homonucléaire du  $^{13}\text{C}$  dans le TOCSY était dipsi3. La structure tertiaire de Llp a été déterminée par Torsten Herrmann (IBS, Grenoble).

### 2.6.1.3 Détermination de structure par RMN

Les données d'entrée pour l'attribution automatique des résonances des chaînes latérales à l'aide de UNIO-ATNOS/ASCAN<sup>87</sup> étaient constituées de la séquence de la protéine, de la liste attribuée des déplacements chimiques du squelette et de trois spectres NOESY, un spectre [ $^1\text{H},^1\text{H}$ ]-NOESY avec une troisième dimension  $^{15}\text{N}$  et deux spectres [ $^1\text{H},^1\text{H}$ ]-NOESY avec une troisième dimension  $^{13}\text{C}$  dans les régions aliphatiques et aromatiques, respectivement.

Les données d'entrée pour l'attribution automatique des pics NOESY à l'aide de UNIO-ATNOS/CANDID<sup>88,89</sup> étaient constituées des attributions des déplacements chimiques du squelette, de la sortie UNIO-ATNOS/ASCAN des attributions des déplacements chimiques des chaînes latérales et des trois spectres NOESY. Le protocole standard UNIO a été utilisé, qui consistait en sept cycles d'identification des pics NOESY, d'attribution et de calcul de structure. Chaque cycle comprenait la sélection automatisée des pics NOESY avec UNIO-ATNOS<sup>90</sup>, l'utilisation des listes résultantes de positions et d'intensités de pics comme données d'entrée pour l'attribution automatisée de NOE (Effet Overhauser nucléaire) par UNIO-CANDID<sup>91</sup> et l'utilisation de l'ensemble final de contraintes de distance



NOE significatives et non redondantes provenant d'UNIO-CANDID comme données d'entrée pour le calcul de structure à l'aide du programme de dynamique moléculaire de contrainte (rMD) CYANA version 3.98<sup>92</sup>. Dans chaque cycle UNIO-ATNOS/CANDID, le résultat consistait en une liste mise à jour des pics croisés NOE attribués pour chaque spectre d'entrée et un ensemble final de contraintes de distance de limite supérieure significatives et non redondantes qui constituaient l'entrée de la routine de recuit simulé de CYANA pour le calcul de la structure 3D. En outre, les restrictions d'angle de torsion pour les angles dièdres du squelette  $\varphi$  et  $\psi$  dérivés de tous les déplacements chimiques du squelette ont été automatiquement générées par UNIO et ajoutées à l'entrée de CYANA pour chaque cycle de calcul de structure. Des contraintes de distance ambiguës ont été utilisées pendant les six premiers cycles UNIO-ATNOS/CANDID. Pour le calcul final de la structure au cycle 7, UNIO n'a retenu que les contraintes de distance qui pouvaient être attribuées sans ambiguïté sur la base de la structure tridimensionnelle de la protéine du cycle 6. Dans chaque cycle UNIO, 80 structures de départ aléatoires ont été soumises à 8000 étapes MD (Molecular dynamics), et les 20 conformères présentant les valeurs résiduelles de fonction cible les plus faibles ont été sélectionnés comme ensemble structural représentatif.

Enfin, les 20 conformères obtenus à partir du cycle UNIO 7 ont été raffinés en énergie dans un solvant explicite à l'aide du programme CNS\_SOLVE version 1.3<sup>93</sup> en utilisant la procédure standard recommandée par l'initiative de la base de données RECOORD (<http://www.ebi.ac.uk/pdbe/recalculated-nmr-data>)<sup>94</sup>.

L'analyse de la qualité stéréochimique des modèles moléculaires a été réalisée à l'aide des outils de validation PDB (<https://validate-rcsb-3.wwpdb.org>)<sup>95</sup>. Les déplacements chimiques  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$  ont été déposés dans la BioMagResBank (<http://www.bmrb.wisc.edu>) sous l'id. BMRB : 34696. Les coordonnées atomiques des faisceaux de 20 conformères RMN sont déposées dans la Protein Data Bank (<http://www.rcsb.org/pdb/>) sous le code d'accèsion : 7QJF.

#### 2.6.1.4 Dynamique de Llp : relaxation de spin

Les données ont été enregistrées et analysées par Emeline Mestdach (CRMN, Lyon) sur un échantillon uniformément marqué au  $^{15}\text{N}$  de sol-Llp dans du tampon Tris 25 mM (pH 6.5), du NaCl 75 mM, 10 %  $\text{D}_2\text{O}$ . Les spectres de relaxation  $^{15}\text{N}$  ont été acquis à 303 K sur un spectromètre Bruker Avance III 600 MHz, équipé d'un cryo-sonde TXI 5mm HCND, pour avoir un aperçu de la dynamique de Llp. Comme mentionné précédemment, les T1, T1 $\rho$ , T2 et NOEs hétéronucléaires donnent des informations sur les mouvements dans l'échelle de temps ps-ns.

Différents délais ont été utilisés pour enregistrer les mesures de temps de relaxation, de manière aléatoire, avec la répétition d'une valeur afin d'évaluer la reproductibilité des expériences : 0, 400, 160, 320, 560, 400, 80, 720, 1120, 880 et 240 ms pour mesurer les T1 et 120, 10, 60, 250, 40, 160, 90, 1, 60, 200, 20 ms pour mesurer les T1 $\rho$ .

Pour les mesures de T1 $\rho$ , le spin-lock radiofréquence (RF) était réglée à 1,4 kHz. L1 $\rho$  a été marquée sur les <sup>15</sup>N seulement, donc aucun découplage <sup>13</sup>C n'a été appliqué. 16 scans par FID ont été enregistrés, avec un délai de récupération entre les scans de 2 s.

Les NOEs hétéronucléaires (hetNOEs) sont basés sur le transfert de l'aimantation d'un noyau à un autre via une relaxation croisée. Un ensemble de deux spectres est enregistré. Les intensités des pics varient entre ces deux spectres en fonction de la proximité entre les noyaux <sup>1</sup>H et <sup>15</sup>N ainsi que leur dynamique : dans l'une des expériences, un train d'impulsions radiofréquence de 180° est utilisé pour irradier les protons et observer l'influence sur le spectre <sup>1</sup>H-<sup>15</sup>N. Les deux spectres ont été enregistrés de manière intercalée. Les porteuses sont restées identiques à celles des expériences de mesure des T1 et T2, mais 96 balayages ont été enregistrés. Un délai de récupération de 5 s a été choisi, et un délai supplémentaire de 3 s a été ajouté en présence ou en l'absence de saturation.

Les spectres ont été traités et analysés à l'aide du logiciel NMRPipe<sup>96</sup>. Les intensités spectrales et les taux de relaxation ont été extraits avec des scripts NMRPipe. Les deux séquences d'impulsions et les scripts de traitement NMRPipe sont disponibles en ligne dans la bibliothèque du groupe de Bax<sup>97</sup>. L'analyse des taux de relaxation et les calculs d'erreur par des simulations de Monte-Carlo ont été effectués à l'aide de scripts Matlab écrits par l'équipe de Loïc Salmon (CRMN, Lyon). Les taux de relaxation longitudinaux R1 et R1 $\rho$  ont été déduits en fittant les intensités décroissantes avec une exponentielle décroissante. Les taux de relaxation transversale R2 ont ensuite été calculés à l'aide de l'équation suivante : avec  $\tan^2(\theta)$  =  $\omega_1/\Omega$  et  $\omega_1$  spin lock RF.

$$R_2 = \frac{R_{1\rho}}{\sin^2(\theta)} - \frac{R_1}{\tan^2(\theta)}$$

Le logiciel Tensor (7-20) a été utilisé pour interpréter les données de relaxation en utilisant le formalisme de modèle libre de Lipari & Szabo<sup>98</sup>.

Tensor est basé sur la théorie de la relaxation, et peut s'adapter soit à des tenseurs isotropiques ou anisotropiques. Le tenseur de diffusion rotationnelle est déterminé en utilisant des résidus qui ne sont pas situés près de la terminaison et qui ne subissent pas d'échange, d'après l'observation des valeurs de R1 $\rho$  ni des valeurs hetNOEs diminuées : généralement, ces résidus sont situés dans des parties structurées de la protéine. Une fois que le tenseur est fixé, des simulations de Monte-Carlo sont mises en œuvre pour tester la pertinence du tenseur. Différents modèles sont utilisés pour analyser les mouvements internes en fonction des échelles de temps des mouvements (et s'ils peuvent être séparés les uns des autres), ainsi que la contribution d'échange des NH du squelette.

Les données expérimentales de relaxation peuvent être analysées en utilisant soit un tenseur isotrope, où la diffusion rotationnelle est indépendante des orientations des vecteurs NH, ou un tenseur anisotrope. Un tenseur de diffusion isotropique a été utilisé, calculé en utilisant les données obtenues sur les résidus dans les parties structurées de la protéine, qui ne présentaient pas de valeurs accrues de

R1 ni de valeurs diminuées de hetNOE ( $< 0,83$ ). La distance  $^1\text{H}-^{15}\text{N}$  a été fixée à 1,02 angström. Les paramètres d'ordre, le temps de corrélation et les constantes d'échange ont ensuite été calculés tout au long de la séquence de Llp. Les paramètres d'ordre des vecteurs  $^1\text{H}-^{15}\text{N}$  sont limités entre les valeurs 0 et 1, 1 étant la limite la plus rigide tandis que 0 est la plus flexible.

Différents modèles peuvent être utilisés pour les données expérimentales : de 1 à 5, avec un nombre croissant de paramètres. En particulier, le modèle 1 contient un paramètre d'ordre alors que le temps de corrélation interne du mouvement (i) correspondant au paramètre d'ordre est considéré comme étant inférieur à 20 ps (clairement plus rapide que le mouvement global). Le modèle 2 est quant à lui l'analyse classique de Lipari-Szabo et « fit » ce i. Le modèle 3 incorpore une composante d'échange  $R_{\text{ex}}$  au modèle 1, tandis que le modèle 4 fait de même au modèle 2. Le modèle 5 inclut un temps de corrélation supplémentaire sur une échelle de temps plus lente, qui peut être séparé, comme dans le modèle étendu de Lipari-Szabo.

Des tests F sont ensuite utilisés pour déterminer si une amélioration significative découle de l'ajout de paramètres, et décider du choix du modèle.

#### 2.6.1.5 *Exploration de la surface d'interaction entre Llp et FhuA*

Le signal RMN dépend de l'environnement chimique des acides aminés. Ainsi si l'environnement chimique change du fait de la liaison d'un ligand, le signal de RMN sera impacté. En réalisant la titration d'un ligand non marqué, on peut suivre les perturbations de déplacements chimiques et les variations d'intensité des pics de la protéine marquée et trouver les acides aminés impliqués dans l'interface d'interaction.

Le détergent choisi pour solubiliser FhuA dans cette expérience est le DDM afin d'être le plus près possible des conditions d'interaction testées entre FhuA et Ac-Llp. Des spectres  $^1\text{H},^{15}\text{N}$ -SOFAS-HSQC (selective optimized flip angle short transient-HSQC) ont été mesurés sur Sol-Llp en présence de DDM seul ou en présence de concentrations croissantes de Sol-Llp. Du DDM est ajouté à la solution mère de Llp de sorte que la concentration totale de DMM dans la solution ne varie pas au cours de la titration : au fur et à mesure que les quantités de Llp sont ajoutées dans le tube de RMN contenant l'échantillon de FhuA, du DDM à 0,4 mM est également ajouté pour éviter la dilution de celui-ci. Un total de 10 rapports FhuA/Llp ont été enregistrés, séparés en 2 sous-titrations (voir Tableau 2-11). Les échantillons ont été mélangés en utilisant des stocks de  $C_{\text{Llp}}=1,2$  mM et  $C_{\text{FhuA}}=0,15$  mM.

$V_{Llp}/\mu\text{L}$	$V_{FhuA}/\mu\text{L}$	$V_{total}/\mu\text{L}$	FhuA-Llp %
5	110	130	0,36
12	110	137	0,87
20	110	145	1,45
30	110	155	2,18
25	65	130	3,08
32	65	137	3,94
40	65	145	4,92
50	65	155	6,15
60	65	165	7,38

**Tableau 2-11 : Titration Llp-FhuA.**

Les pourcentages de complexe sont calculés à l'aide du  $K_{Dsol-Llp/FhuA}$  déterminé par MST.

Les spectres  $^1\text{H},^{15}\text{N}$ -SOFAS-HSQC ont été enregistrés pour chaque ajout de Llp. Les intensités des pics ont été normalisées par le nombre de scans dans l'expérience et la concentration de Llp. La RMN est aussi une technique qui permet d'accéder au  $K_d$  d'un complexe.

### 2.6.2 Microscopie électronique

Les protéines doivent être purifiées puis congelées de façon très rapide afin d'obtenir de la glace vitreuse permettant de capturer la protéine dans plusieurs orientations. Les particules sont ensuite imagées en 2D dans un microscope électronique. Chaque particule est utilisée pour réaliser des moyennes en classe 2D puis un modèle 3D est réalisé à partir des projections des classes 2D.

La résolution de la structure de FhuA-pb5 s'est faite en deux étapes et avec deux jeux de données différents, tous les deux pris sur le Titan Krios de l'ESRF (European Synchrotron Radiation Facility) à 300 kV en utilisant une caméra Gatan K2, par Grégory Effantin (ESRF-IBS, Grenoble).

Un premier jeu d'image d'environ 800 films/micrographies a été utilisé pour faire la preuve de concept du projet. Les données ont été prises avec une plaque de phase afin d'augmenter le contraste des images. L'analyse des images a mené à une structure 3D autour de 6 à 7 Å de résolution, dans laquelle la structure cristallographique de FhuA pouvait être adaptée sans ambiguïté.

Un deuxième jeu de données de 8000 films pris cette fois sans plaque de phase. La structure du complexe a été résolue en utilisant le premier jeu de données comme modèle initial pour l'alignement des particules du deuxième jeu de données. L'ensemble de l'analyse d'image a été réalisé avec le logiciel RELION<sup>99</sup>, par Grégory Effantin (ESRF-IBS, Grenoble).

La protéine pb5 a été tracée *de novo* dans la carte de microscopie électronique sur le logiciel Coot version 0.9.2. FhuA a été tracée à partir de la structure de FhuA obtenue en cristallographie (PDB 2GRX)<sup>100</sup>. Les deux structures ont été affinées séparément puis ensemble grâce au logiciel Phenix

version 1.18.2-3874 avec le module Real-space refinement Cryo-EM. Enfin la structure a été validée à l'aide des outils de contrôle de coot et le site MOL-Probity.

### 2.6.3 Cristallographie

La résolution des structures en cristallographie dépend de la qualité de la diffraction des cristaux. Dans le processus de cristallisation, les conditions physico-chimiques sont modifiées afin de réduire la solubilité de la protéine, ce qui permet d'explorer son diagramme de phase, et éventuellement de franchir la limite de sursaturation et d'atteindre la zone de nucléation puis de croissance des cristaux. La méthode la plus utilisée pour obtenir des cristaux est la diffusion de vapeur avec des gouttes en suspension (condition à la main) ou une goutte assise (condition du HTXLab). Si des cristaux sont obtenus ils sont irradiés par rayons X ce qui permet d'obtenir une carte de diffraction. Après résolution de la phase on obtient une carte de densité électronique permettant de tracer la protéine à l'intérieur.

Les tests de cristallisation sont menés en présence de cristallophore de Lanthanides (qui permet d'augmenter la nucléation)<sup>101</sup>. Le complexe FhuA:Ac-Llp a été réalisé avec 1 mol de FhuA pour 4 mol de Ac-Llp. Le criblage de cristallisation a été réalisé avec la plateforme de cristallisation haut débit de l'EMBL CRIMS (Crystallographic Information Management System) sur les plaques commerciales Molecular Dimension : MemChannel\_MD, MemStartSys\_MD, MemGold\_MD, Midas-plus\_MD, MemTrans\_MD, MemMeso\_MD. La cristallisation a été menée dans le robot à 4°C et suivie aux jours 2, 5, 9, et 17. Les cristaux ont été pêchés au 28e jour et envoyés à la ligne automatique ID30A-1 de l'ESRF pour cribler la diffraction. Des conditions de cristallisation ont également été reproduites à la main dans des plaques 48 puits à 20°C avec des gouttes de 1 µL de complexe et 1 µL de tampon de cristallisation pour des macros cristaux, mais aussi pour des microcristaux.

## 3 Contexte & Objectifs de thèse

L'infection des bactéries par le bactériophage T5 a été étudiée depuis plusieurs années. Néanmoins, la transduction du signal pour déclencher l'éjection d'ADN reste inconnue pour le phage T5 qui cible un récepteur protéique. Comme décrit dans l'introduction, la liaison de pb5 sur FhuA est capable de déclencher l'injection de l'ADN dans l'hôte. L'hypothèse est que, par cette liaison, pb5 subit des changements de conformation qui sont transmis, tout au long de la queue, pour provoquer l'éjection d'ADN. Afin d'explorer cette hypothèse, un très long travail a été mené pour obtenir la structure du complexe FhuA-pb5 depuis le début des années 2000<sup>60-62</sup>. Jusqu'à aujourd'hui, aucune méthode d'étude n'a permis de comprendre le mécanisme de déclenchement de l'éjection de l'ADN. Les progrès récents de la microscopie électronique ont enfin permis d'obtenir la structure atomique de FhuA-pb5, travail que je présenterai dans la première partie de cette thèse. La résolution de cette structure est la première structure de RBP liant un récepteur protéique. Elle permet l'analyse de la structure de pb5 ainsi que son interaction avec FhuA. De plus, nous avons exploré la spécificité d'interaction entre pb5 et FhuA, en testant la sensibilité de souches portant FhuA mutée ou de souches proches de *E. coli* avec des séquences de FhuA proche (similarité de 96 à 100%). Nous avons montré que T5 était capable d'infecter une souche de *Salmonelle*, et nous avons essayé de comprendre pourquoi il n'infectait pas des souches de *Shigella*, pourtant plus proche génétiquement d'*E. coli*.

Les phénomènes d'exclusion de surinfection sont beaucoup étudiés chez les bactériophages tempérés. Le phage T5 a la particularité de posséder une protéine d'exclusion de surinfection, alors qu'il est strictement lytique. Néanmoins, Llp partage une grande similitude avec les protéines Cor retrouvées chez plusieurs phages lysogéniques. Llp se lie du côté périplasmique de FhuA et empêche la liaison du côté extracellulaire de pb5. Nous avons fait l'hypothèse que l'interaction entre FhuA et Llp provoque des changements de conformation à longue distance dans FhuA. De plus, des mutations de FhuA, aussi bien du côté périplasmique que du côté extracellulaire, ont montré des résistances à l'inactivation par Llp<sup>66</sup>, confirmant cette hypothèse. L'objectif principal de ma thèse est donc de comprendre les mécanismes d'action de la protéine Llp dans l'inhibition des liaisons de FhuA avec ses ligands. La question est : comment la liaison de Llp du côté périplasmique de FhuA peut empêcher la liaison de pb5 du côté extracellulaire ? Comment Llp bloque le transport de toutes les substances transportées par FhuA ? Pour répondre à ces questions nous avons résolu la structure de Llp par RMN ce qui a permis de découvrir une nouvelle structure de phage. Je détaillerai ces résultats dans la deuxième partie de ma thèse. Nous avons également étudié l'interaction entre FhuA et Llp par des méthodes biochimiques, biophysiques et structurales. Je présenterai nos tentatives pour résoudre la structure du complexe FhuA-Llp. Une équipe ayant récemment résolu la structure du complexe FhuA-Llp, je proposerai un mécanisme d'inhibition au regard des résultats de la caractérisation biochimique du complexe FhuA-Llp. Ces résultats seront présentés dans la troisième partie de ma thèse.

## Résultats





## 4 Chapitre I : Structure du complexe FhuA-pb5

---

## 4.1 Introduction

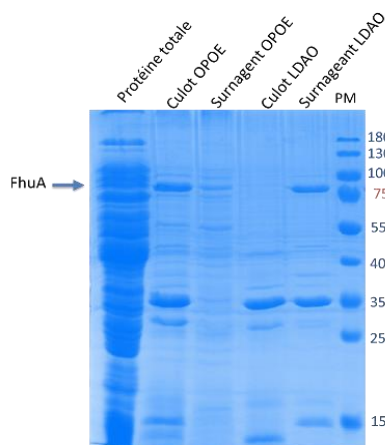
Comme décrit dans l'introduction, la liaison de pb5 à FhuA est capable de déclencher l'injection de l'ADN dans l'hôte. L'hypothèse est que par cette liaison, pb5 subit des changements de conformation. Grâce à ces changements de conformation, le signal d'éjection de l'ADN est transmis jusqu'à la capsid. Afin d'explorer cette hypothèse, un très long travail a été mené pour obtenir la structure du complexe FhuA–pb5 depuis le début des années 2000<sup>60</sup> sans succès. La formation de cristal implique des contacts cristallins sur la surface des protéines. Or dans le cas des protéines membranaires, la surface utile pour la formation de contacts cristallins est réduite car il y a une forte proportion de surface hydrophobe qui est masquée par le détergent. Une technique récente pour augmenter les surfaces d'interaction disponibles pour la cristallisation des protéines est l'utilisation de VHH (camelid heavy-chain variable domains), dérivés d'anticorps de lama. Le complexe FhuA-pb5 a été utilisé pour produire ce type d'anticorps. Bien que des cristaux aient été obtenus avec cette nouvelle méthode, diffractant à une résolution raisonnable (~4Å), la collecte de données n'a pas été optimale et la phase non résolue. Le principal problème est la reproductibilité de la cristallisation.

Pour augmenter les chances d'obtenir des cristaux reproductibles du complexe FhuA-pb5, nous avons utilisé un complexe de lanthanide. Ce composé est un cristallographe, c'est-à-dire qu'il aide au processus de nucléation. En outre il possède des atomes lourds qui permettent d'obtenir la phase par le signal anomal. Des cristaux ont pu être obtenus avec le complexe FhuA-pb5-VHH64 en présence ou en absence de lanthanide. La présence du complexe de lanthanide n'a pas permis d'obtenir de résultats supplémentaires de cristallisation. En revanche il a ralenti la croissance des cristaux et a permis d'obtenir des cristaux légèrement plus grands. Tous les cristaux ont été pêchés grâce à la ligne automatisée de récolte de cristaux et passés sur la ligne de faisceaux de rayons X MASSIF 1 de l'ESRF, mais aucun cristal n'a donné de diffraction. Pour augmenter la qualité des cristaux, des nouvelles plaques de cristallisation ont été réalisées avec un ensemencement issu des cristaux précédents. Une grande quantité de cristaux a été obtenue dans différentes conditions, malheureusement tous les cristaux testés ne présentaient pas de diffraction. Une autre méthode pour obtenir la structure du complexe FhuA-pb5 était donc indispensable. Les progrès réalisés par la microscopie électronique, directement liée à l'amélioration des microscopes (plus stables), des détecteurs, et du traitement d'image ont permis d'envisager cette méthode pour résoudre la structure du complexe FhuA-pb5 (environ 150 kDa) à une résolution atomique. En revanche pb5 seule (environ 68,7 kDa) est vraiment à la limite basse pour être étudiée par microscopie électronique. Je présenterai donc ici l'analyse de la structure du complexe FhuA-pb5 obtenue en microscopie électronique avec l'aide de Gregory Effantin et de Romain Linares. Je présenterai la préparation des échantillons en français, puis la structure du complexe FhuA-pb5 sera présentée dans un format article.

## 4.2 Formation du complexe FhuA-pb5

### 4.2.1 Production de FhuA

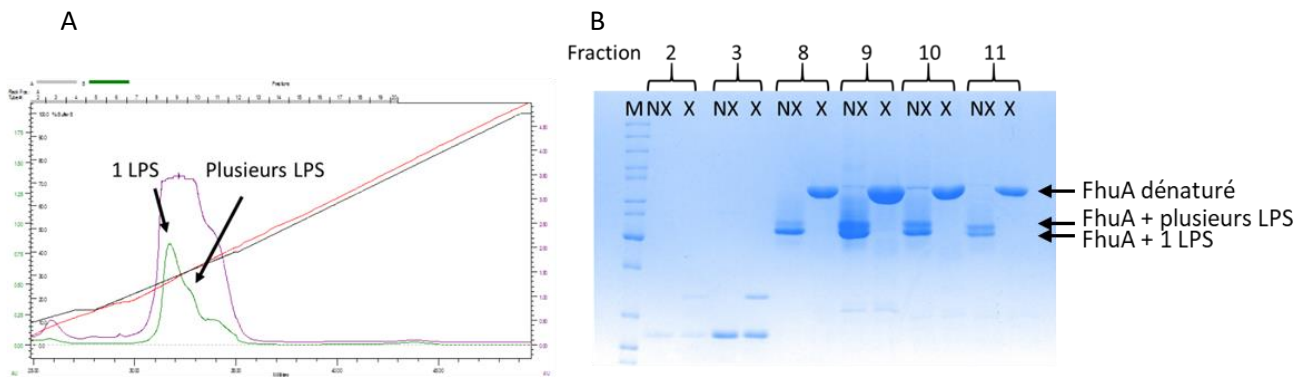
La purification de FhuA était déjà maîtrisée quand je suis arrivée au laboratoire. La solubilisation différentielle des membranes issues des cultures de AW740 surexprimant FhuA en permet d'effectuer une pré-purification. La surexpression de FhuA au milieu des protéines totales sur gel SDS n'est pas évidente à distinguer. Une première solubilisation des membranes totales avec du OPOE solubilise préférentiellement les protéines de la membrane interne et FhuA est retrouvée dans le culot de cette première solubilisation. On note une légère bande de FhuA dans le surnageant OPOE. Cela peut arriver si les membranes sont un peu moins concentrées que d'habitude car la solubilisation dépend de la nature de la protéine, de son environnement lipidique, mais aussi de la concentration de membrane et de détergent. La seconde solubilisation avec du LDAO solubilise les protéines de la membrane externe. Comme prévu, FhuA est présente dans le surnageant de la solubilisation LDAO (Figure 4-1).



**Figure 4-1 : Étapes de solubilisation différentielle de FhuA sur gel d'acrylamide 10%, coloré au bleu de Coomassie.**

Les échantillons sont chauffés 2 min à 100°C en présence de SDS avec réducteur. Protéines membranaires totales, OPOE culot et surnageant, LDAO culot et surnageant, PM : marqueurs moléculaires. FhuA est pointée par la flèche bleue.

La colonne NiNTA a été réalisée pour enlever les contaminants. Pour les études structurales, une délipidation a été réalisée (cristallographie), en revanche les purifications pour les études biochimiques (interaction avec Llp voir Résultat - chapitre III) ont été réalisées sans délipidation. La colonne échangeuse d'anions permet d'affiner la purification de FhuA et de séparer FhuA lipidée de celle délipidée quand cette étape a été réalisée. On distingue en effet deux pics sur le chromatogramme de la colonne échangeuse d'anions, correspondant à FhuA comigrant avec une seule molécule de LPS pour le premier pic et FhuA avec plusieurs LPS pour le second pic. On sait qu'il reste une molécule de LPS sur FhuA délipidée, mais cela représente le plus petit nombre de LPS lié à FhuA et la distribution de LPS est homogène (figure 4-2 A). Quand FhuA n'est pas délipidée on obtient un seul pic large et dissymétrique car le nombre de LPS sur FhuA n'est pas homogène (figure 4-2 B). La délipidation est importante pour la cristallisation de FhuA car elle augmente l'homogénéité de l'échantillon.

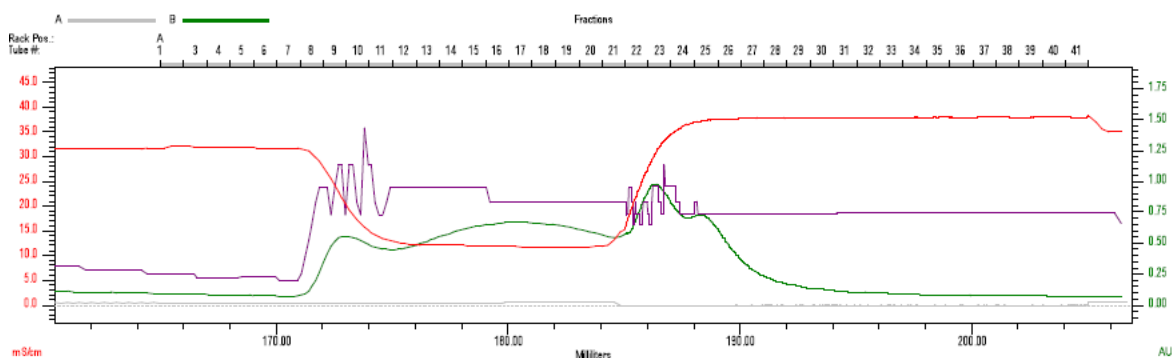


**Figure 4-2 : comparaison des profile d'éluion de FhuA sur une colonne échangeuse d'anions avec ou sans delipidation.**

A. chromatogame La courbe rouge correspond à la conductivité (mS/cm), la courbe verte correspond à l'absorbance à 280 nm (AU) et la courbe violette correspond à l'absorbance à 214 nm (AU). B. Gel SDS correspondant aux fractions d'éluion. X : échantillon chauffé, NX : échantillon non chauffé.

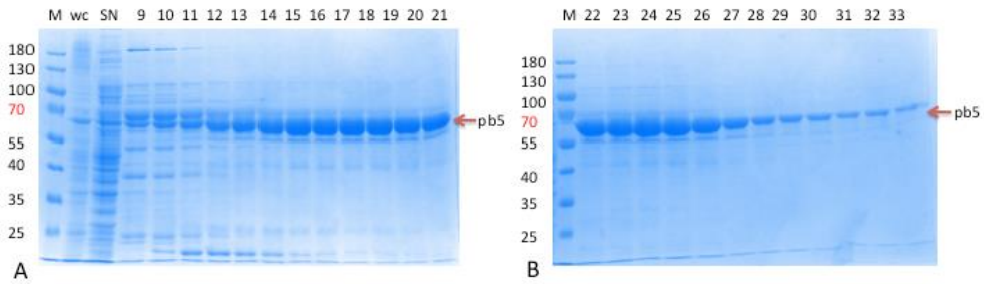
#### 4.2.2 Production de pb5

La protéine pb5 est produite à 20 °C sur une nuit grâce à la fuite du promoteur. La purification de pb5 a été réalisée afin de former le complexe FhuA-pb5. Après la lyse cellulaire, une partie de la pb5 est insoluble et est retirée par ultracentrifugation (Figure 4-4 (A)). Le surnageant contenant la protéine pb5 soluble est chargé sur une colonne NiNTA. La protéine est éluée en deux étapes. La première étape est une élution à faible pH, avec EDTA mais sans sel. Au cours de cette étape, la majorité des contaminants est éliminée et une partie de la pb5 également (Figure 4-3 et 4-4 (A)). La deuxième étape est une élution à faible pH, avec EDTA et avec du NaCl. Dans cette étape, la pb5 est bien éluée dans un deuxième pic. Pour éluer complètement la protéine pb5 de la colonne, un grand volume d'élution est nécessaire (Figure 4-3 et 4-4 (B)). Nous suggérons que pb5 interagit avec la matrice de la colonne de façon aspécifique, probablement par des interactions électrostatiques. La colonne d'échange cationique permet d'obtenir une protéine pure avec un rendement final de  $5.0 \pm 0,2$  mg/l de culture.



**Figure 4-3 : Chromatogramme NiNTA de pb5.**

La courbe rouge correspond à la conductivité (mS/cm), les courbes vertes correspondent à l'absorbance à 280 nm (AU) et la courbe violette correspond à l'absorbance à 214 nm (AU). En haut, le numéro de la fraction collectée (1 à 41) et en abscisse, le volume (ml).

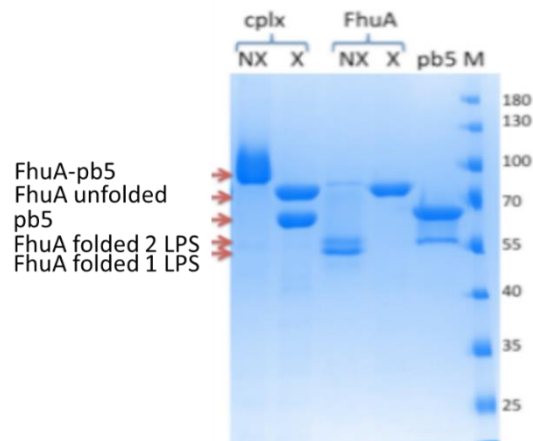


**Figure 4-4 : A : Gel SDS (10 % d'acrylamide) de l'éluion NiNTA de pb5 à faible pH et EDTA puis avec du NaCl (B) colorée au bleu de Coomassie.**

*M : poids moléculaire, wc : débris cellulaires, SN : protéines solubles, fractions 9 à 33. Pb5 est pointé par la flèche rouge.*

#### 4.2.3 Formation du complexe FhuA-pb5

Le complexe FhuA-pb5 est formé par le mélange équimolaire de FhuA et de pb5. La formation du complexe pour la microscopie électronique est contrôlée par un gel SDS (Figure 4-5). Le complexe non chauffé migre comme une bande unique, l'excès d'un des deux partenaires serait visible. En effet le SDS n'est pas suffisant pour dénaturer le complexe FhuA-pb5 dont l'interaction est très forte. Quand le complexe est chauffé en présence de SDS avant dépôt, on distingue 2 bandes, correspondant à FhuA et pb5. FhuA non chauffée n'est également pas complètement dénaturée par le SDS et migre donc à une taille apparente plus petite. Les 2 bandes de FhuA non chauffée montrent que FhuA possède 1 LPS et 2 LPS. FhuA chauffée est à la taille attendue. On distingue également une petite dégradation de pb5.



**Figure 4-5 : FhuA-pb5, FhuA et pb5 sur gel SDS 10 % coloré avec du bleu de Coomassie.**

*Complexe FhuA-pb5 non chauffé (NX) ou chauffé (X), FhuA non chauffé (NX) ou chauffé (X), pb5 (non chauffé) et M : poids moléculaire.*

### 4.3 Structure du complexe FhuA-pb5 : Article

La suite de ce chapitre est rédigée au format d'un article scientifique : l'article «Structure of siphophage T5 Receptor Binding Protein pb5 in complex with its *E. coli* receptor FhuA: deciphering host recognition mechanism and trigger for infection » en soumission.

Nous détaillons la structure de pb5 (RBP<sub>pb5</sub>) en complexe avec son récepteur FhuA. C'est la première structure d'un RBP en complexe avec un récepteur protéique. La structure de pb5 est un nouveau repliement partageant un repliement homologue avec des RBP de phage pour un de ses sous domaine. Grâce à des expériences antérieures de CD<sup>62</sup> et à la structure de pb5 seule prédite par AlphaFold, nous montrons que la liaison de pb5 à FhuA induit des changements de conformation dans pb5. En complétant ces résultats avec les cartes de la queue de T5 en interaction avec du FhuA en nanodisque permettant d'adapter le complexe, nous proposons un mécanisme de déclenchement de l'infection de T5.

La résolution du complexe FhuA-pb5 permet de comprendre le déclenchement de l'infection du phage T5 par la liaison avec son récepteur. Il permet aussi de savoir où se fixe pb5 ce qui n'était pas encore connu. Dans le cadre de l'étude de Llp cette information est importante, pour comprendre l'inhibition de la liaison entre FhuA et Llp. Cette discussion sera présentée dans le chapitre 3 complexe FhuA-Llp.

## Deciphering bacteriophage T5 host recognition mechanism and infection trigger

S raphine Degroux<sup> 1</sup>, Gr gory Effantin<sup> 1</sup>, Romain Linares<sup>1</sup>, Guy Schoehn<sup>1</sup> and  
C cile Breyton<sup>1\*</sup>

<sup>1</sup>Univ. Grenoble Alpes, CNRS, CEA, IBS, F-38000, Grenoble, France.

<sup> </sup>These authors contributed equally to the work

\*Corresponding author's email: Cecile.Breyton@ibs.fr.

**Keywords:** Bacteriophage, electron cryo-microscopy, host recognition, conformational changes

### Abstract

Bacteriophages, viruses infecting bacteria, recognise their host with high specificity, either binding to saccharide motifs or proteins of the cell wall of their host. In the majority of bacteriophages, this host recognition is performed by Receptor Binding Proteins (RBPs) located at the extremity of a tail. Interaction between the RBPs and the host is the trigger for bacteriophage infection, but the molecular details of the mechanisms are unknown for the majority of bacteriophages. Here, we present the electron cryo-microscopy structure of bacteriophage T5 RBP<sub>pb5</sub> in complex with its *E. coli* receptor, the iron ferrichrome transporter FhuA. Monomeric RBP<sub>pb5</sub> is located at the extremity of T5 long flexible tail, and its irreversible binding to FhuA commits T5 to infection. Analysis of RBP<sub>pb5</sub> structure within the complex, comparison with its AlphaFold2 predicted structure, and its fit into a previously determined map of T5 tail tip in complex with FhuA allow us to propose a mechanism of transmission of RBP<sub>pb5</sub> receptor binding to the straight fibre, initiating the cascade of events that commits T5 to DNA ejection.



## Introduction

Bacteriophages, bacterial viruses also called phages, are the most abundant and diversified biological entity on Earth (Dion *et al*, 2020). As bacterial killers, but also as a means of genetic material exchange, they shape and drive microbial population evolution and diversity (Suttle, 2007). The vast majority of phages indexed in databases are composed of a capsid protecting the DNA and bear a tail, which can be long and contractile (*Myoviridae*), long and flexible (*Siphoviridae*) or short (*Podoviridae*). The tail extremity serves to recognise the host surface, perforate its cell wall and safely channel the DNA from the capsid to the bacterial cytoplasm (Nobrega *et al*, 2018; Linares *et al*, 2020). The distal tip of the tail bears Receptor Binding Proteins (RBPs) that reversibly or irreversibly bind saccharides or proteins. Depending on their morphology, RBPs can be fibres (more than 1000 residues long, requiring chaperones to fold) or more globular RBPs often called Tail Spike Proteins (TSP)(Broeker & Barbirz, 2017). Fibres and TSP of known structure share an elaborately interwoven often trimeric state, rich in  $\beta$ -helices. TSP are multidomain proteins, containing at least a phage binding domain and a receptor binding domain (Goulet *et al*, 2020). The receptor binding domain can also bear enzymatic activity, digesting the saccharide receptors (Broeker & Barbirz, 2017). Because they recognise the bacterial host surface with high specificity and because of their enzymatic activities, phage RBPs can be used to identify bacterial strains for diagnosis (Filik *et al*, 2022), to engineer phages with new host range (Dams *et al*, 2019) or in biotechnology (Santos *et al*, 2018). The stoichiometry of RBP per virion can vary from one copy (*e.g.* RBP<sub>pb5</sub> in colisiphophage T5) to 54 (*e.g.* in lactosiphophage TP901-1) and maybe more; phages can also combine RBPs with different binding specificities (Sørensen *et al*, 2021). Multiple binding is proposed to increase affinity of saccharide binding by avidity. It also has been proposed to position the tail tube correctly for cell wall perforation (Taylor *et al*, 2018) and has been shown to trigger infection (*e.g.* T4 (Hu *et al*, 2015; Yap *et al*, 2016) and T7 (Hu *et al*, 2013; González-García *et al*, 2015)). All RBP structures determined to date are saccharide binding RBPs.

Phage T5, a *Siphoviridae* infecting *E. coli*, is a model phage belonging to the T-series introduced by Delbrück and co-workers in the 1940s (Demerec & Fano, 1945). It bears an icosahedral capsid that protects the 121 kb DNA and a 250-nm long flexible tail (Arnaud *et al*, 2017). At the distal end of the tail is located the host recognition apparatus that consists of three side tail fibres and a central straight fibre. The central fibre is composed of a trimer of the C-terminus part of the Baseplate Hub Protein (or Tail protein) pb3 followed by a trimer of pb4, at the tip of which is found a monomer of the 640-residues RBP, RBP<sub>pb5</sub> (Heller & Bryniok, 1984; Zivanovic *et al*, 2014; Linares *et al*, 2020). The side tail fibres reversibly bind Lipopolysaccharide (LPS) O-

polymannose (Heller & Braun, 1982; Garcia-Doval *et al*, 2015), allowing the phage to ‘walk’ at the surface of the bacterium until RBP<sub>pb5</sub> irreversibly binds to FhuA, an *E. coli* iron-ferrichrome transporter (Braun, 2009). *Shigella* and *Salmonella* strains also bear FhuA, sharing high identity with *E. coli* FhuA, and at least *Salmonella paratyphi* can also be hosts to T5 (Graham & Stocker, 1977). T5 side tail fibres are however dispensable, as the T5*hd1* mutant that lacks them is only affected in host adsorption rates (Heller & Bryniok, 1984; Saigo, 1978). *In vitro*, T5 mere interaction with purified FhuA commits T5 to infection and triggers DNA ejection (Boulanger *et al*, 1996), making T5 an excellent model to study phage-host interaction at the cellular and molecular level (Arnaud *et al*, 2017; Linares *et al*, 2022). Both RBP<sub>pb5</sub> and FhuA can be purified and can *in vitro* bind in a 1:1, irreversible complex (Plançon *et al*, 2002; Flayhan *et al*, 2012). Biophysical analysis of the individual FhuA and RBP<sub>pb5</sub> proteins compared to the complex showed that some secondary structure changes occur in RBP<sub>pb5</sub> upon binding to FhuA (Flayhan *et al*, 2012). Also, a low resolution envelop of the proteins obtained by SANS (Small Angle Neutron Scattering) did not show large conformational changes in neither FhuA nor RBP<sub>pb5</sub> upon interaction (Breyton *et al*, 2013), but showed that RBP<sub>pb5</sub> is an 8 nm long elongated protein. We recently solved T5 tail tip structure before and after interaction with FhuA reconstituted into a nanodisc, and built atomic models for all the tail tip proteins except RBP<sub>pb5</sub> and FhuA (Linares *et al*, 2022). These structures allowed us to describe the sequence of events underwent by T5 tail tip upon FhuA-RBP<sub>pb5</sub> binding, namely straight fibre bending, opening of the tail tube, its anchoring to the membrane and finally formation of a transmembrane channel. Thus, unravelling RBP<sub>pb5</sub> structure before and after interaction with FhuA is the last piece of the puzzle to fully understand the infection trigger mechanism of phage T5 following FhuA binding. This structure is the first one of a non-saccharide binding RBP.

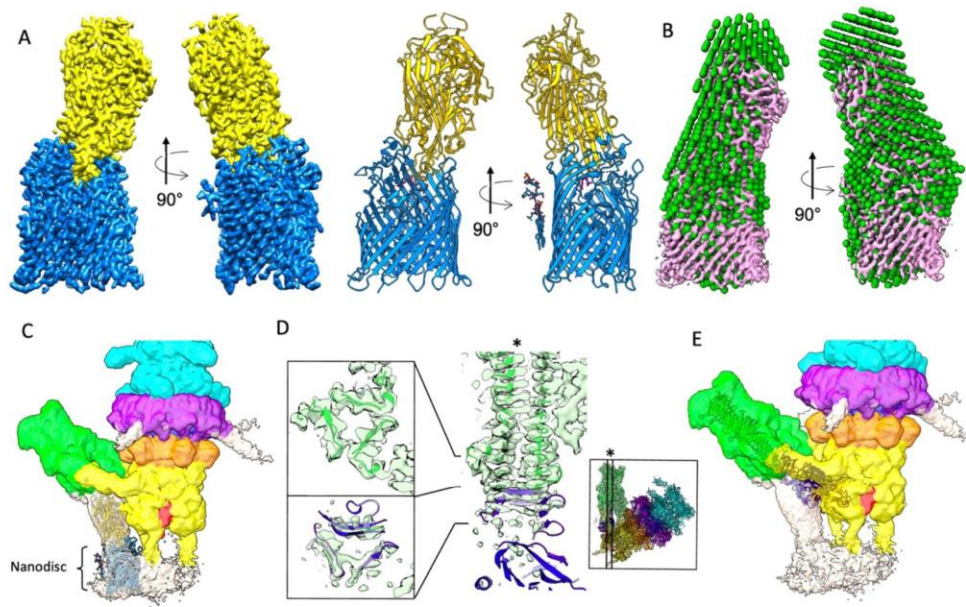
## Results and discussion

### *Overall structure of the FhuA-RBP<sub>pb5</sub> complex*

The FhuA-RBP<sub>pb5</sub> complex being so stable and FhuA having crystallised in different conditions and detergents (*e.g.* (Locher *et al*, 1998; Breyton *et al*, 2019)), we actively tried to crystallise it. We could only obtain non-reproducible anisotropically 8-Å diffracting crystals (Flayhan, 2008). To improve the crystals, we produced nanobodies against the complex. This strategy of increasing the hydrophilic domain of membrane proteins has lead previously to success in crystallisation (Desmyter *et al*, 2015). We could isolate four good binders, and submitted the four FhuA-RBP<sub>pb5</sub>-nanobody complexes to crystallisation. In addition to nanobodies, the lanthanide complex XO4 (nucleating and phasing agent (Engilberge *et al*, 2017)) was used to increase the chances of obtaining well-diffracting crystals. In neither condition could



reproducibly diffracting crystals be obtained. Reproducible 2-Å diffracting 2D crystals were obtained and analysed in electron diffraction and image analysis in collaboration with Mohamed Chami and Henning Stahlberg (Basel)(Flayhan, 2008). The crystals were however too thick, rendering structure determination too complicated. Finally, taking advantage of the ‘resolution revolution’ in electron cryo-microscopy (cryo-EM)(Kühlbrandt, 2014), we investigated directly the FhuA-RBP<sub>pb5</sub> structure as a complex in solution by cryo-EM. Even though the complex is relatively small (150 kDa), we obtained a 2.6-Å resolution 3D structure (Figs. 1A, S1, Table S1).



**Figure 1: Overall structure of isolated FhuA-RBP<sub>pb5</sub> and fit in the Tip-FhuA map.** **A.** Isosurface view of the FhuA-RBP<sub>pb5</sub> complex in two 90° related side views (left), and ribbon representation of the modelled proteins with FhuA in blue and RBP<sub>pb5</sub> in gold (right). **B.** FhuA-RBP<sub>pb5</sub> map fitted into the SANS envelop of FhuA-RBP<sub>pb5</sub> (Breyton *et al*, 2013). **C.** Isosurface view at high contour level of the unmasked and unfiltered cryo-EM map of T5 Tip-FhuA (Linares *et al*, 2022) (EMD-14799) with the structure of FhuA-RBP<sub>pb5</sub> manually fitted into the empty density below pb4 spike and anchoring into the nanodisc. Cyan: density corresponding to pb6, purple: p132, orange: pb9, yellow: pb3, green: pb4 and red: pb2. Unattributed densities are in white. **D.** Manual fit of AF-RBP<sub>pb5</sub> pb4-binding subdomain into the density extending pb4 spike, with pb4 structure shown in ribbon (EMD-14800, PDB 7ZN4). Top left: Section showing the two last β-strands of pb4 spike trimer, bottom left: section of the density extending pb4 spike. The position of the sections is localised on the longitudinal thin section of pb4 spike map and model (right). The position of the latter section is indicated on the Tip-FhuA model by an \* (inset, PDB 7ZN2). **E.** Overall position of AF-RBP<sub>pb5</sub> in the Tip-FhuA map (same as in C), after manual fitting of pb4-binding subdomain as in D.

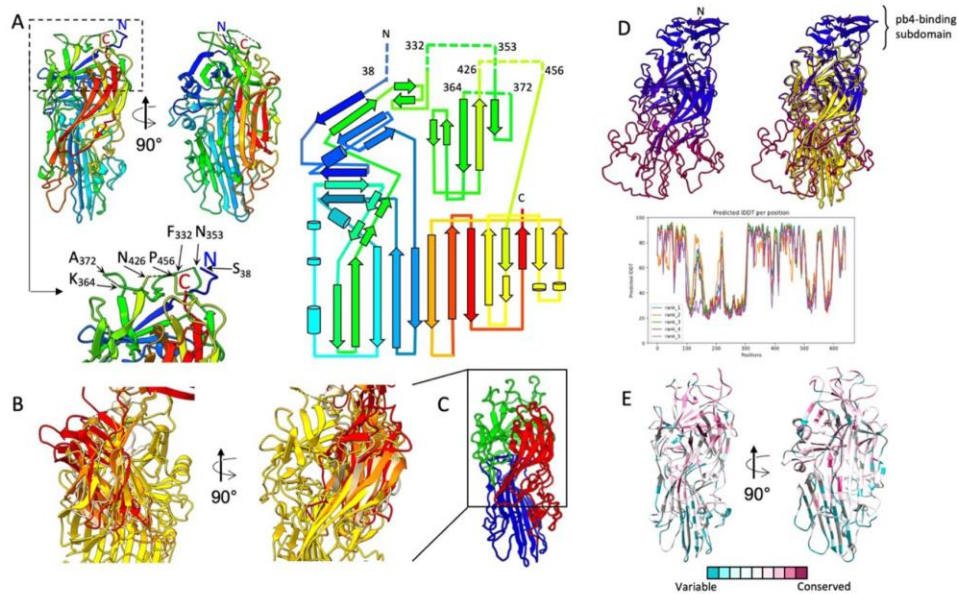
The complex and RBP<sub>pb5</sub> within it have an oblong structure, the interaction surface occurring through the distal end of RBP<sub>pb5</sub> that penetrates into FhuA barrel down to FhuA plug. The overall FhuA-RBP<sub>pb5</sub> structure fits well in the SANS envelop (Breyton *et al*, 2013)(Fig. 1B). It also fits well in the map of T5 tail tip interacting with its receptor reconstituted into a nanodisc (Tip-FhuA, Fig. 1C). In the FhuA-RBP<sub>pb5</sub> structure, both proteins are very well resolved, except for FhuA His<sub>6</sub>-tag located after residue 405 in loop 5 and residues 1-17 including FhuA TonB box. These are disordered in all FhuA structures except when FhuA is crystallised with TonB C-terminal domain (Braun, 2009; Pawelek *et al*, 2006). The lipopolysaccharide that copurifies with FhuA could also be fitted. In RBP<sub>pb5</sub>, the 37 first N-terminal residues, as well as residues 333-352, 365-371 and 427-455 could not be traced in the cryo-EM map. These four sequences point to RBP<sub>pb5</sub> apex, suggesting, from the fit in Tip-FhuA map, that they form a phage-binding subdomain in interaction with pb4 spike C-terminus. It seems that these loops are thus unstructured without the pb4 spike partner. Unfolding of this small subdomain could also be favoured by the presence of the His-tag in C-terminus of RBP<sub>pb5</sub>, which also points in that direction (Fig. 2A,B).

#### *Analysis of the FhuA-RBP<sub>pb5</sub> structure*

##### *RBP<sub>pb5</sub>*

As previously proposed (Flayhan *et al*, 2012), RBP<sub>pb5</sub> folds as a unique domain rich in  $\beta$ -structures. A large and long central 9-stranded  $\beta$ -sheet serves as the spine of the protein. It is surrounded by shorter  $\beta$ -sheets in the apical half and by random coil loops and small  $\alpha$ -helices in the distal half of the protein (Fig. 2A). A DALI analysis (Holm, 2020) does not align RBP<sub>pb5</sub> whole structure with any protein of known structure. However, a  $\beta$ -sandwich subdomain shows structural similarity with other proteins: a close analysis of the aligned DALI domains highlights three proteins with the same topology (Fig. 2B). These three matches are all from phage trimeric proteins, two from RBPs (*S. aureus* phage K gp144, PDB code 5M9F, Dali Z-score = 5.1 and phage T7 gp17, PDB code 4A0U, Dali Z-score = 3.2) and one from a T4 baseplate protein (gp9, PDB code 1QEX, Dali Z-score = 4.9, with however a slight topological difference on the last  $\beta$ -strand). In RBP<sub>pb5</sub>, the  $\beta$ -strands are however longer. This domain displays an original fold, similar however to the knob domain of other viral RBPs, which is believed to interact with the host receptor (Garcia-Doval & van Raaij, 2012). This again illustrates the 'Lego' strategy of phages and virus that use and exchange functional modules (Cardarelli *et al*, 2010). RBP<sub>pb5</sub> structure analysis with the Sword software indeed proposes RBP<sub>pb5</sub> to be composed of three subdomains, including the common RBP subdomain, as first alternative to a unique domain partitioning (Fig. 2C).





**Figure 2: RBP<sub>pb5</sub> structure analysis.** **A.** Top left: ribbon representation of RBP<sub>pb5</sub> structure in the FhuA-RBP<sub>pb5</sub> complex context, in two 90° related side views (left) and topology diagram (right), both rainbow coloured (N-terminal: blue, C-terminal: red, indicated)). Bottom left: blow up of the apical region of RBP<sub>pb5</sub> with indication of the residues bounding the unresolved pb4-binding subdomain. **B.** Dali alignment of RBP<sub>pb5</sub> (gold) with domains of 5M9F (orange), 1QEX (red), 4A0U (beige). **C.** Subdomain partitioning proposed by the Sword software (<https://www.dsimb.inserm.fr/sword/>). The position of the aligned subdomain in B is boxed. **D.** Top: best ranked AF-RBP<sub>pb5</sub> ribbon structure coloured by prediction confidence score (high pLDDT: blue, low pLDDT: dark magenta) alone (left) or aligned with RBP<sub>pb5</sub> structure within the FhuA-RBP<sub>pb5</sub> complex (gold)(right). Pb4-binding subdomain is clearly identified at the apex of AF-RBP<sub>pb5</sub>, unresolved in our cryo-EM structure. Bottom: pLDDT scores of the five AF-RBP<sub>pb5</sub> predictions. **E.** Cryo-EM RBP<sub>pb5</sub> ribbon representation coloured by sequence similarity / divergence (Ashkenazy *et al*, 2016)(colour key as shown) after ClustalO alignment of a Blast search.

Structure prediction software have done tremendous progress over the last few years (Alexander *et al*, 2021), AlphaFold2 in particular (Jumper *et al*, 2021). To investigate the structure of RBP<sub>pb5</sub> missing phage-binding subdomain and to compare isolated RBP<sub>pb5</sub> prediction with our structure, we submitted RBP<sub>pb5</sub> sequence to AlphaFold2 and obtained five different predicted models (AF-RBP<sub>pb5</sub>)(Fig. 2D). For all models, AF-RBP<sub>pb5</sub> structure can be divided into two halves: the distal half predicted with a poor confidence (pLDDT < 50)(pLDDT: predicted local-distance difference test), while the apical half is predicted with high confidence (pLDDT > 70). This includes the N-terminus and disordered loops of RBP<sub>pb5</sub>, which are predicted to form a small  $\beta$ -helix, attached to RBP<sub>pb5</sub> core by five linkers (residues 37-45, 327-336, 353-357, 426-432 and 456-463). This  $\beta$ -helix would extend perfectly pb4 spike: indeed, it fits very well in the density present in Tip-FhuA cryo-EM map extending pb4- $\beta$ -helix (Fig. 1D). However, the rest of AF-RBP<sub>pb5</sub> does not fit at all the rest of the low-resolution density

available for RBP<sub>pb5</sub> (Fig. 1E). Indeed, our Tip-FhuA map rather suggests a different angle between pb4 binding subdomain and the rest of RBP<sub>pb5</sub>, made possible by the flexibility given by the five linkers (see also below). This flexibility, visible on negative stain images (see Fig. 5 in Zivanovic *et al*, 2014) explains in part why cryo-EM image analysis of T5 tip and Tip-FhuA did not generate a high resolution map for RBP<sub>pb5</sub>.

Considering the rest of AF-RBP<sub>pb5</sub> structure prediction, the apical half aligns very well with our RBP<sub>pb5</sub> structure (rmsd ranging from 0.811 to 0.979 Å over 222 to 291 residues depending on the AF model). This is also coherent with the good level of confidence calculated by AlphaFold2. On the other hand, as expected, alignment of the poorly predicted distal half, corresponding to FhuA binding domain, is pretty bad (overall rmsd ranging from 7.6 to 18.2 Å over all 572 residues for the five AF-RBP<sub>pb5</sub> models). pLDDT values below 50 are indicated to be strong predictors of disorder (Tunyasuvunakool *et al*, 2021), suggesting that the regions are unstructured in physiological conditions: they would only fold upon interaction with FhuA. This is consistent with biophysical and biochemical data that allowed to conclude that isolated RBP<sub>pb5</sub> is less structured and sees an increase in its  $\beta$ -sheet content at the expense of other structures, as well as a rigidification of its tertiary structure, upon binding to FhuA (Flayhan *et al*, 2012). This also explains RBP<sub>pb5</sub> strong thermal stabilisation upon binding to FhuA ( $T_m = 43^\circ\text{C}$  for RBP<sub>pb5</sub> alone and  $89^\circ\text{C}$  for FhuA-RBP<sub>pb5</sub>)(Flayhan *et al*, 2012).

We also submitted the FhuA-RBP<sub>pb5</sub> complex to AlphaFold2: the prediction resulted to be far from correct, exhibiting a wrong orientation between FhuA and RBP<sub>pb5</sub> and a very small interacting surface. RBP<sub>pb5</sub> does however appear a bit more structured, in particular its large central sheet (Fig. 3D). We have proposed the FhuA-RBP<sub>pb5</sub> complex to CASP15, let us see the progress they will have made!

Whereas there are no RBP<sub>pb5</sub> homologues in the PDB, a BLAST search fishes a large number of proteins, with identities ranging from 29.0 to 96.3%. They are described as *E. coli*, *Salmonella*, *Shigella* or *Klebsiella* phage hypothetical RBPs or RBPs. This is consistent with the fact that the T5-like genus is largely represented in the databases. Clustal alignment and the ConSurf Server (Ashkenazy *et al*, 2016) allow to visualise the distribution of conserved vs. variable residues on RBP<sub>pb5</sub> structure (Fig. 2E). The dichotomy of distribution is quite striking: the apical half of RBP<sub>pb5</sub> is largely conserved, in particular for the unresolved N-terminus and loops, while the distal half, which is proposed by AlphaFold2 to be largely unstructured when unbound to FhuA and which mediates FhuA interaction, is much more variable. This trend – phage binding domain conserved / receptor binding domain much more variable – is quite general in RBPs, fibres or TSP, highlighting that the receptor binding domains are under heavy evolutionary pressure to diversify as they are in interaction with a large



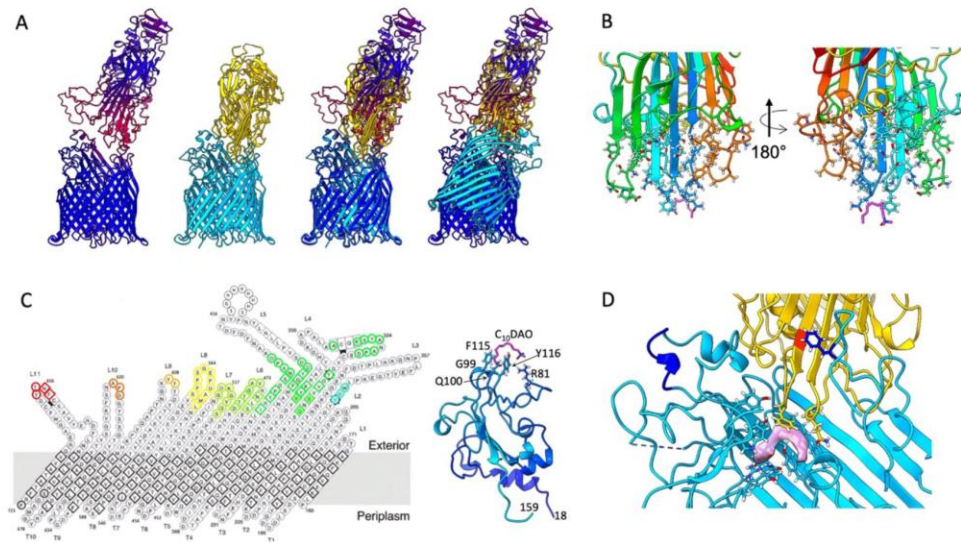
variety of host targets, while the phage binding domain, bearing a structural function, is more constrained by interaction with its phage partner-s (Jonge *et al*, 2019).

### *FhuA*

The structure of FhuA has been solved in its apo form, in complex with many ligands and with the C-terminal domain of TonB (Braun, 2009). The different structures are remarkably similar, with rmsd between the eleven different structures below 1.93 Å over all 696 common residues. The main difference between the different structures is the N-terminal: in FhuA apo structures, residues 19-30 are structured as an  $\alpha$ -helix, whereas in structures binding TonB-dependant ligands it adopts a conformation which is superimposable to the TonB bound conformation (Braun, 2009). In agreement with the fact that T5 infection is TonB independent, the conformation of RBP<sub>pb5</sub>-bound FhuA residues 19-30 is that of apo FhuA.

### *FhuA-RBP<sub>pb5</sub> interface*

RBP<sub>pb5</sub> interaction with FhuA occurs through the distal  $\beta$ -turns and loops of the central sheet and through random coil loops and small  $\alpha$ -helices in the distal half of RBP<sub>pb5</sub>, with a total of 63 residues involved, as calculated by PISA (Krissinel & Henrick, 2004)(Fig. 3A). In FhuA, interactions are mediated by the apical tip of the plug, residues in loops 4-11 and in  $\beta$ -strands 7-12, with a total of 64 residues involved (Fig. 3B). The surface of interaction is large, with a buried area of 2,034 Å<sup>2</sup>, an energy of -21.4 kcal/mol and it is stabilised by 13 hydrogen bonds and hydrophobic interactions, explaining the quasi-irreversibility of the interaction (Flayhan *et al*, 2012). Interestingly, a C<sub>10</sub>DAO (n-Decyl N,N-dimethylamine N-oxide) molecule, contained in the buffer, could be modelled at the interface between FhuA plug and RBP<sub>pb5</sub> (Figs. 3C, S1E). It overlaps with ferrichrome hydrophobic binding pocket and interacts, on FhuA side, with 13 residues – four from the plug, four from loops 3 and 4, and five from  $\beta$ -strands 7 to 9, with a hydrogen bond between Y116 and the carboxyl of the aminoxyde head group. The C<sub>10</sub>DAO binding pocket is further completed by four residues from loops of the tip of RBP<sub>pb5</sub>  $\beta$ -sheet, further stabilising the interaction between the two proteins (Fig. 3C).



**Figure 3: Surface of interaction between FhuA and RBP<sub>pb5</sub>.** **A.** From left to right: AF-FhuA-RBP<sub>pb5</sub> best prediction, coloured by prediction confidence score (high pLDDT: blue, low pLDDT: dark magenta), cryo-EM FhuA-RBP<sub>pb5</sub> structure (FhuA: cyan, RBP<sub>pb5</sub>: gold), alignment of AF-FhuA-RBP<sub>pb5</sub> on FhuA and on RBP<sub>pb5</sub> of the cryo-EM structure. **B.** Two orientations of a rainbow-coloured ribbon representation of RBP<sub>pb5</sub> distal part, with residues involved in the interactions, as determined by PISA, shown in sticks. The C<sub>10</sub>DAO molecule is shown in pink. **C.** Left: topology diagram of FhuA barrel (from Locher *et al*, 1998), on which are indicated the residues involved in the interaction (dotted line and filled symbols: those interacting exclusively with the C<sub>10</sub>DAO). Right: ribbon diagram of FhuA plug, with residues involved in the interactions and the C<sub>10</sub>DAO molecule shown in sticks. **D.** Zoom on the FhuA-RBP<sub>pb5</sub> interface. FhuA is in cyan, RBP<sub>pb5</sub> in gold. Both proteins are in ribbon, with the residues involved in C<sub>10</sub>DAO binding in sticks. The C<sub>10</sub>DAO density is shown in pink. FhuA-PADKGH is highlighted in blue. In RBP<sub>pb5</sub>, G166 is highlighted in orange and FhuA-F566 in blue sticks.

T5 is a historic phage and a prototype of siphophages, for which host binding has been extensively studied in the past. The presence of an RBP binding to FhuA has been proposed while studying the *oad* mutant that had a reduced affinity to FhuA (Heller & Bryniok, 1984; Krauel & Heller, 1991). Later, Mondigler *et al*, 1996 identified the *oad* mutation: Gly166 is switched to a Trp. In our structure, Gly166 participates in the binding interface, making C-H... $\Pi$  interaction with FhuA Phe566 (Fig. 3C). A Gly->Trp mutation would thus indeed destabilise the protein-protein interaction interface. Mondigler *et al* also investigated the ability of different RBP<sub>pb5</sub> deletion and insertion mutants to bind to FhuA. Apart from one (deletion of the 153 C-terminus residues), none retained binding abilities, which is explained by the tight fold of FhuA-bound RBP<sub>pb5</sub>.

FhuA binding to RBP<sub>pb5</sub> has also been studied from the FhuA perspective: before FhuA structure was known, Killmann *et al*, 1995 investigated T5 binding loops in FhuA by competitive peptide mapping. They identified a peptide –<sub>334</sub>PADKGH<sub>339</sub>–

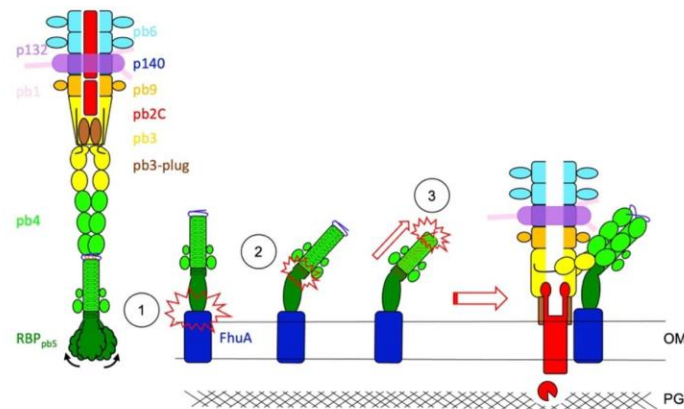


of FhuA loop 4 that induced DNA ejection when incubated with T5, in a temperature dependant manner. Given the large interaction area and the many loops involved to stabilise RBP<sub>pb5</sub> bound conformation, it is surprising that interaction of RBP<sub>pb5</sub> with such a small peptide would induce the large conformational changes resulting in DNA ejection. Furthermore, as seen in our structure, this sequence does not appear to be involved in RBP<sub>pb5</sub> binding (Fig. 3C). We could not reproduce the results with a slightly longer peptide (APADKGHY). Point mutation impairing T5 binding could not be isolated unless they perturb FhuA overall structure (Braun, 2009) and deletion of loops three to eleven, except for loop 8, did not prevent T5 binding (Endriss & Braun, 2004). This can be explained by the large interacting surface between the two proteins, which would compensate point mutations or loop deletions. There does not seem to be a rationale for loop 8 being more important for RBP<sub>pb5</sub> binding than the other ones. Deletion of this loop also impairs T1 and  $\Phi$ 80 binding, and ferrichrome transport, suggesting its deletion might rather prevent FhuA proper folding. A FhuA plug-less mutant was also investigated for its ability to bind T5: an *E. coli* strain deleted for endogenous FhuA and expressing FhuA $\Delta$ plug is resistant to T5, thus RBP<sub>pb5</sub> does not bind to FhuA $\Delta$ plug (Braun *et al*, 2003). Given the surface of interaction provided by FhuA barrel loops, this result suggests that the latter are either flexible or adopt a different conformation in FhuA $\Delta$ plug than in WT FhuA, preventing RBP<sub>pb5</sub> binding. Indeed, the plug does not provide a larger surface of interaction than individual FhuA barrel loops. The role of the plug is however important in the overall stabilisation of the interaction between FhuA and RBP<sub>pb5</sub>, as previously suggested (Flayhan *et al*, 2012): whereas FhuA displays two T<sub>m</sub>, one for the plug (64°C) and one for the barrel (75°C), the complex displays a unique T<sub>m</sub> which is shifted to 89°C.

T5 is described as infecting *E. coli*; however, some other strains also bear FhuA. Thus, we tested T5 sensitivity on different *E. coli*, *Shigella* and *Salmonella* strains, which were chosen because of the different level of FhuA identity with *E. coli* lab strain FhuA, ranging from 96 to 100% (Fig. S2). T5 sensitivity tests were performed in the presence of dipyriddy, and iron chelator that boosts FhuA expression. Surprisingly, only *Salmonella bongori* is sensitive to T5. FhuA sequence alignment shows however some differences in the residues interacting with RBP<sub>pb5</sub>, confirming the resilience of the FhuA-RBP<sub>pb5</sub> interaction. We checked T5 irreversible adsorption on all strains, but only *Salmonella bongori* showed adsorption, comparable to that of *E. coli*. Interestingly, *E. coli* EPEC and the *Shigella* strains are resistant to T5 despite no difference in the interacting residues (Fig. S2). Resistance to T5 may be explained by FhuA being blocked by a superinfection exclusion protein coded by a prophage in their genome or by the synthesis of a capsule, shielding FhuA from the phage.

## Conclusion

We present here the first structure of an RBP, RBP<sub>pb5</sub>, in complex with its protein receptor, FhuA. Analysis of this structure compared to that predicted of RBP<sub>pb5</sub> and fitted in the T5 tail tip map after interaction with FhuA (Tip-FhuA)(Linares *et al*, 2022) allow us to propose a mechanism that would transmit FhuA-binding to pb4 through RBP<sub>pb5</sub> (Fig. 5): the apical subdomain of AF-RBP<sub>pb5</sub> shows a small structured  $\beta$ -helix that fits well in our Tip-FhuA map, in the densities extending pb4  $\beta$ -helix spike. This RBP<sub>pb5</sub> subdomain would serve as an adaptor between the trimeric pb4 and the monomeric RBP<sub>pb5</sub>. The presence of linkers between this subdomain and RBP<sub>pb5</sub> core confers flexibility between these two parts of the protein. AF-RBP<sub>pb5</sub> predicted structure and biophysical analysis of RBP<sub>pb5</sub> (Flayhan *et al*, 2012) suggest that the distal half of RBP<sub>pb5</sub> is unstructured when unbound to FhuA. Upon binding to FhuA, RBP<sub>pb5</sub> distal half is structured, rigidifying the protein. In our Tip-FhuA map, RBP<sub>pb5</sub> is stabilized with a  $\sim 45^\circ$  angle between pb4-binding subdomain and the protein core. We suggest that rigidification of RBP<sub>pb5</sub> distal half upon FhuA binding is transmitted to the pb4-binding subdomain, imposing a kink within RBP<sub>pb5</sub>. This rigidification and kink would constrain RBP<sub>pb5</sub> pb4-binding subdomain, in turn constraining pb4  $\beta$ -helix spike, inducing the different twist in pb4  $\beta$ -helix that initiates the conformational change domino resulting in tail tube opening (Linares *et al*, 2022). Definite confirmation of this hypothesis will require the structure of pb4-bound RBP<sub>pb5</sub> before and after interaction with FhuA.



**Figure 4: Proposed mechanism of signal transduction of FhuA binding to pb4 spike through RBP<sub>pb5</sub>.** Left: within T5 tail, RBP<sub>pb5</sub> distal half is unstructured and RBP<sub>pb5</sub> core is flexible with respect to the pb4-binding subdomain. Upon binding to FhuA, RBP<sub>pb5</sub> distal half gets structured (1). This would induce rigidification of the whole protein, in particular the linkers between the core of RBP<sub>pb5</sub> and pb4-binding subdomain, which is stabilised at an angle of  $\sim 45^\circ$  (2). This in turn would induce a modification of the twist of pb4  $\beta$ -helix (3), which induces the opening of the tail tube, its anchoring to the membrane, the expulsion of the Tape Measure Protein and the formation of a channel across the host outer-membrane (Linares *et al*, 2022).



## Materials and Methods

**Bacterial and phage strains.** The heat-stable deletion mutant T5st0 was used for sensitivity and adsorption test. *E. coli* F, a fast-adsorbing strain for T5, was used for the production of T5st0 and for control of free T5 in adsorption tests. *Shigella flexneri* (CIP 106171), *Shigella sonnei* (CIP 106204), *Salmonella enterica serovar* Typhimurium (CIP 104474), *Salmonella enterica indica* (CIP 102501T), *Salmonella bongori* (CIP 82.33T) and *E. coli* EPEC (Enteropathogen *E. coli*)(CIP 52.170) were used for sensitivity to T5 and adsorption tests. *E. coli* strain AW740 and BL21(DE3) were used for FhuA and RBP<sub>pb5</sub> expression respectively.

**FhuA and RBP<sub>pb5</sub> production and purification** – FhuA was purified from the *E. coli* strain AW740 transformed with a plasmid encoding the *fhuA* gene in which a His-tag coding sequence has been inserted in extracellular loop L5. RBP<sub>pb5</sub> was purified from the *E. coli* strain BL21(DE3) carrying the *oad* gene encoding RBP<sub>pb5</sub> fused to a His-tag coding sequence in 3' of *oad*, in a pET-28 vector (Plançon *et al*, 2002). Cells were grown in LB medium at 37 °C in the presence of 100 µM of the iron-chelating agent dipyrindine for AW740 cells and at 20 °C without induction for BL21(DE3) cells. FhuA and RBP<sub>pb5</sub> purifications were carried out as described (Flayhan *et al*, 2012). The FhuA-RBP<sub>pb5</sub> complex was formed by adding equimolar amounts of the two proteins, which results in 100% complex formation as described in (Plançon *et al*, 2002). Detergent exchange on the FhuA-RBP<sub>pb5</sub> complex was performed by a 10 times dilution in deionised water and centrifugation at 100,000 *g* for 45 min at 4°C. The pellet was rinsed with water and resuspended in 20 mM Tris, 1.6% C<sub>10</sub>DAO at a protein concentration of 4.3 mg/mL.

**Cryo-EM sample preparation** – Typically, 3.5 µL of the FhuA-RBP<sub>pb5</sub> complex were deposited on a freshly glow discharged (25 mA, 30 sec) Cu/Rh 400 mesh Quantifoil R 2/1 EM grids and flash-frozen in nitrogen-cooled liquid ethane using a ThermoFisher Mark IV Vitrobot device (100% humidity, 20°C, 2 s blotting time, blot force 1). Preliminary screening of freezing conditions has been performed on an F20 electron microscope.

Two different datasets were acquired on the same grid. For both dataset, 60-frame movies with a total dose of 60 e<sup>-</sup>/Å<sup>2</sup> were acquired on a ThermoFisher Scientific Titan Krios G3 TEM (European Synchrotron Radiation Facility, Grenoble, France) (Kandiah *et al*, 2019) operated at 300 kV and equipped with a Gatan Quantum LS/967 energy filter (slit width of 20 eV used) coupled to a Gatan K2 summit direct electron detector. Automated data collection was performed using ThermoFisher Scientific EPU software. A nominal magnification of 130.000x was used, resulting in a calibrated pixel size at the specimen level of 1.052 Å/pixel. For the first dataset, 777 movies were acquired with a phase plate, close to focus (between -0.5 and -1.0 µm) while, for the second dataset, 8752 movies were acquired without phase plate and with a defocus ranging between -1.2 and -2.6 µm

**EM image processing** – For both datasets, the processing was done with Relion (Zivanov *et al*, 2018). Motion-correction using 5×5 patches was done with Motioncor2 (Zheng *et al*, 2017) while CTF estimation was done with Gctf (Zhang, 2016).

For dataset 1, an initial set of particles was selected with the Laplacian of Gaussian (LoG) picker in Relion. After 2D classification with 2 times binned particles, nicely defined 2D classes were used as references for a template-based picking. Following another 2D classification, an initial model was obtained in Relion, which was then used to get a first 3D reconstruction. 3D classification (without a mask) was performed to select the best particles which were then used to compute a final 3D reconstruction at 6.7 Å (at FSC = 0.143) into which the atomic model of FhuA (PDB 1QFG) could be un-ambiguously fitted (not shown).

For dataset 2, 2D projections of the 3D model obtained from dataset 1 were used as a template to automatically pick all micrographs. 2,191,586 particles were obtained, binned two times and further subjected to two rounds of 2D classification to remove obvious outliers. A first 3D classification was done without any mask and the selected particles were re-extracted without binning. Following a 3D refinement step, two successive 3D classifications with a mask following the contour of the complex were performed without any image alignment, leading to a homogenous subset of 109,350 particles. Then, particle polishing and two rounds of CTF refinements (magnification anisotropy, beam tilt and per-particle defocus) were performed before a last 3D refinement yielded the final 3D reconstruction at 2.6 Å average resolution (at FSC = 0.143). The final map was sharpened using DeepEMhancer (Sanchez-Garcia *et al*, 2021).

#### *Model building*

The RBP<sub>pb5</sub> protein model was built *de novo* in FhuA-RBP<sub>pb5</sub> cryo-EM map using Coot (version 0.9.2)(Morin *et al*, 2013). FhuA was adapted from a FhuA structure solved by X-ray crystallography (PDB 2GRX). The two individual proteins and the complex were then refined using PHENIX (version 1.18.2-3874) Real Space Refine tool (Morin *et al*, 2013). Structure validation was done using the Molprobitry online tool.

#### *Phage sensitivity tests*

*Shigella*, *Salmonella* and *E. coli* strains were grown in LB medium at 37°C, 180 rpm until an OD<sub>600nm</sub> of 0.5. 1.2 ml of LB soft agar are inoculated with 300 µL of culture, and spread on an LB Petri dish. Serial dilutions of phage T5, at 0 to 10<sup>9</sup> PFU/mL, were directly spotted on the soft agar. The titre was estimated by counting lysis plaques in the spotted area after overnight incubation at 37°C.

#### *Irreversible phage adsorption tests*

To 5 mL of each strain grown in LB medium at 0.2 OD<sub>600nm</sub> is added 0.01 M CaCl<sub>2</sub>, 0.04 M MgCl<sub>2</sub> and T5 at a multiplicity of infection of 0.5 and incubated at 37°C. 200 µL of infected cells are collected at different times, vigorously vortexed for 10 s and centrifuged 10 min at 15,000 *g*. The supernatant containing free phages (unabsorbed and reversibly bound particles) are titrated on *E.coli* F using the double layer agar method described above.

## **Acknowledgments**

We would like to acknowledge the thesis work of Ali Flayhan and the shorter or longer internship work of Flavien Grégoire, Charles-Adrien Arnaud, Marie-Ange Marrel, Sarra Landri, Emmi Mikkola and Annelise Vermot who tried to optimise RBP<sub>pb5</sub> construct or crystallise the complex with nanobodies, and Edine Hammouche who performed some T5 adsorption tests. We also acknowledge Mohamed Chami and Henning Stahlberg (Biozentrum, Basel) for the 2D



crystal adventure, and Alain Roussel and Aline Desmyter (Marseille, <https://nabgen.org/>) for production of nanobodies directed against the FhuA-RBP<sub>pb5</sub> complex. We thank Eric Faudry for a kind gift of the *Salmonella typhidium*, *Shigella flexneri*, and *E. coli* EPEC strain, Daphna Fenel for quality control by negative staining and Emmanuelle Neumann for help with the F20 electron microscope.

This research was funded by the Agence Nationale de la Recherche, grant numbers ANR-16-CE11-0027 and ANR-21-CE11-0023. It used the EM facility at the Grenoble Instruct-ERIC Center (ISBG; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB). IBS platform access was supported by FRISBI (ANR-10-INBS-05-02) and GRAL, a project of the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche), CBH-EUR-GS (ANR-17-EURE-0003). The electron microscope facility is supported by the Auvergne-Rhône-Alpes Region, the Fondation pour la Recherche Médicale (FRM), the Fonds FEDER and the GIS-Infrastructures en Biologie Santé et Agronomie (IBISA). We acknowledge the provision of in-house experimental time from the CM01 facility at the ESRF. IBS acknowledges integration into the Interdisciplinary Research Institute of Grenoble (IRIG, CEA).

**Disclosure and Competing Interests Statement:** The authors declare no conflict of interest.

### Author Contributions

CB conceived the project. SD prepared the FhuA-RBP<sub>pb5</sub> complex. GS and RL optimised cryo-grid preparation. GE recorded and processed the cryo-EM data. SD performed T5 sensitivity and adsorption experiments, and built the atomic model with help from RL. CB analysed and interpreted the structure. CB wrote the paper with contributions from GE, GS and SD. All authors contributed to the editing of the manuscript.

### Data Availability

Cryo-EM density maps of FhuA-RBP<sub>pb5</sub> and the associated atomic coordinate have been respectively deposited in the Electron Microscopy Data Bank (EMDB) and in the Protein Data Bank (PDB) under the following accession codes: PDB 8B14, EMD 15802.

### References

- Alexander LT, Lepore R, Kryshtafovych A, Adamopoulos A, Alahuhta M, Arvin AM, Bomble YJ, Böttcher B, Breyton C, Chiarini V, *et al* (2021) Target highlights in CASP14: Analysis of models by structure providers. *Proteins* 89: 1647–1672
- Arnaud C-A, Effantin G, Vivès C, Engilberge S, Bacia M, Boulanger P, Girard E, Schoehn G & Breyton C (2017) Bacteriophage T5 tail tube structure suggests a trigger mechanism for Siphoviridae DNA ejection. *Nat Commun* 8: 1953
- Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T & Ben-Tal N (2016) ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Research* 44: W344–W350
- Boulanger P, le Maire M, Bonhivers M, Dubois S, Desmadril M & Letellier L (1996) Purification and structural and functional characterization of FhuA, a transporter of the *Escherichia coli* outer membrane. *Biochemistry* 35: 14216–14224
- Braun M, Endriss F, Killmann H & Braun V (2003) In vivo reconstitution of the FhuA transport protein of *Escherichia coli* K-12. *J Bacteriol* 185: 5508–5518
- Braun V (2009) FhuA (TonA), the career of a protein. *J Bacteriol* 191: 3431–3436
- Breyton C, Flayhan A, Gabel F, Lethier M, Durand G, Boulanger P, Chami M & Ebel C (2013) Assessing the Conformational Changes of pb5, the Receptor-binding Protein of Phage T5, upon Binding to Its *Escherichia coli* Receptor FhuA. *J Biol Chem* 288: 30763–30772

- Breyton C, Javed W, Vermot A, Arnaud C-A, Hajjar C, Dupuy J, Petit-Hartlein I, Le Roy A, Martel A, Thépaut M, *et al* (2019) Assemblies of lauryl maltose neopentyl glycol (LMNG) and LMNG-solubilized membrane proteins. *Biochim Biophys Acta Biomembr* 1861: 939–957
- Broeker NK & Barbirz S (2017) Not a barrier but a key: How bacteriophages exploit host's O-antigen as an essential receptor to initiate infection. *Mol Microbiol* 105: 353–357
- Cardarelli L, Pell LG, Neudecker P, Pirani N, Liu A, Baker LA, Rubinstein JL, Maxwell KL & Davidson AR (2010) Phages have adapted the same protein fold to fulfill multiple functions in virion assembly. *Proc Natl Acad Sci U S A* 107: 14384–14389
- Dams D, Brøndsted L, Drulis-Kawa Z & Briers Y (2019) Engineering of receptor-binding proteins in bacteriophages and phage tail-like bacteriocins. *Biochemical Society Transactions* 47: 449–460
- Demerec M & Fano U (1945) Bacteriophage-Resistant Mutants in Escherichia Coli. *Genetics* 30: 119–136
- Desmyter A, Spinelli S, Roussel A & Cambillau C (2015) Camelid nanobodies: killing two birds with one stone. *Curr Opin Struct Biol* 32: 1–8
- Dion MB, Oechslein F & Moineau S (2020) Phage diversity, genomics and phylogeny. *Nat Rev Microbiol* 18: 125–138
- Endriss F & Braun V (2004) Loop deletions indicate regions important for FhuA transport and receptor functions in Escherichia coli. *J Bacteriol* 186: 4818–4823
- Engilberge S, Riobé F, Pietro SD, Lassalle L, Coquelle N, Arnaud C-A, Pitrat D, Mulatier J-C, Madern D, Breyton C, *et al* (2017) Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chem Sci* 8: 5909–5917
- Filik K, Szermer-Olearnik B, Oleksy S, Brykała J & Brzozowska E (2022) Bacteriophage Tail Proteins as a Tool for Bacterial Pathogen Recognition-A Literature Review. *Antibiotics (Basel)* 11: 555
- Flayhan A (2008) Etude strctutale du complexe FhuA-pb5Δ1
- Flayhan A, Wien F, Paternostre M, Boulanger P & Breyton C (2012) New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its Escherichia coli receptor FhuA. *Biochimie* 94: 1982–1989
- Garcia-Doval C, Castón JR, Luque D, Granell M, Otero JM, Llamas-Saiz AL, Renouard M, Boulanger P & van Raaij MJ (2015) Structure of the Receptor-Binding Carboxy-Terminal Domain of the Bacteriophage T5 L-Shaped Tail Fibre with and without Its Intra-Molecular Chaperone. *Viruses* 7: 6424–6440
- Garcia-Doval C & van Raaij MJ (2012) Structure of the receptor-binding carboxy-terminal domain of bacteriophage T7 tail fibers. *Proc Natl Acad Sci USA* 109: 9390–9395
- González-García VA, Pulido-Cid M, Garcia-Doval C, Bocanegra R, van Raaij MJ, Martín-Benito J, Cuervo A & Carrascosa JL (2015) Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *J Biol Chem* 290: 10038–10044
- Goulet A, Spinelli S, Mahony J & Cambillau C (2020) Conserved and Diverse Traits of Adhesion Devices from Siphoviridae Recognizing Proteinaceous or Saccharidic Receptors. *Viruses* 12
- Graham AC & Stocker BA (1977) Genetics of sensitivity of Salmonella species to colicin M and bacteriophages T5, T1, and ES18. *J Bacteriol* 130: 1214–1223
- Heller K & Braun V (1982) Polymannose O-antigens of Escherichia coli, the binding sites for the reversible adsorption of bacteriophage T5+ via the L-shaped tail fibers. *J Virol* 41: 222–227
- Heller KJ & Bryniok D (1984) O antigen-dependent mutant of bacteriophage T5. *J Virol* 49: 20–5
- Holm L (2020) DALI and the persistence of protein shape. *Protein Sci* 29: 128–140
- Hu B, Margolin W, Molineux IJ & Liu J (2013) The bacteriophage t7 virion undergoes extensive structural remodeling during infection. *Science* 339: 576–579
- Hu B, Margolin W, Molineux IJ & Liu J (2015) Structural remodeling of bacteriophage T4 and host membranes during infection initiation. *Proc Natl Acad Sci USA* 112: E4919–4928
- Jonge PA de, Nobrega FL, Brouns SJJ & Dutilh BE (2019) Molecular and Evolutionary Determinants of Bacteriophage Host Range. *Trends in Microbiology* 27: 51–63

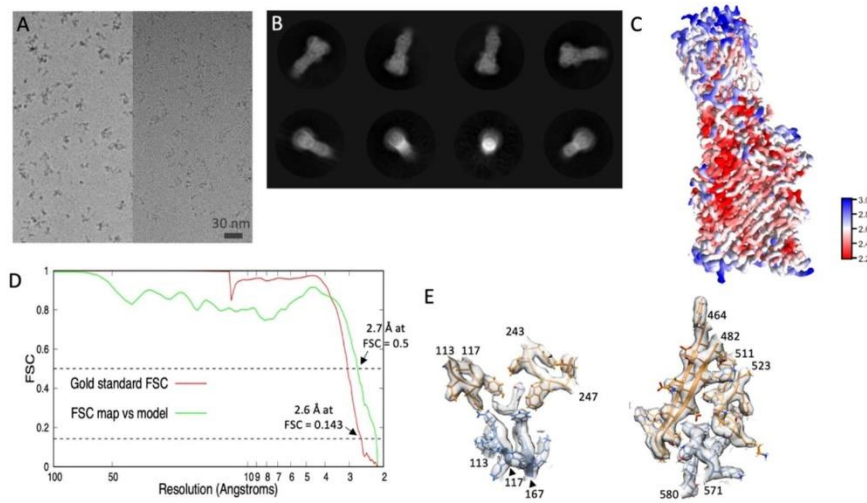


- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, *et al* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596: 583–589
- Kandiah E, Giraud T, de Maria Antolinis A, Dobias F, Effantin G, Flot D, Hons M, Schoehn G, Susini J, Svensson O, *et al* (2019) CM01: a facility for cryo-electron microscopy at the European Synchrotron. *Acta Crystallogr D Struct Biol* 75: 528–535
- Killmann H, Videnov G, Jung G, Schwarz H & Braun V (1995) Identification of receptor binding sites by competitive peptide mapping: phages T1, T5, and phi 80 and colicin M bind to the gating loop of FhuA. *J Bacteriol* 177: 694–8
- Krauel V & Heller KJ (1991) Cloning, sequencing, and recombinational analysis with bacteriophage BF23 of the bacteriophage T5 oad gene encoding the receptor-binding protein. *J Bacteriol* 173: 1287–97
- Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60: 2256–2268
- Kühlbrandt W (2014) Biochemistry. The resolution revolution. *Science* 343: 1443–1444
- Linares R, Arnaud C-A, Degroux S, Schoehn G & Breyton C (2020) Structure, function and assembly of the long, flexible tail of siphophages. *Curr Opin Virol* 45: 34–42
- Linares R, Arnaud C-A, Effantin G, Darnault C, Epalle NH, Erba EB, Schoehn G & Breyton C (2022) Structural basis of bacteriophage T5 infection trigger and *E. coli* cell wall perforation. 2022.09.20.507954 doi:10.1101/2022.09.20.507954 [PREPRINT]
- Locher KP, Rees B, Koebnik R, Mitschler A, Moulinier L, Rosenbusch JP & Moras D (1998) Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell* 95: 771–8
- Mondigler M, Holz T & Heller KJ (1996) Identification of the receptor-binding regions of pb5 proteins of bacteriophages T5 and BF23. *Virology* 219: 19–28
- Morin A, Eisenbraun B, Key J, Sanschagrin PC, Timony MA, Ottaviano M & Sliz P (2013) Collaboration gets the most out of software. *Elife* 2: e01456
- Nobrega FL, Vlot M, de Jonge PA, Dreesens LL, Beaumont HJE, Lavigne R, Dutilh BE & Brouns SJJ (2018) Targeting mechanisms of tailed bacteriophages. *Nat Rev Microbiol* 16: 760–773
- Pawelek PD, Croteau N, Ng-Thow-Hing C, Khursigara CM, Moiseeva N, Allaire M & Coulton JW (2006) Structure of TonB in complex with FhuA, *E. coli* outer membrane receptor. *Science* 312: 1399–1402
- Plançon L, Janmot C, le Maire M, Desmadril M, Bonhivers M, Letellier L & Boulanger P (2002) Characterization of a high-affinity complex between the bacterial outer membrane protein FhuA and the phage T5 protein pb5. *J Mol Biol* 318: 557–569
- Saigo K (1978) Isolation of high-density mutants and identification of nonessential structural proteins in bacteriophage T5; dispensability of L-shaped tail fibers and a secondary major head protein. *Virology* 85: 422–433
- Sanchez-Garcia R, Gomez-Blanco J, Cuervo A, Carazo JM, Sorzano COS & Vargas J (2021) DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun Biol* 4: 874
- Santos SB, Costa AR, Carvalho C, Nóbrega FL & Azeredo J (2018) Exploiting Bacteriophage Proteomes: The Hidden Biotechnological Potential. *Trends Biotechnol* 36: 966–984
- Sørensen AN, Woudstra C, Sørensen MCH & Brøndsted L (2021) Subtypes of tail spike proteins predicts the host range of Ackermannviridae phages. *Comput Struct Biotechnol J* 19: 4854–4867
- Suttle CA (2007) Marine viruses—major players in the global ecosystem. *Nat Rev Microbiol* 5: 801–812
- Taylor NMI, van Raaij MJ & Leiman PG (2018) Contractile injection systems of bacteriophages and related systems. *Mol Microbiol* 108: 6–15
- Tunyasuvunakool K, Adler J, Wu Z, Green T, Zielinski M, Židek A, Bridgland A, Cowie A, Meyer C, Laydon A, *et al* (2021) Highly accurate protein structure prediction for the human proteome. *Nature* 596: 590–596
- Yap ML, Klose T, Arisaka F, Speir JA, Veessler D, Fokine A & Rossmann MG (2016) Role of bacteriophage T4 baseplate in regulating assembly and infection. *Proc Natl Acad Sci USA* 113: 2654–2659

- Zhang K (2016) Gctf: Real-time CTF determination and correction. *Journal of Structural Biology* 193: 1–12
- Zheng W, Wang F, Taylor NMI, Guerrero-Ferreira RC, Leiman PG & Egelman EH (2017) Refined Cryo-EM Structure of the T4 Tail Tube: Exploring the Lowest Dose Limit. *Structure* 25: 1436-1441.e2
- Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E & Scheres SH (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* 7: e42166
- Zivanovic Y, Confalonieri F, Ponchon L, Lurz R, Chami M, Flayhan A, Renouard M, Huet A, Decottignies P, Davidson AR, *et al* (2014) Insights into bacteriophage T5 structure from analysis of its morphogenesis genes and protein components. *J Virol* 88: 1162–1174



## Supplementary Figures and Tables



**Figure S1:** **A.** Cryo-EM field of views of FhuA-RBP<sub>pb5</sub> molecules with (left) and without (right) phase plate. Scale bar: 30 nm. **B.** Exemplary 2D class averages of FhuA-RBP<sub>pb5</sub> obtained from the dataset without phase plate. **C.** Local resolution map of the 3D reconstruction of FhuA-RBP<sub>pb5</sub> obtained with the dataset without phase plate. **D.** Fourier Shell Correlation (FSC) curves: gold standard FSC between two independent 3D reconstructions of FhuA-RBP<sub>pb5</sub> (red) and FSC between the cryo-EM coulomb potential map and the refined atomic model (green). The two dotted horizontal lines represent FSC = 0.143 and 0.5 which are used as cutoffs to determine the resolutions between the two different sets of maps. **E.** Illustrations of the quality of the obtained 3D reconstruction and atomic model. The coulomb potential map is in transparent gray while FhuA, RBP<sub>pb5</sub> and C<sub>10</sub>DAO are colored, orange, blue and grey respectively. The left panel is centered around the interaction of FhuA and RBP<sub>pb5</sub> with C<sub>10</sub>DAO while the right panel represents one of the interacting areas between FhuA and RBP<sub>pb5</sub> centered around loop 571-580 of RBP<sub>pb5</sub>.



**Figure S2** : Sequence alignment of FhuA in *E.coli*, *Shigella* and *Salmonella* strains. The alignment was performed with Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) and visualised with ESPrnt (<https://esprnt.ibcp.fr/ESPrnt/ESPrnt/>). Conserved residues are highlighted in red. Stars show residues involved in the FhuA-RBP<sub>pb5</sub> interaction. Residues interacting exclusively with RBP<sub>pb5</sub> (full), with both C10DAO and RBP<sub>pb5</sub> (empty) or with C10DAO only (hatched) are coloured with the same colour code as in Figure 4. Note that the 9 residues of FhuA His-tag are not included in the alignment nor in the numbering.

**Table S1:** Summary of data collection and atomic model statistics

<b>Data collection</b>	
Microscope	Titan Krios G3
Voltage (kV)	300
Magnification	130,000 x
Unbinned pixel size	1.052 Å/pixel
Camera	K2 Summit (Gatan Inc)
Exposure time	8 s
Number of frames	60
Total dose (e <sup>-</sup> /Å <sup>2</sup> )	60
<b>Image processing</b>	
EMDB	
Rotational symmetry	C1
Final number of Particles	109,350
Average map resolution in Å (FSC 0.143)	2.6
PDB	
Map to Model resolution in Å (FSC 0.5)	2.7
Ramachandran favoured (%)	97.2
Ramachandran outliers (%)	0.0
Rama Z score	0.3
Rotamer outliers (%)	0.8
C-beta deviations	0
Rms on bond lengths	0.0072
Rms on bond angles	1.05
Clashscore	7.1
Molprobrity score	1.53

## 5 Chapitre II : Structure de la protéine Sol-Llp

---



## 5.1 Introduction

Dans cette partie je présenterai la structure de la Llp résolue par RMN. La Llp est une lipoprotéine qui est ancrée à la membrane par 2 acylations sur la cystéine terminale de la protéine et doit être manipulée avec des détergents. La taille des micelles de détergent est de l'ordre de 40 à 60 kDa. La technique de RMN pour résoudre des structures de protéine est particulièrement adaptée pour des petites protéines en solution, de taille typiquement inférieure à 40 kDa. La protéine Llp fait 7,9 kDa mais avec du détergent, le DDM par exemple, elle apparaît comme une particule de 60 kDa en solution. La délétion de la cystéine N-terminale permet d'obtenir une protéine soluble (Sol-Llp) de 7,06 kDa. Je détaillerai la mise au point de la purification Sol-Llp. Afin de s'assurer que la structure obtenue par RMN soit valide, nous avons vérifié que la troncation de la cystéine terminale n'affecte pas le repliement global de la protéine par dichroïsme circulaire et RMN. La forme soluble de Llp est donc principalement produite pour la détermination de la structure de Llp par RMN. Néanmoins, nous avons caractérisé l'interaction entre Sol-Llp et FhuA qui sera présentée dans la troisième partie de la thèse.

## 5.2 Mise au point de la production et purification de Sol-Llp

### 5.2.1 Construction du plasmide pour la surexpression de Llp sous forme soluble

Comme décrite dans le paragraphe I.5, la surexpression de Llp, sous forme acétylée ou non, a été décrite en 1998<sup>51</sup> et 2003<sup>71</sup>, respectivement. Les laboratoires ayant mené ces études ont été contactés. L'équipe à laquelle appartenait Kaspar Locher à l'institut Pasteur a été dissoute, plus aucun plasmide n'est disponible. La production décrite dans l'article de Pedrucci et al. a montré un rendement important pour la protéine Llp sous forme soluble par expression en corps d'inclusion puis repliement *in vitro*. Cependant la protéine possède un pont disulfure interne et nous avons conçu 2 constructions pour optimiser la production de Llp sous forme soluble dans le périplasma.

Schéma de la construction	Nom de la protéine	Objectif de la construction
	Sol-Llp	Production de la protéine sans le site d'acylation, dans le cytoplasme bactérien (en corps d'inclusion ou dans la souche Shuffle)
	MBP-Llp MBP-His-Llp	Production de la protéine non acétylée, en fusion avec la MBP. Adressage au périplasma pour la formation des ponts disulfures. La MBP augmentera la solubilité de la protéine et servira d'étiquette de purification.

**Tableau 5-1 : Constructions et stratégie pour l'expression de Llp sous forme soluble.**

Le tableau présente les différentes constructions essayées dans le projet et leur intérêt. Le schéma de construction est représenté graphiquement. Abréviations utilisées : SS : séquence signal, C : cystéine, His : étiquette polyhistidine, TEV : site de coupure de la protéase TEV, MalE : gène de la MBP, llp : gène de Llp

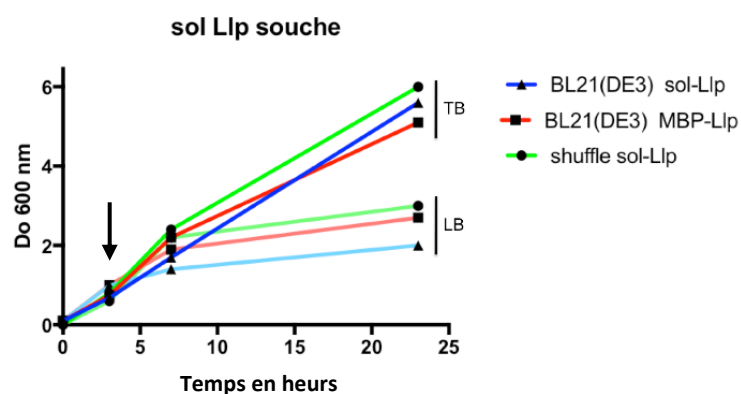
Le clonage réalisé est un clonage par double PCR. Le principe est de faire une première PCR (PCR1) qui permet de récupérer les séquences d'intérêt (*lfp* dans l'ADN génomique du phage T5) en y ajoutant des séquences d'hybridation pour le plasmide de destination (pET20b ou pMAL-p2E). La seconde PCR permet d'amplifier le plasmide choisi en y insérant le gène d'intérêt en se servant des produits de PCR 1 comme amorce. Le plasmide utilisé comme modèle pour la deuxième PCR est issu d'une purification après amplification bactérienne, il est donc méthylé et peut être digéré par la DpnI après la deuxième PCR (PCR2). Seuls les plasmides contenant l'insert sont alors présents après la digestion enzymatique à la DpnI et peuvent transformer les bactéries.

La présence de l'insert est vérifiée par digestion enzymatique (BamHI) avant l'envoi en séquençage (Annexe 3). Le séquençage pour les plasmides Sol-Llp et MalE-Llp a confirmé l'absence de mutation dans la séquence codant pour la protéine.

## 5.2.2 Production de la protéine Sol-Llp

### 5.2.2.1 Culture bactérienne, expression et localisation

La construction Sol-Llp ne possède pas de séquence signal. La croissance bactérienne a été contrôlée en présence d'IPTG pour vérifier l'éventuelle toxicité de la surexpression de la protéine. La souche BL21(DE3) a été testée pour les constructions Sol-Llp (en corps d'inclusion) et MBP-Llp (dans le périplasme). La souche Shuffle a également été testée pour la construction Sol-Llp. En effet elle a un cytoplasme réducteur ce qui permet de produire des protéines possédant des ponts disulfures sans export vers le périplasme. Dans toutes les conditions, l'ajout d'IPTG n'a pas freiné la croissance bactérienne (Figure 5-1). On note que le milieu TB permet de doubler les densités optiques ( $Do_{600nm}$ ) en fin de culture ( $Do_{culture_{600TB}} \sim 6$  vs  $Do_{600LB} \sim 3$ ).



**Figure 5-1 : courbe de croissance bactérienne de souche BL21 et Shuffle transformées avec les plasmides Sol-Llp ou MBP-Llp**  
 Courbes de la densité optique à 600 nm sans unité, en fonction du temps en heure. Comparaison entre le milieu LB et TB pour la croissance bactérienne. La flèche indique à quel moment les bactéries ont été induites.



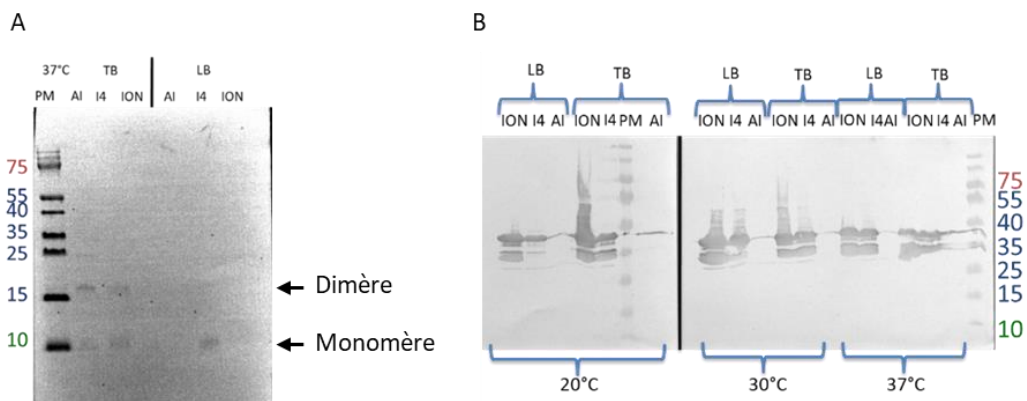
Le niveau d'expression a été testé dans 14 conditions, en milieu TB ou LB, à 37 °C, 30 °C ou 20 °C, en fonction de la souche d'expression et de construction (Tableau 5-2).

plasmide	sol-Llp : pET20b Llp soluble				MalE-Llp : pMalE-p2E-Llp	
souche	BL21DE3		Shuffle		BL21DE3	
condition						
milieu de culture	TB	LB	TB	LB	TB	LB
température	37				37	37
			30	30	30	30
			20	20	20	20

**Tableau 5-2 : Condition des tests d'expression pour la Llp sous forme soluble.**

*BL21(DE3) avec le plasmide Sol-Llp (objectif : corps d'inclusion) / Shuffle avec plasmide Sol-Llp (objectif : production lente & pont disulfure dans le cytoplasme) / BL21(DE3) avec le plasmide MalE-Llp (objectif : export au périplasme & augmentation de ...)*

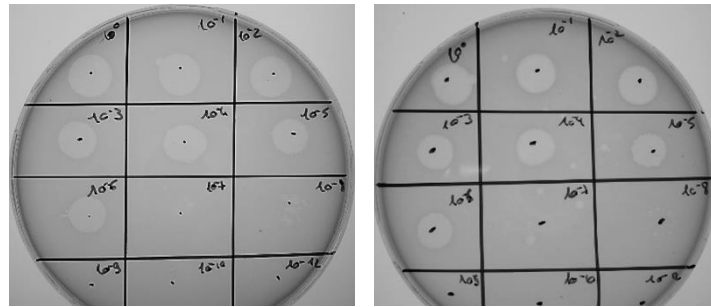
Le niveau d'expression de Llp a été contrôlé par immuno-révélation sur les bactéries totales déposées sur gel, avec des anticorps dirigés contre l'étiquette poly-histidine pour la construction Sol-Llp ou des anticorps dirigés contre la MBP pour la construction MalE-Llp. Il n'y a eu aucune expression de Sol-Llp dans la souche Shuffle (non montré). L'expression de Sol-Llp est très faible dans la souche BL21(DE3) et montre une protéine à la taille attendu, 7 kDa, et un dimère à 14 kDa. En présence de réducteur on s'attend à obtenir une seule bande à la taille attendue. Dans le cas de Llp la présence de réducteur à des concentrations classiques ne permet pas de réduire tous les ponts disulfures et permet la formation de dimère. Pour obtenir un comportement classique en présence de réducteur il faut être à de fortes concentrations de réducteur (environ 10 % de  $\beta$ -mercaptoéthanol). Sur les gels avec réducteur Llp présente donc un dimère artefactuel qui résulte d'une ré-oxydation des ponts disulfures (Figure 5-2 (A)). L'expression en BL21 contenant le plasmide de MalE-Llp a été satisfaisante. L'immuno-révélation révèle deux bandes, une pour la taille attendue pour la fusion MBP-Llp et l'autre pour la taille attendue pour la MBP seule, suggérant qu'il y a dégradation ou arrêt de traduction de la protéine de fusion. La condition permettant d'obtenir le plus de protéines de fusion MBP-Llp est une expression dans le milieu LB à 30 °C avec une induction sur la nuit (Figure 5-2 (B)).



**Figure 5-2 : Tests d'expression de la protéinée Sol-Llp (A) ou MalE-Llp (B) dans les souches BL21(DE3).**

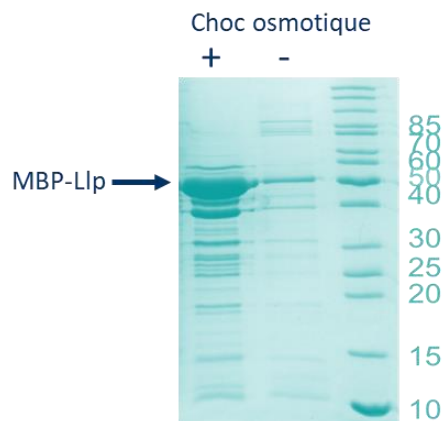
*PM : poids moléculaires, AI : avant induction, I4 : 4 heures et ION : une nuit d'induction avec 1mM d'IPTG. A. Immunorévélation avec un anticorps anti poly histidine sur des bactéries totales exprimant la Llp soluble. Toutes les expressions ont été faites à 37°C. À gauche, tests d'expression sur BL21(DE3) en TB. À droite, tests d'expression sur BL21(DE3) en LB. B. Immunorévélation avec un anticorps anti MBP sur des bactéries totales exprimant la protéine de fusion MBP-Llp. Expression à 20°C / 30°C et 37°C en TB ou LB. La taille attendue pour la protéine de Fusion MBP-Llp et de 52,4 kDa et la taille attendue de la MBP seule et de 44,5 kDa.*

La protéine MBP-Llp étant correctement exprimée, nous avons vérifié si elle était capable de jouer son rôle physiologique, c'est-à-dire protéger la cellule d'une infection au bactériophage T5. La sensibilité au phage T5 de la souche BL21(DE3) portant le plasmide MalE-Llp a été testée. Comme publié précédemment<sup>71</sup>, les résultats montrent que la souche BL21(DE3) n'est pas protégée contre le bactériophage T5 par la protéine MBP-Llp (Figure 5-3). Nous avons fait l'hypothèse que, la protéine MBP étant une grosse protéine, pouvait empêcher la liaison de Llp à FhuA par encombrement stérique.



**Figure 5-3 : test phénotypique par sensibilité au phage T5 de la souche BL21 contenant ou non le plasmide d'expression codant pour Sol-Llp.**  
De gauche à droite : souche BL21(DE3) / souche BL21(DE3) avec plasmide MBP-Llp.

Une température de culture bactérienne de 30°C ou 20°C pouvant affecter les systèmes de transport au périplasma, la localisation de la protéine de fusion MBP-Llp a été vérifiée pour s'assurer que le pont disulfure est bien formé. Un choc osmotique a été appliqué sur les cellules surexprimant MBP-Llp. On retrouve un peu de MBP-Llp dans le surnageant pendant l'incubation avec le saccharose, probablement due à des bactéries déjà lysées, ou plus fragiles, qui libèrent le contenu de leur périplasma dans la solution contenant le saccharose. En revanche après dilution dans l'eau, on voit que la MBP-Llp est largement libérée (Figure 5-4). La protéine a donc bien été adressée au périplasma. Les ponts disulfures sont donc présents et l'hypothèse de l'encombrement stérique validé.

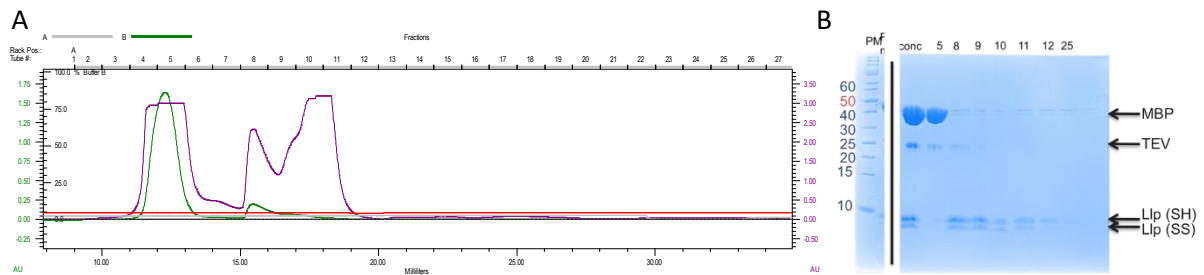


**Figure 5-4 : Vérification de la localisation subcellulaire de la protéine MBP-Llp.**  
PM : poids moléculaires. - : surnageant après incubation des bactéries dans 0,5 M saccharose, + : surnageant après incubation des bactéries dans l'eau. Une majorité de la protéine se trouve au niveau de la taille attendue de MBP-Llp (52,4kDa). On note également une bande plus faible un peu en dessous de 40 kDa qui pourrait correspondre à de la MBP seule (soit MBP endogène, soit MBP-Llp dégradée).



### 5.2.2.2 Mise au point de la purification

La protéine de fusion MBP-Llp est purifiée par une colonne d'affinité à la MBP permettant d'obtenir uniquement la protéine de fusion et de la MBP seule. La MBP est séparée de Llp par une coupure à la TEV, permettant d'obtenir un mélange de protéine MBP, Sol-Llp et TEV. La TEV et la MBP sont séparées de Sol-Llp par un tamis moléculaire. Le gel d'acrylamide de la fin de purification montre que la Llp est bien séparée de la protéine TEV, en revanche la MBP est présente dans toutes les fractions contenant Llp (Figure 5-5).



**Figure 5-5 : Résultat de la première purification de MBP-Llp**

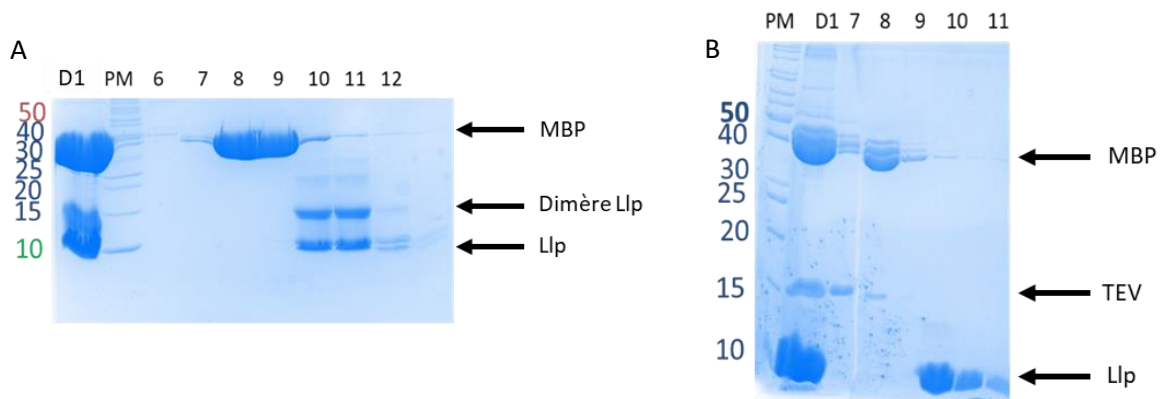
A. profil d'élution de la colonne de tamis moléculaire. En vert, absorbance à 280 nm, en violet absorbance à 320 nm, en rouge conductivité. B. gel d'acrylamide 15% coloré au bleu de Coomassie. De gauche à droite. PM : poids moléculaire ; conc : échantillon concentré, injecté sur la colonne de tamis moléculaire ; 5, 8, 9, 10, 11, 12, 25 : fractions de purification.

La purification de la forme soluble de Llp, Sol-Llp, a été optimisée pour augmenter la pureté et le rendement afin d'obtenir suffisamment d'échantillons pour les études structurales. La construction MalE-his-Llp a été réalisée pour faciliter la purification de Llp sous forme soluble en permettant d'enlever la MBP et la TEV par une colonne au nickel négative. La construction a été réalisée à l'aide d'une seule PCR pour insérer une étiquette polyhistidine en C-terminal de la protéine MBP sur la construction MalE-Llp.

Malgré l'insertion de l'étiquette histidine sur la MBP, la purification de Sol-Llp a été optimisée avec 4 étapes :

- MBP-Trap pour purifier la protéine de fusion MBP-Llp
- Coupure entre la MBP et Sol-Llp par la protéase TEV
- NiNTA pour retenir la TEV et la MBP
- Tamis moléculaire pour séparer la MBP restante et Sol-Llp.

Le rendement de purification a également pu être augmenté. D'abord avec un rendement à 1,8 mg de protéine pour 1 L de culture bactérienne (grâce à une meilleure maîtrise du protocole). Cependant, il reste un contaminant en fin de purification, la MBP. On suppose que la MBP est retardée sur la colonne de tamis moléculaire, car celle-ci est composée de sucre. Une nouvelle série d'optimisation a permis d'obtenir un rendement à 17 mg de protéine pour 1L de culture en réalisant la culture bactérienne en milieu riche (TB), mais aussi en purifiant à l'aide de 2 colonnes MBP en série. Pour améliorer la pureté de Sol-Llp, les échantillons issus de la colonne de nickel négative ont été plus concentrés avant la dernière étape de purification par un tamis moléculaire, et du maltose a été ajouté dans le tampon de migration de cette dernière étape pour que la MBP ne soit pas retardée sur la colonne (Figure 5-6).



**Figure 5-6 : comparaison de la purification de Sol-Llp avant et après optimisation.**

*Contrôle des fractions du tamis moléculaire sur un gel d'acrylamide 18% coloré au bleu de Coomassie. A : purification avant optimisation et échantillon préparé en présence de réducteur. B : purification après optimisation et échantillon préparé sans réducteur. A et B : PM : poids moléculaire. 6 à 12 : numéros de fraction à la sortie de la colonne D1 : échantillons chargés sur le tamis moléculaire*

La production de la protéine Sol-Llp en milieu minimum en présence d'isotopes  $^{15}\text{N}$  et  $^{13}\text{C}$  a également été réalisée pour les études structurales par RMN. Le protocole appliqué est le même que pour la protéine non marquée. En tout, 4 purifications ont été nécessaires pour l'ensemble des études par RMN avec un rendement moyen de  $3,4 \pm 0,6$  mg de protéine pour 1 L de culture.

Llp a été décrite comme soluble à une concentration inférieure à 1 mg/ml<sup>51</sup>. La protéine purifiée a été concentrée par ultrafiltration (Amicon, MWCO 3 kDa). L'absence d'agrégation de la protéine a été contrôlée par un spectre UV-visible. 125  $\mu\text{l}$  de protéine Llp soluble a été obtenue à une concentration de 1,14 mg/ml soit 160  $\mu\text{M}$ , avec un rendement de concentration de 60% pour la première purification. Pour la RMN des concentrations de protéines proche de 1 mM sont recommandées. La concentration a donc été poussée et toujours contrôlée par des spectres UV-Visible. Finalement nous avons pu atteindre une concentration maximum de 3 mM sans signe d'agrégation de la protéine avec un rendement moyen de concentration de 80%.

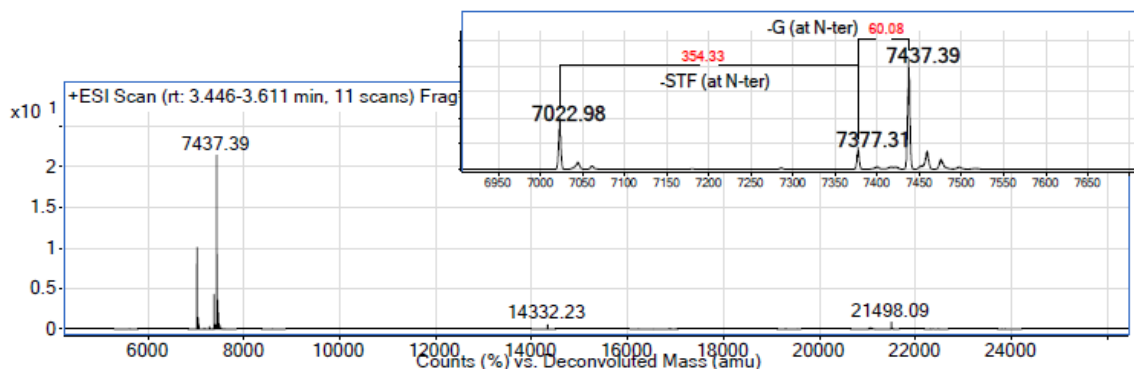
### 5.3 Caractérisation biochimique de Sol-Llp et validation pour la RMN

#### 5.3.1 Vérification de l'intégrité des protéines par spectrométrie de masse

Sol-Llp a été contrôlée en spectrométrie de masse après purification. La masse mesurée est de 7043,99 Da indiquant que le pont disulfure est formé (masse attendue 7046,01 Da -2 Da correspondant à la perte des deux hydrogènes résultant de la formation du pont disulfure).

La stabilité de Llp a été vérifiée pour les études de RMN. Après 2 semaines à 4 °C la masse de Llp est identique à la masse de la protéine purifiée avec la présence du pont disulfure soit 7118.61 Da pour une protéine marquée à l'isotope  $^{15}\text{N}$ . En revanche après 2 semaines à 20°C, deux produits de dégradation sont visibles, dont la différence de masse peut correspondre à la perte d'une glycine puis d'une serine en N-terminal. Le pic à la masse attendue étant majoritaire, la protéine a été jugée suffisamment stable pour les mesures par RMN.

Au cours des enregistrements RMN une dégradation a été clairement visible sur les spectres les 2 premiers jours d'enregistrement puis les spectres étaient stables. Il a été décidé de contrôler la protéine à la suite de la campagne d'enregistrement des spectres 3D de RMN. Les spectres de masse ont confirmé une dégradation des 4 premiers acides aminés GSTF. Bien que la forme majoritaire reste la protéine complète, la proportion de protéine dégradée est non négligeable (environ la moitié des protéines présentes) et sera donc pris en compte dans l'analyse des spectres de RMN (Figure 5-7).



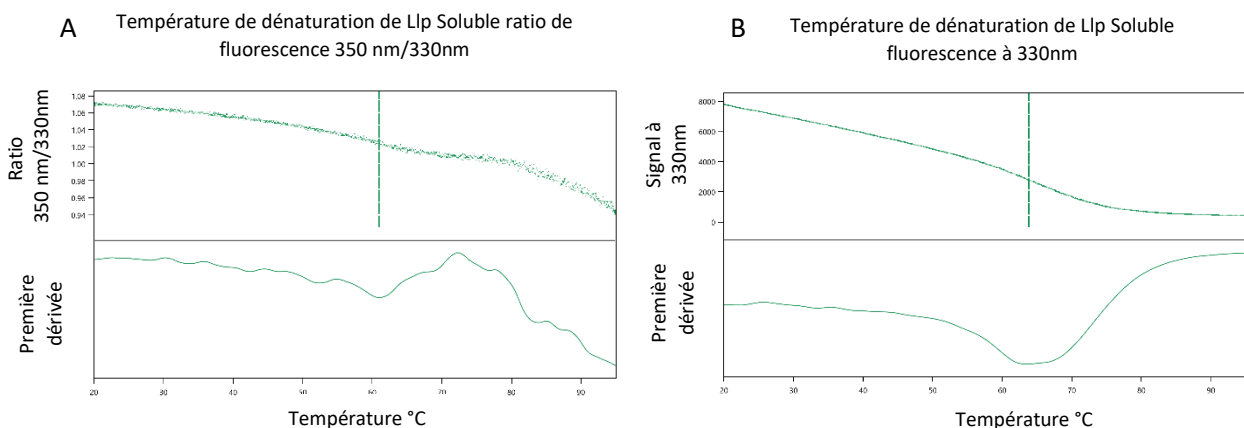
**Figure 5-7 : spectrométrie de masse sur la protéine Sol-Llp après enregistrement des spectres RMN.**

Abcisse : m/z, ordonné : intensité en A.U, chaque pic présente la masse en Da, et les différences de masse entre les pics sont indiquées en rouge.

### 5.3.2 Stabilité thermique

Étudier la stabilité thermique des protéines permet de comparer l'effet de la composition de différents tampons sur la protéine. En biologie structurale cela permet d'ajuster les conditions pour les mesures biophysique en optimisant la stabilité des protéines. L'étude de la stabilité thermique est faite par Nano-DSF. C'est une méthode simple, sans marquage et rapide. Les mesures de la stabilité thermique sont mesurées sur les protéines purifiées.

La technique nano-DSF permet de mesurer la température de dénaturation en suivant le changement de fluorescence des tryptophanes enfouis dans la structure par un changement de longueur d'onde de fluorescence de ces tryptophanes quand ils sont exposés au solvant. L'inflexion dans le ratio de fluorescence 350 nm/330 nm permet donc de définir la  $T_m$ . Les mesures effectuées au laboratoire ne montraient pas de transition pour le ratio de longueur d'onde utilisé (Figure 5-8 (A)). En revanche pour une seule longueur d'onde l'inflexion permettait de trouver la température de dénaturation que ce soit à 330 nm (Figure 5-8 (B)) ou à 350 nm (non montré). Ce phénomène s'explique par le fait que la protéine ne possède qu'un seul tryptophane, déjà exposé au solvant et trois tyrosines. Or, l'exposition des tyrosines au solvant change drastiquement l'intensité de fluorescence, mais pas la longueur d'onde. Les résultats sont cohérents avec les observations du constructeur et peuvent être utilisés. L'ensemble des températures de dénaturation seront étudié à 330 nm pour pouvoir être comparées.

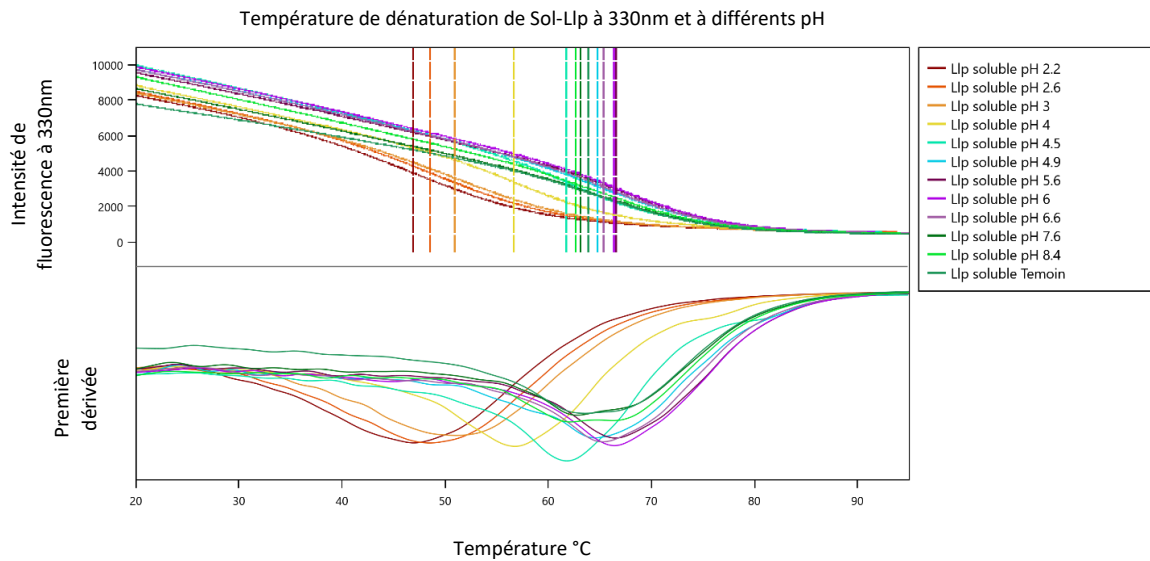


**Figure 5-8 : courbe de suivi de dénaturation de la protéine Sol-Llp par le nano-DSF.**

Abscisse : température en °C, ordonnés panneaux du haut : intensité de fluorescence sans unité en A et nombre d'évènements par minute en B. Panneaux du bas, en ordonné première dérivée, sans unité. A. Température de dénaturation de Llp Soluble pour ratio de fluorescence 350 nm/330 nm. B. Température de dénaturation de Llp Soluble suivie de fluorescence à 330 nm.

La stabilité de Sol-Llp pour la RMN a été contrôlée dans le temps (voir paragraphe précédent), mais également à différent pH, car un pH acide, permettant les échanges de proton dans la protéine, permet d'obtenir un meilleur signal en RMN. Sol-Llp est initialement purifiée dans un tampon à pH 8,5, est très stable puisque le  $T_m$  est de  $63,9 \pm 0,8$  °C (Figure 5-9). La protéine est stable sur une grande gamme de pH puisque le  $T_m$  commence à baisser pour un pH de 4,5. On peut même noter qu'un pH

légèrement acide, entre 5,6 et 6,6 augmente légèrement la stabilité de Sol-Llp d'environ 2,4°C en moyenne avec un maximum de 2,7°C supplémentaires à pH 6,6 (Figure 5-9). Le pH du tampon utilisé pour les mesures de RMN a donc été fixé à 6,5.



**Figure 5-9 : courbe de suivi de dénaturation de la protéine Sol-Llp à différent pH.**

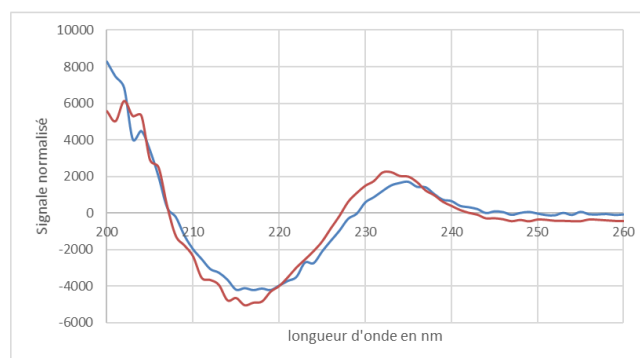
*Abscisse : température en °C, ordonnés bloc du haut : intensité de fluorescence en nombre d'évènements par minute ; Bloc du bas : première dérivée, sans unité. Témoin : Sol-Llp dans le tampon de purification a pH 8,5.*

L'impact du pont disulfure sur la stabilité de Ac-Llp et Sol-Llp a également été mesuré par le nano-DSF. Les Tm des protéines sont mesurés avec 5% ou sans bêta-mercaptoéthanol. En présence de  $\beta$ -mercaptoéthanol, le Tm de Sol-Llp diminue de 57.6°C à 43.5°C. Il y a une différence de 14,1°C, ce qui veut dire que la présence du pont disulfure est très importante pour la stabilité de la protéine. Cette vérification nous a permis de valider le protocole de purification au cours duquel la coupure a la TEV est réalisé à l'aide d'un couple oxido-réducteur pour préserver le pont disulfure plutôt que l'utilisation du DTT comme source réducteur pour la TEV.

### 5.3.3 Comparaison de structure de Ac-Llp et Sol-Llp par dichroïsme circulaire et RMN

Le CD (dichroïsme circulaire) permet de définir la proportion des différentes structures secondaire (hélice alpha, feuillet bêta et parties désordonnées) à partir de spectres mesurés avec de la lumière polarisée dans les UV lointains. Afin d'étudier la protéine Llp par RMN, la cystéine terminale portant les acylations a été enlevée par un clonage. Pour vérifier si l'absence de la cystéine terminal a un impact sur la structure globale de la protéine, les spectres CD de la forme acylée et soluble de la protéine sont comparés. Les spectres CD sont enregistrés sur les protéines purifiées (section 6.2.2 pour la purification de Ac-Llp). Les spectres de CD sont très semblables entre Ac-Llp et Sol-Llp (Figure 5-10). On note une petite différence entre les deux spectres aux alentours de 230 nm. En revanche, le marquage  $^{15}\text{N}$  comme attendu, n'a aucune incidence sur le spectre (Figure 5-10). Les prédictions de la proportion de structure confirment l'aspect visuel des courbes, puisque pour Ac-Llp et Sol-Llp la prédiction de feuillet bêta est d'environ 40 %. Il y a une légère différence sur la prédiction des hélices alpha avec une prédiction à 0,17 % pour la protéine acylée et 1,2 % pour la protéine soluble.

#### Spectre de dichroïsme circulaire



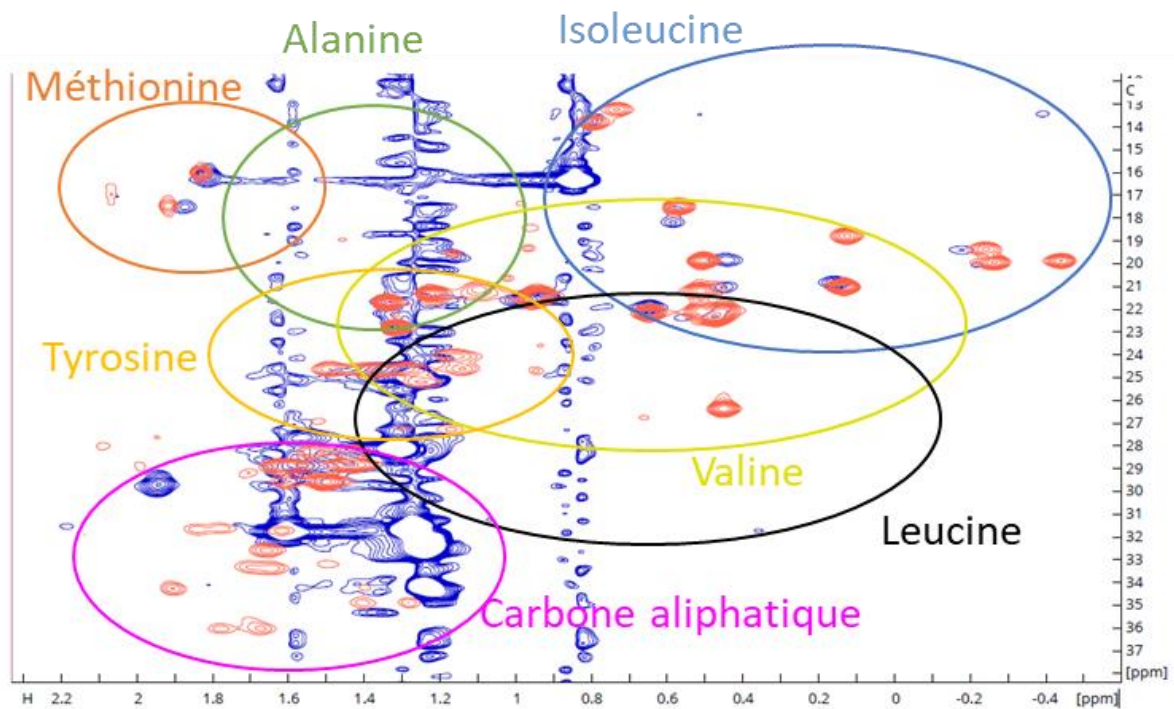
**Figure 5-10 : Comparaison des spectres de dichroïsme circulaire de Ac-Llp ou Sol-Llp.**

Les courbes présentées sont une moyenne de 10 acquisitions, auxquelles sont retranchées la moyenne de 10 acquisitions sur le tampon et normalisées par la quantité de protéine. Courbe bleu : Ac-Llp, Courbe rouge : Sol-Llp.

Afin de confirmer que la différence dans le spectre du CD est mineure, et pour valider l'utilisation de Sol-Llp pour la détermination de structure, un spectre HSQC des deux protéines Ac-Llp et Sol-Llp a été mesuré. La mesure des spectres pour la protéine Ac-Llp a été réalisée en OG (Octyl-beta-Glucoside Detergent), car la micelle de ce détergent est plus petite que celle du DDM (Dodécyl- $\beta$ -D-maltoside). Malheureusement le signal en RMN n'est pas suffisant pour obtenir un spectre HSQC pour la protéine Ac-Llp, car la relaxation est trop lente : la particule protéine – détergent est trop grosse (~50 kDa).

Pour résoudre ce problème, un spectre des méthyles en abondance naturelle de  $^{13}\text{C}$  a été enregistré sur Sol-Llp et Ac-Llp en OG. Le spectre de Ac-Llp est moins bien défini que celui de Sol-Llp dû à la relaxation plus lente de Ac-Llp. La protéine Llp contient 26 méthyles, parmi les acides aminés contenant des méthyles certains possèdent des attributions doubles pour l'HSQC (voir section 5.4.2), rajoutant ainsi 6 méthyles possiblement observables. En tout dans le spectre de Sol-Llp on compte 27 pics

dans la zone des méthyles. Dans les spectres TROSY méthyl chaque type de méthyl a une zone de déplacement chimique spécifique (Figure 5-11). Les méthyls des méthionines sont facilement observables car ils sont bien séparés des autres déplacements chimiques. Dans les spectres de Sol-Llp et Ac-Llp on retrouve la superposition de 2 pics intenses et un pic plus faible pouvant correspondre au 3 méthionines présentes dans la protéine dont 2 sont attribuées (voir section 5.4.2). Par ailleurs, une grande partie des méthyles des deux protéines sont superposés ou proches. Certains pics dans Sol-Llp ne sont pas visibles dans le spectre de Ac-Llp probablement dû à un signal faible. Inversement un pic dans la zone des méthyles se trouvent uniquement dans le spectre de Ac-Llp peut être dû à l'influence du détergent sur certains acides aminés (Figure 5-11). Dans la zone des carbones aliphatique, des pics intenses sont présents pour Ac-Llp, correspondant aux chaines des acides gras et des détergents. Dans le spectre Sol-Llp les pics présents dans cette zone peuvent être dus au tampon ou aux chaines latérales des acides aminés (Figure 5-11). Les spectres sont suffisamment proches pour conclure que Sol-Llp et Ac-Llp ont une structure globale très proche et viennent confirmer les résultats de dichroïsme circulaire, validant l'utilisation de Sol-Llp pour la détermination de structure de la Llp par RMN.



**Figure 5-11 : TROSY-méthyl-HSQC sur les méthyles de Sol-Llp (rouge) et de Ac-Llp en OG (bleu).** Spectres mesurés sur la nuit en abondance naturelle de  $^{13}\text{C}$  sur le spectromètre 600 MHz sur une protéine concentrée à 0,91 mM (Sol-Llp) et 1,15 mM (Ac-Llp) dans un tampon Tris pH 6,5. Les zones de déplacement chimique des différents méthyles et carbones aliphatiques sont mis en évidence.

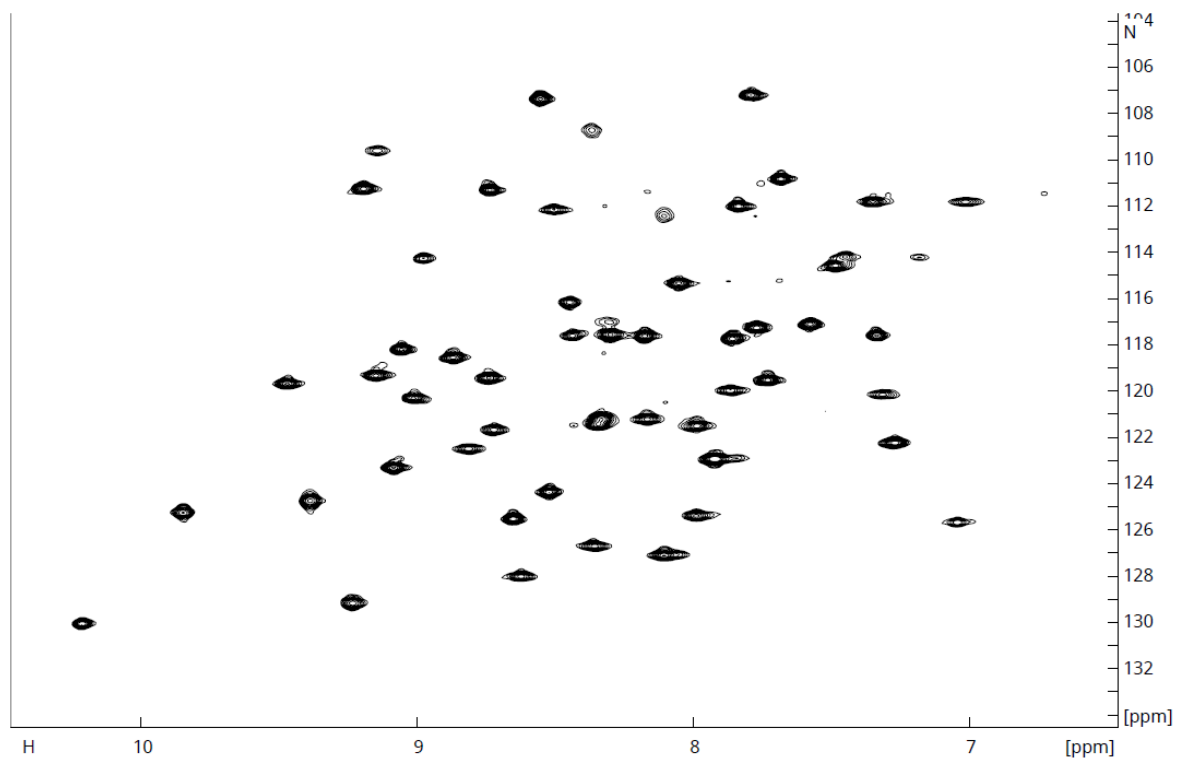
## 5.4 Résolution de la structure de Sol-Llp.

### 5.4.1 Détermination des conditions de mesure :

#### 5.4.1.1 Spectre 2D HSQC

La première étape de la détermination de la structure d'une protéine par RMN est la réalisation d'un spectre 2D HSQC. Dans cette séquence d'impulsion, le déplacement chimique d'un proton est corrélé avec le déplacement chimique de l'azote auquel il est lié. On attend donc un pic pour chaque couple NH de la protéine. On doit donc avoir autant de pics que d'acides aminés à l'exception des prolines pour les NH de la chaîne peptidique, mais dans des régions spécifiques, on doit aussi retrouver les NH des chaînes latérales des tryptophanes et des arginines. On peut également trouver des doublets de même déplacement chimique en azote mais différents en proton pour le NH<sub>2</sub> de chaîne latérale des asparagines et glutamines.

Pour Sol-Llp on attend donc 58 pics pour les acides aminés, un pic pour le tryptophane, un pic pour la chaîne latérale de l'arginine et six pics pour les deux asparagines et la glutamine, soit un total possible de 66 pics. Le premier enregistrement 2D sur une heure donne un spectre HSQC de Sol-Llp qui est bien dispersé dans les dimensions proton et azote, ce qui indique que la protéine est bien structurée (Figure 5-12). De plus on compte 53 pics sur les 66 pics attendus (Figure 5-12), ce qui est un bon résultat car on peut s'attendre à ce que certains pics 2D soient confondus, mais ils seront résolus sur les spectres 3D.



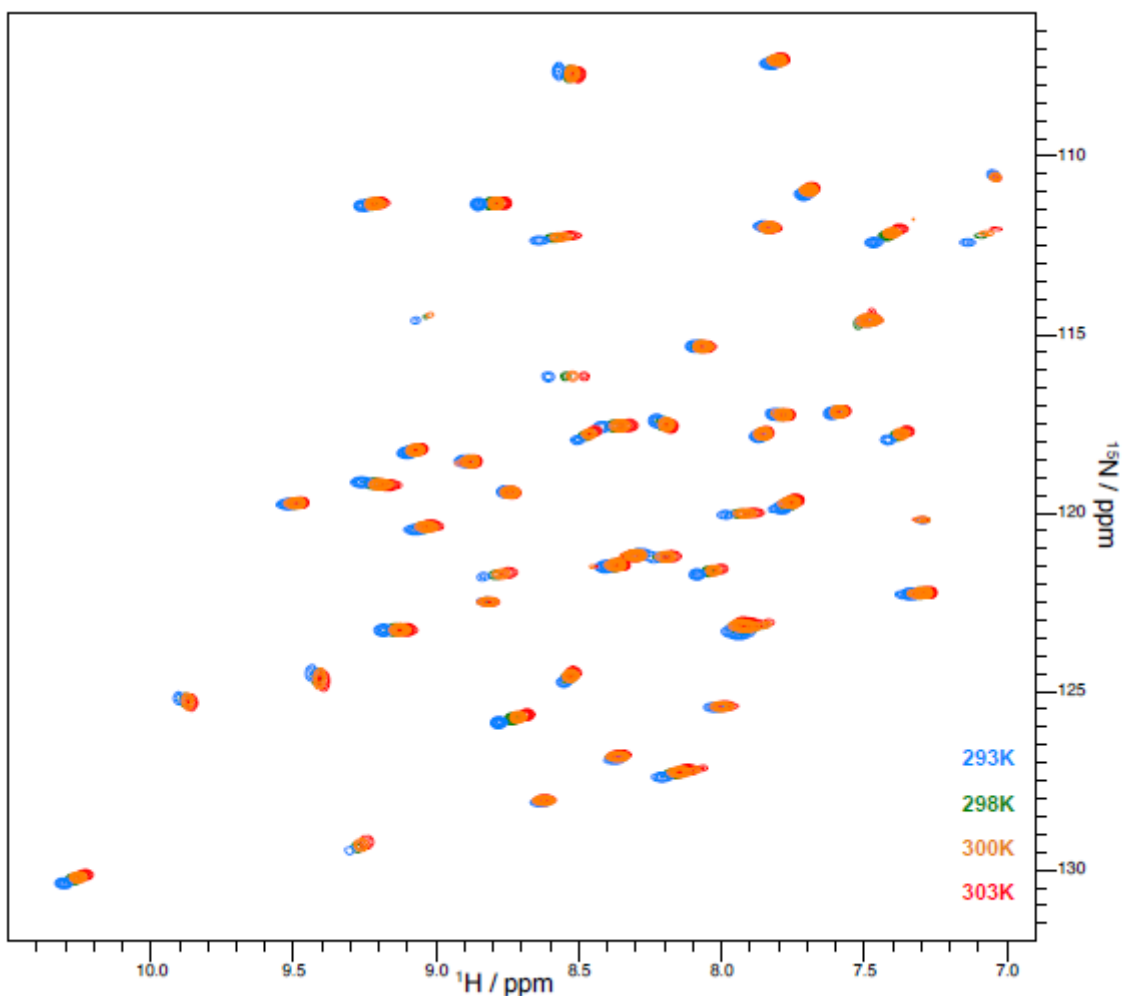
**Figure 5-12 : Spectre <sup>1</sup>H, <sup>15</sup>N-HSQC de Sol-Llp.**

Spectre mesuré en 1h sur le spectromètre 600 MHz sur une protéine concentrée à 900 μM. Travail en collaboration Emeline Mestdach.



#### 5.4.1.2 Effet de la température

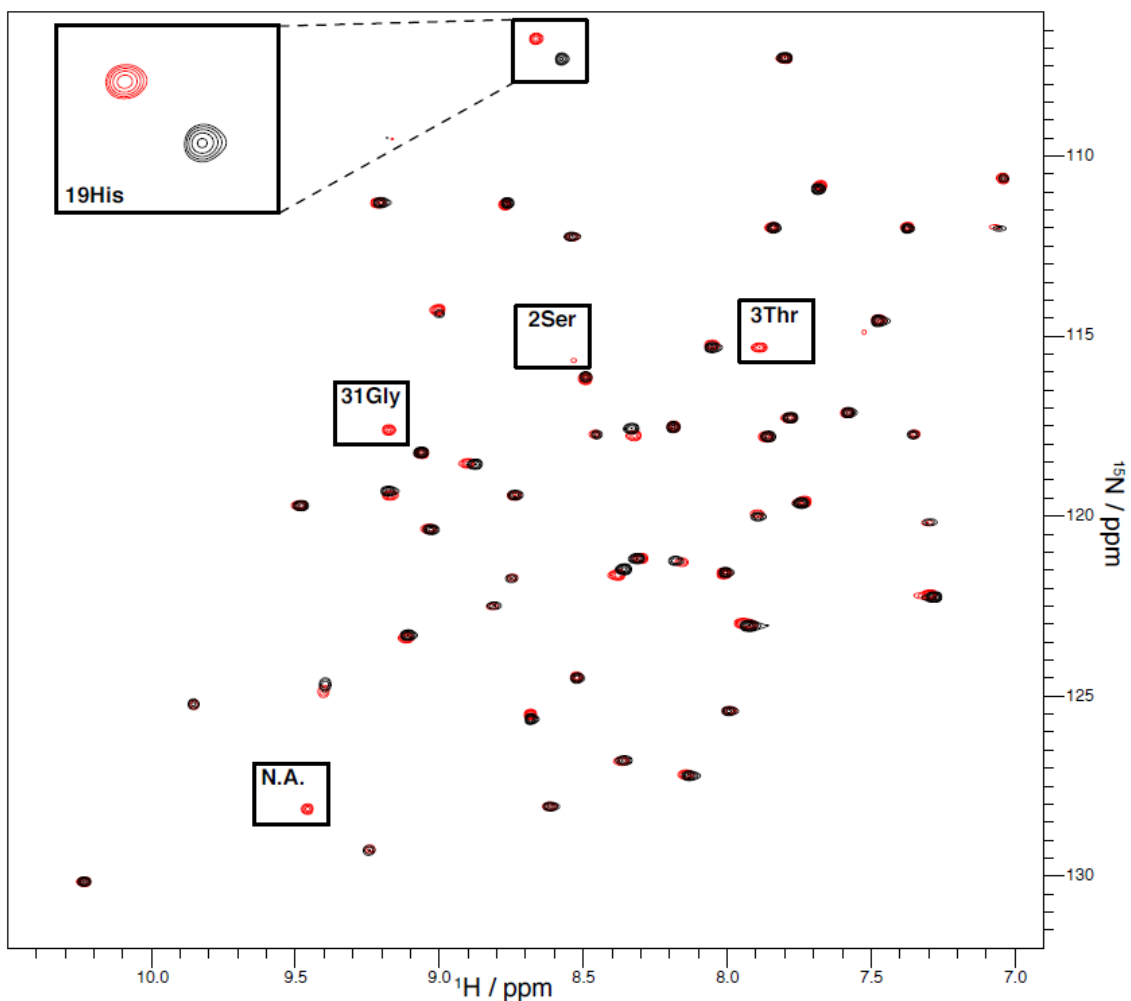
Plus la température de mesure des spectres de RMN est élevée, plus le « culbutage » (« tumbling ») des protéines est importante. Ceci permet d'avoir un temps de relaxation transversal plus long, donc des pics plus étroits et plus intenses, ce qui augmente la résolution et le signal sur bruit des spectres. L'effet de la température a donc été testé sur Sol-Llp. Quelle que soit la température de mesure, les spectres HSQC contiennent le même nombre de pics avec un bon signal sur bruit. Comme attendu certains pics sont affectés dans leurs déplacements chimiques par la température (Figure 5-13). La protéine est stable sur plusieurs jours à des températures de 20°C (293K), mais plus la température est élevée plus les résultats des mesures des spectres en trois dimensions sont bons. Le compromis retenu entre stabilité et fiabilité des mesures a été de 303K (30°C).



**Figure 5-13 : Investigation de l'effet de température sur les spectres  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC de Sol-Llp.**  
Mesure sur le spectromètre à 600 MHz à des températures allant de 293K à 303K sur une protéine a 0,9 mM.

#### 5.4.1.3 Stabilité de la protéine

Comme discuté dans la section 5.3.1 (Vérification de l'intégrité des protéines par spectrométrie de masse) la protéine Sol-Llp se dégrade au cours du temps et perd d'abord la glycine N-terminale puis les trois acides aminés suivants, Ser-Thr-Phe. Cette dégradation a un impact sur les spectres HSQC de Sol-Llp. A 8 jours d'intervalle à 303K (Figure 5-14), certaines variations de déplacement chimique sont perceptibles, ainsi que l'apparition et la disparition de certains pics dans différentes régions : les résidus autour de la partie N-terminale (Ser2 et Thr3) et dans la première moitié de la protéine (His19 et Gly31). Cependant la variation du pic His19 peut être due à un changement de pH dans le temps mais n'a pas été testé. L'analyse de l'ensemble des spectres HSQC de contrôle pris entre chaque expérience 3D (toutes les 24h environ) montre que la dégradation et la possible variation de pH survient rapidement (dans les 2 premiers jours) puis les spectres HSQC ne montrent plus de changement. La plus grande partie des informations structurales étant enregistrée après 2 jours d'acquisition (total de 14 j), l'attribution des NH a été réalisée à partir d'un spectre HSQC dans lesquels la protéine est déjà dégradée.



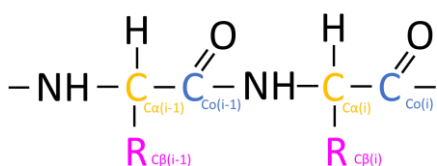
**Figure 5-14 : Spectres  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC de Sol-Llp au cours du temps acquis à 303K à 600 MHz.** Comparaison des spectres au début des expériences (en noir) et après 8 jours de mesure à 303K (en rouge). Les différences les plus visibles sont mises en évidence (type et numéro de résidu, N.A. étant non attribué).

### 5.4.2 Attribution du squelette carboné

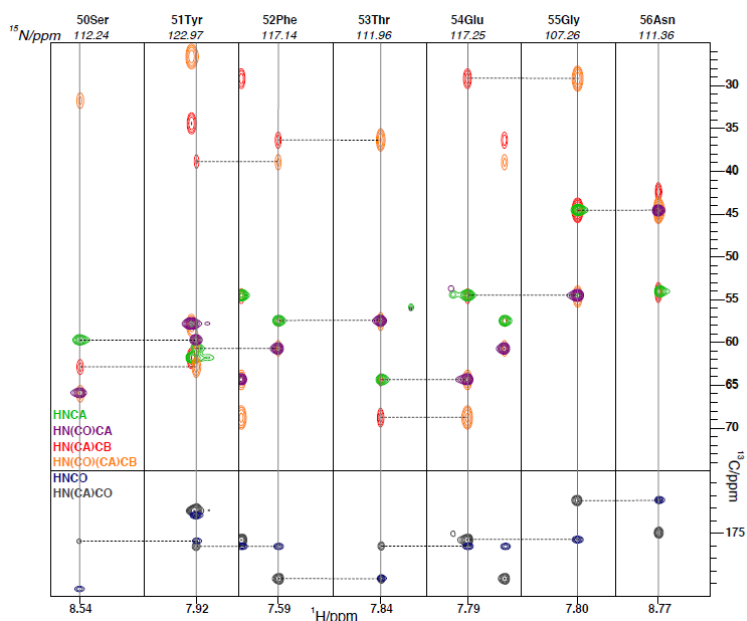
La stratégie d'attribution est basée sur la corrélation des déplacements chimiques des NH avec ceux des C $\alpha$  et C $\beta$  intra- et inter-résidus. Ces déplacements sont mesurés par des spectres en trois dimensions (NHcaCO, NHCACB, NHCA). Pour faciliter l'identification des carbones inter-résidus, des spectres contenant uniquement les carbones -1 sont enregistrés (NHCacoCB, NHcoCA et NHCO) (Figure 5-15 et Tableau 5-3). Les C $\alpha$  et C $\beta$  sont reliés avec leurs carbones -1 respectifs, permettant de lier les acides aminés séquentiellement (Figure 5-16). Les déplacements chimiques des C $\beta$  sont très sensibles à leur environnement et sont donc souvent un bon indicateur de la nature de l'acide aminé. Ceci permet de corréler la séquence théorique avec les données expérimentales. Si le nombre d'acides aminés séquentiellement attribués est trop petit, l'attribution n'est pas prise en compte.

Nom des expériences	Atomes observés
NHcaCO	Co(i) & Co(i-1)
NHCACB	Ca(i), C $\beta$ (i) & Ca(i-1), C $\beta$ (i-1)
NHCA	Ca(i) & Ca(i-1)
NHCacoCB	Ca(i-1), C $\beta$ (i-1)
NHcoCA	Ca(i-1)
NHCO	Co(i-1)

**Tableau 5-3 : atomes de carbone observés sur les spectres pour chaque expérience 3D d'attribution.**  
Noms des atomes voir figure 5-15.



**Figure 5-15: Noms des atomes de carbone sur la structure d'un acide aminé général.**

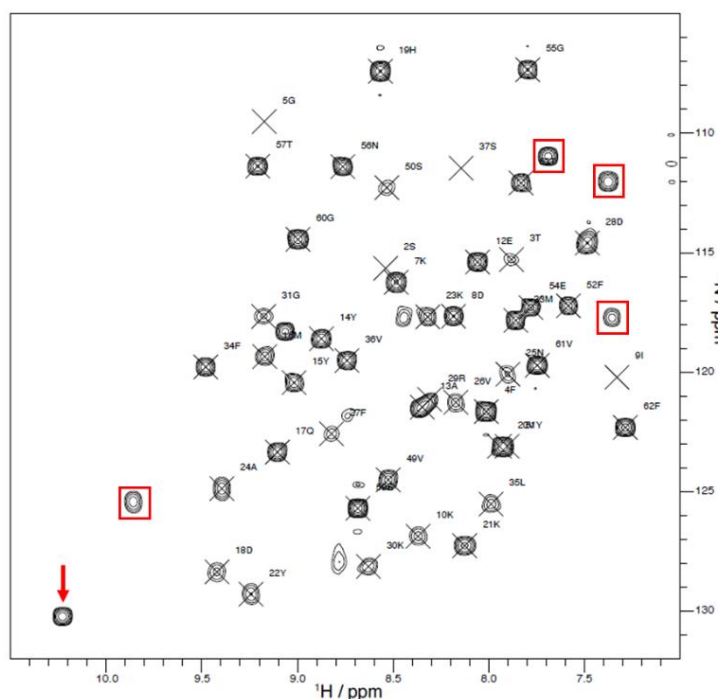


**Figure 5-16 : Stratégie d'attribution de la protéine Sol-Llp pour les résidus S50 à N56.**

Les pics des différents spectres (légende de couleur sur la figure) contiennent les résonances inter- et intra-résidus pour les carbones C $\alpha$ , C $\beta$  et CO pour chaque groupe NH identifié en HSQC (fréquence 15N écrite en italique au-dessus). Les lignes pointillées montrent la connexion entre les carbones *i* et *i-1*. Les lignes verticales séparent différentes fréquences d'hydrogène.

L'attribution de la chaîne principale a initialement été prévue pour être réalisée avec une attribution automatique et des spectre spécifiques APSI. Des problèmes de format de mesure n'ont pas permis l'attribution automatique dans un premier temps. L'attribution a donc été réalisée manuellement grâce aux expériences 3D, et a permis d'attribuer 47 acides aminés de la protéine soit 76% de la séquence de Sol-Llp. Seul un pic de l'HSQC est superposé à un autre couple NH. Les acides aminés 39 à 47 (séquence entre les deux prolines) n'ont pas pu être attribués. Certains pics pourraient correspondre à la séquence 38-42, mais ils peuvent également être attribués à une séquence ailleurs dans la protéine. Aucune résonance n'a pu être identifiée pour les résidus 1, et 43 à 45. Deux résonances pourraient être attribuées pour les acides aminés 46 et 47, mais elles n'ont pas été retenues, la chaîne attribuée étant trop courte pour être statistiquement valide. Des pics supplémentaires permettent de faire des attributions alternatives sur 35% de la séquence de la protéine pour les acides aminés 3 à 10, 17 et 18, 30, 22 à 28, et 50 à 53. Ces pics sont tous plus faibles en intensité que ceux ayant permis la première attribution, qui a été choisie pour la détermination de la structure. Les pics plus faibles indiquent que les attributions alternatives appartiennent à une sous population de conformation de la protéine. Cette sous population de conformation peut être soit due à la dégradation des premiers acides aminés, soit à une conformation alternative. La structure de cette seconde attribution n'a pas été explorée. Enfin, 4 pics intenses dans l'HSQC (sans compter le pic dû au tryptophane) n'ont pas pu être attribués (Figure 5-17). L'attribution automatique a finalement pu être réalisée et a confirmé les résultats obtenus. Comme pour l'attribution manuelle, l'attribution automatique a proposé plusieurs attributions possibles pour les mêmes acides aminés.

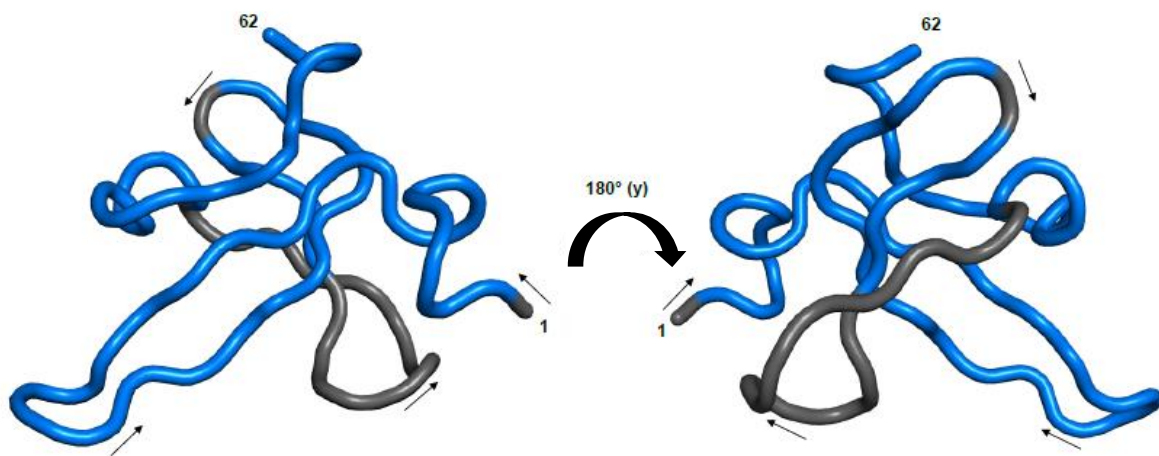
Sur le même principe que l'attribution du squelette carboné, les protons liés au carbone  $\alpha$ , HA et les protons liés au carbone  $\beta$ , HB ont été attribués à l'aide des spectre HBHA(CO)NH et HBHANH.



**Figure 5-17 : Spectre HSQC de Sol-Llp avec l'attribution.** Chaque pic est nommé avec sa position dans la séquence et la nature de l'acide aminé. 76 % de la séquence de la protéine a été attribuée. La flèche rouge indique la chaîne latérale du tryptophane et les pics entouré d'un carré rouge sont les pics intenses qui n'ont pas pu être attribués

### 5.4.3 La structure de Sol-Llp

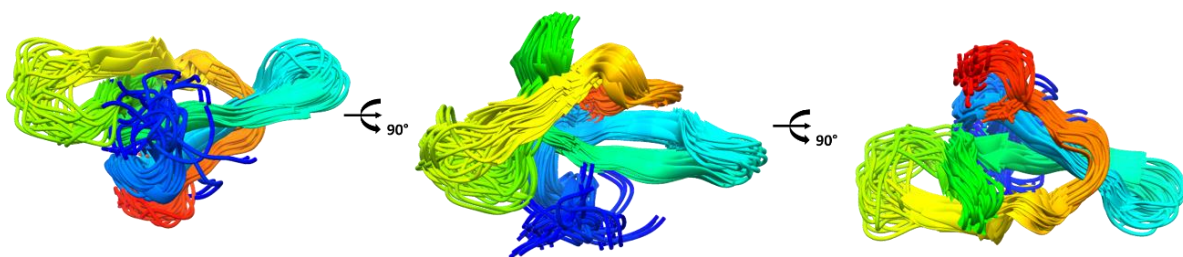
Le calcul de structure a été réalisé par Torsten Hermann à l'aide de l'attribution automatique de la chaîne principale et de l'attribution manuelle des HA et HB. Le programmes d'analyse RMN UNIO<sup>88</sup> en combinaison avec le programme de calcul de structure RMN CYANA<sup>92</sup> a généré automatiquement 511 restrictions de distance significatives et non redondantes à partir des spectres NOE (Nuclear Overhauser Enhancement). La structure finale de RMN a été calculée dans un solvant explicite à l'aide du programme CNS<sup>93</sup>. Les 20 conformères générés par le calcul de structure présentent tous une très bonne stéréochimie, avec 95,2 % des résidus dans les régions plus favorables du diagramme de Ramachandran, 3,1 % de résidus dans des régions généralement permissives, et 1,7% dans des régions non autorisées du diagramme. En outre, l'ensemble de la structure présente une excellente géométrie, sans violation systématique des contraintes de distance supérieure à 0,2 Å et sans violation systématique des angles dièdres supérieure à 5°. À l'exception des cinq premiers résidus N-terminaux et de la région en boucle englobant les résidus 39 à 44, les coordonnées atomiques de la structure RMN sont bien définies. La valeur de l'écart quadratique moyen (RMSD) par rapport aux positions des coordonnées moyennes des atomes du squelette pour les résidus 6-38 et 45-62 est de  $1,02 \pm 0,19$  Å (Figure 5-18). Les statistiques détaillées de la structure RMN et de la minimisation de l'énergie sont présentées en annexe (Annexe 2).



**Figure 5-18 : structure RMN moyenne de Sol-Llp.**

*Les acides aminés attribués sont colorés en bleu et les parties non attribuées sont en gris. Les flèches indiquent le sens croissant de la séquence d'acide aminé.*

La structure globale de Llp est bien définie et la présence de nombreux brins bêta explique la stabilité de cette protéine ( $T_m$   $63,9 \pm 0,8$  °C). La structure de Llp présente 2 feuillets bêta composés de 3 brins bêta chacun. Le premier feuillet est formé par les brins bêta C11-M16, V20- V26, et E54- T57. Les brins bêta 1 et 5 sont liés par un pont disulfure entre C11 et C58. Le second feuillet bêta comprend les brins Asp28-Lys30, M33-V36 et F45-P48. Il y a également 2 petites hélices, une présente dans la partie N-terminale (P6 -I9) et l'autre dans la partie C-terminale (V49-Y51). L'ensemble des 20 conformères se superposent bien sur la majorité de la structure. Le N terminal et une boucle non définie de I39 à A47 sont moins bien alignés, suggérant que ces parties sont flexibles (Figure 5-19). Ces résultats sont confirmés par une analyse de relaxation réalisée par Émeline Mestdach, CRMN, Lyon.



**Figure 5-19 : structure RMN des 20 conformères les plus probables.**  
Structure de Sol-Llp déterminée par RMN coloré en arc en ciel de N terminal (bleu) au C terminal (rouge).

## 5.5 Discussion

La production de l'échantillon en quantité et qualité suffisante est un enjeu majeur dans la biologie structurale. Le clonage et la mise au point de la purification de Sol-Llp ont été rapides, environ 2 mois et demi. Le rendement de la purification de Sol-Llp a été suffisant pour produire Sol-Llp à partir d'un litre de culture pour chaque expérience, y compris en marquage  $^{15}\text{N}$ - $^{13}\text{C}$ . Comme dans l'étude <sup>51</sup>, les résultats de dichroïsme circulaire montrent que Sol-Llp a un contenu en structures secondaires très proche de Ac-Llp. Ces résultats ont été complétés par une étude RMN qui confirme les similarités de structure entre Sol-Llp et Ac-Llp. Sol-Llp peut donc être utilisé pour résoudre la structure de la protéine seule.

Une recherche BLAST a été réalisée et aligne la séquence de Llp notamment avec les protéines « lytic conversion lipoprotein precursor » des phages S131 de *Salmonella*, VB\_EcoS\_FFH et Slur09 d'*Echerichia coli*. Étonnamment, on trouve aussi une protéine hypothétique de *Shigella Flexenari* et le lysozyme du phage EC100 infectant *E. coli*. La protéine hypothétique de *S. Flexenari* est issue d'un prophage (n° de génome / ref). L'ensemble de ces protéines sont donc des protéines d'exclusion de super infection ciblant les transporteurs de la membrane externe.

Comme présenté dans le paragraphe 1.1.5 décrivant le phénomène d'exclusion de surinfection, les protéines Cor sont des protéines d'exclusion de super infection qui ciblent FhuA et qui sont produites par les coliphages mEp167, mEp213, N15, HK022 et  $\phi$ 80. Ces protéines ne sont pas présentes



dans les résultats du BLAST. Des alignements de séquence ont été réalisés sur Clustal Omega (Figure 5-20). L'alignement des séquences des protéines Llp-like et des protéines Cor montre que la séquence signal est très conservée pour l'ensemble de ces protéines. Les 3 cystéines sont présentes dans l'ensemble de ces protéines ce qui indique que leur rôle sont essentiels pour la fonction de la protéine. En effet la première cystéine est acylée au cours de la maturation ce qui lui permet d'être ancrée dans la membrane externe et de cibler FhuA. Pour la protéine Llp, les deux autres cystéines forment un pont disulfure indispensable pour la protection de la cellule <sup>71</sup>. Cet alignement montre aussi que Llp de T5 et les protéines Cor ont une origine commune.

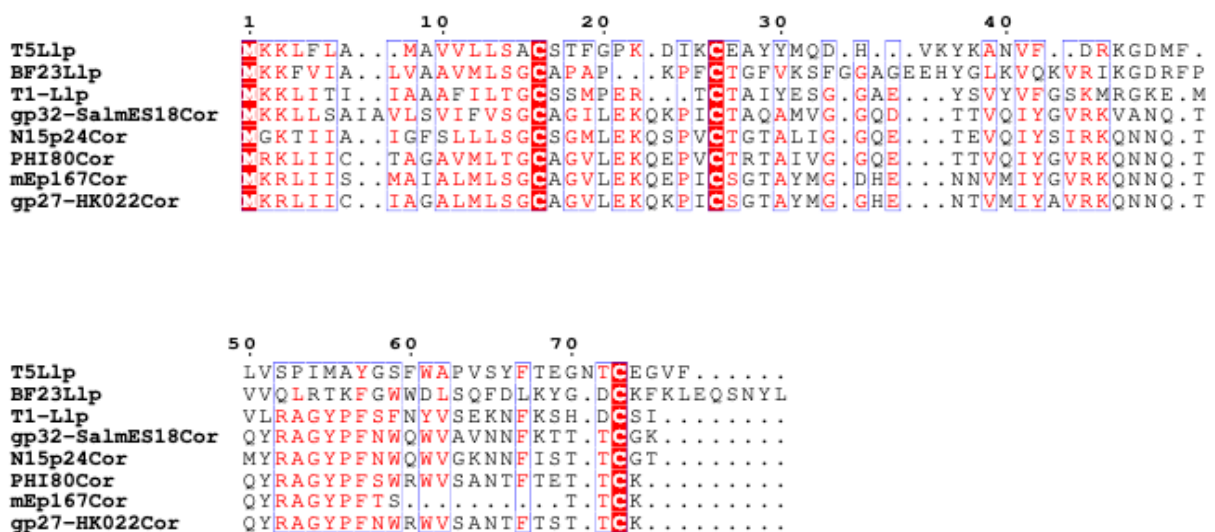
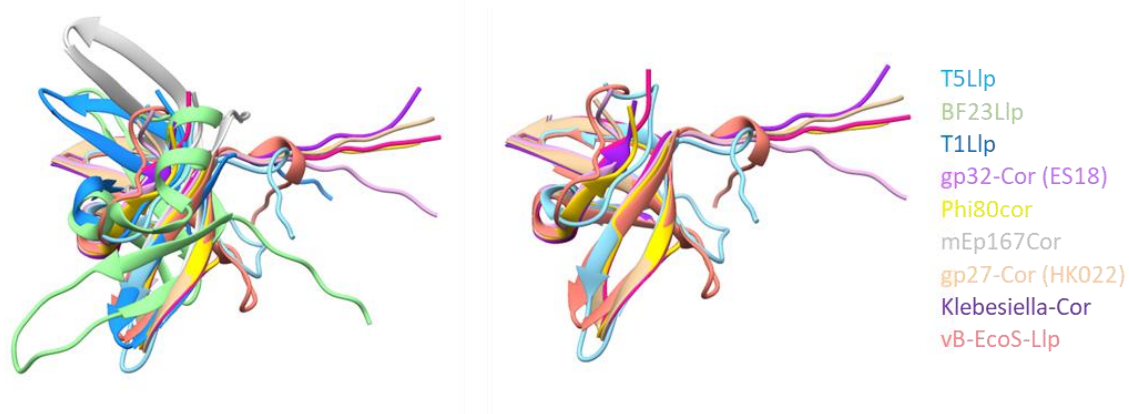


Figure 5-20 : Alignement des protéines Cor et Llp-like avec Llp de T5.

Les acides aminés conservés sont surlignés en rouge, ceux de même nature sont écrits en rouge et les cadres bleus indiquent les séquences consensus.

Quand on s'intéresse à la structures des protéines de phage, on retrouve des repliements communs à plusieurs phages de familles différentes et ce malgré une grande diversité de séquences <sup>102</sup>. Nous avons utilisé Alphafoldcollab <sup>103</sup> pour prédire les structures de certaines protéines Cor et Llp-like à partir des séquences des protéines matures (c'est-à-dire à partir de la première cystéine présente dans la séquence). Les phages contenant les gènes de Cor choisis pour cette analyse ne ciblent pas tous FhuA comme récepteur. BF23 cible ButB, Es18 cible un FhuA mais de Salmonella, T5, T1, φ80, mEp167 et HK022 ciblent FhuA. Les récepteurs des phages N15 et vB-EcoS ne sont pas connus, et la protéine Cor de Klebsiella est probablement issue d'un prophage non identifié. Dans l'ensemble, le niveau de confiance des prédictions est bon, entre 80 et 100 %, sauf pour la partie N-terminale de toutes les séquences soumises (environ 40%). Les deux protéines les plus éloignées de Llp en structure sont aussi les deux protéines avec le niveau de confiance le plus bas (Llp<sub>BF23</sub> et Llp<sub>mEp167</sub>) avec un RMSD sur 9 acides aminés d'environ 1,2 Å pour les deux structures et de 8,9 Å et 12,7 Å respectivement sur l'ensemble des acides aminés (63 résidus). Pour les autres protéines testées, le RMSD varie de 0,9 à 1,09 Å sur 24 à 32 acides aminés, et sur l'ensemble des résidus de 6,4 à 7,5 Å (Figure 5-21).

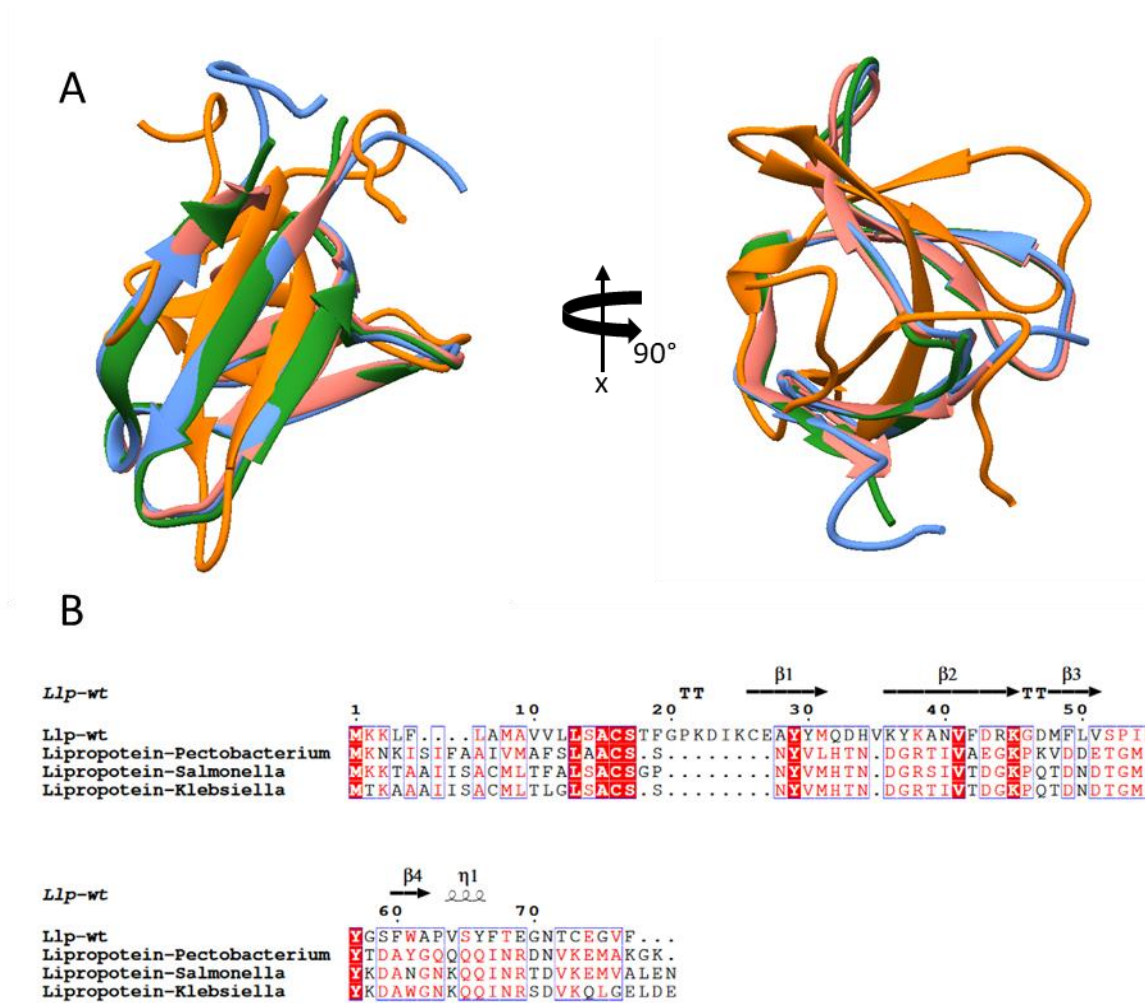


**Figure 5-21 : Superposition des prédictions de structure de protéine Cor et Llp-like par rapport à la structure de Llp obtenue en RMN.**

A. superposition de l'ensemble des structures. B. superposition des 5 structure les plus proche de Llp (ont été exclues BF23Llp, T1Llp, mEp167Cor).

Pour obtenir des informations fonctionnelles sur la structure, nous avons recherché des similitudes avec des structures connues. Quand on soumet la structure de Llp à DALI<sup>104</sup> pour rechercher des protéines au repliement similaire, aucune protéine de phage n'est présente dans les résultats. En revanche les résultats montrent des similarités avec des lipoprotéines hypothétiques chez *Salmonella enterica subsp. enterica serovar Typhimurium str. LT2*, *Klebsiella pneumoniae* et *Pectobacterium atrosepticum*. La fonction de ces protéines n'est pas connue. Les gènes adjacents autour de ces gènes de lipoprotéine semble indiquer qu'ils n'appartenaient pas à un prophage pour toutes les souches. Les trois structures de ces lipoprotéines supposées sont extrêmement proche les unes des autres, mais aussi proches de Sol-Llp (Figure 5-22 (A)). L'alignement de séquence montre que la lipobox est très conservée ainsi que la première cystéine après la séquence signal. Le reste de la protéine est aussi relativement proche avec 4 acides aminés très conservés et des séquences avec des acides aminés de même nature (Figure 5-22 (B)). En revanche Llp est la seule à contenir 2 autres cystéines qui forment un pont disulfure. Étonnamment le pont disulfure se trouve dans la région où l'alignement de structure est le meilleur.

Nous avons donc résolu une nouvelle structure de protéine de bactériophage. La structure seule ne donne pas beaucoup d'information sur le fonctionnement de Llp pour bloquer le récepteur FhuA. En revanche, grâce à la résolution de cette structure, les résultats d'interaction pourront donner des informations précieuses sur la formation du complexe FhuA-Llp (section 6 Chapitre III : Etude du complexe FhuA-Llp).



**Figure 5-22 : Comparisons structurales entre Sol-Llp et des lipoprotéines aux repliments similaires.**

A. superposition des structures des lipoprotéines (2RA2 bleu, 5CA1 vert et 3BDU saumon) et de la structure de Sol-Llp (orange).  
 B. alignement des séquences primaires des protéines avec leur séquence signal.

## 6 Chapitre III : Etude du complexe FhuA-Llp

---




## 6.1 Introduction

Cette partie de ma thèse traitera de la caractérisation du complexe formé entre FhuA et Llp. Les études biochimiques sont menées sur les protéines purifiées. Dans cette partie je détaillerai la purification de Llp sous sa forme acylée (Ac-Llp) (voir section 1 pour la purification de FhuA et section 2 pour la purification de Sol-Llp). La difficulté dans l'étude de ce complexe tient au fait que les deux protéines étudiées sont des protéines membranaires. Pour limiter les variables il est préférable de mettre les deux protéines dans le même détergent qui doit donc convenir aux deux protéines. De plus, la présence de détergent influence beaucoup les techniques biophysiques qui permettent d'étudier ce complexe. Néanmoins, la convergence des données obtenues par un large panel de techniques peut permettre de conforter les résultats quand ceux-ci ne sont pas d'une très bonne qualité.

## 6.2 Mise au point de la production et purification de Ac-Llp

### 6.2.1 Construction des plasmides de surexpression de Ac-Llp

Comme déjà mentionné, la surexpression de la protéine Llp, sous forme acylée ou non, a été décrite respectivement en 1998<sup>51</sup> et 2003<sup>71</sup>. Les laboratoires ayant mené ces études ont été contactés. L'équipe à laquelle appartenait le Kaspar Locher, institut Pasteur, Paris, a été dissoute, plus aucun plasmide n'est aujourd'hui disponible. En revanche, un plasmide était encore disponible à l'Institut Pasteur, pCHAP6507 (CD-Llp)<sup>71</sup>, et nous a gracieusement été envoyé par Olivera Francetic, institut Pasteur, Paris. La production décrite dans l'article de Pedruzzi et al.<sup>51</sup> a montré un rendement assez faible (1,8 mg de Llp acylée pour 30 g de culot bactérien), et le plasmide pCHAP6507 n'est pas optimisé pour la surexpression. En outre, la surexpression de Llp a été décrite comme toxique pour les bactéries. Le vecteur d'expression retenu pour Ac-Llp a été le pET20b, car, grâce à une fuite du promoteur, il permet une production plus lente et plus importante de protéines bien repliées<sup>51</sup>. Dans ce contexte, nous avons choisi de multiples stratégies de clonage (Tableau 6-1).

Schéma de la construction	Nom de la protéine	Objectif de la construction
	Ac-Llp	Production de la protéine acylée dans sa forme WT à l'exception de l'étiquette histidine
	Opéron Llp	Production de l'opéron Llp, avec une étiquette polyhistidine sur Llp. Les autres Orfs pourraient être des chaperonnes et aider l'expression et le repliement de Llp.
	Ac-Llp (SS-PuIA)	Production de la protéine acylée. Utilisation de la séquence signal de PuIA qui pourrait améliorer la production de Llp.

**Tableau 6-1 : Construction et stratégie pour l'expression de Llp acylée.**

Le tableau présente les différentes constructions essayées dans le projet et leur intérêt. Le schéma de construction est représenté. SS : séquence signal d'adressage au périplasme, C : cystéine, His : séquence codant pour une étiquette polyhistidine, Orf : Open reading frame.

Comme pour Sol-Llp (voir 5.2.1) le clonage est réalisé grâce à 2 PCRs. Les séquences du gène *lIp*, ou l'opéron entier de *lIp* en ont été amplifiées par une première PCR à partir du génome du phage T5. Pour la construction SS-PuIA Ac-Llp cette première PCR a permis de récupérer l'ensemble de la séquence signal de PuIA et du gène de *lIp* à partir de la construction présente dans le plasmide pCHAP6507. Pour toutes les constructions, les PCRs ont permis l'obtention d'un fragment d'ADN à la taille attendue. Une deuxième PCR permet d'insérer les produits des premières PCRs dans le vecteur d'expression pET20b grâce à des séquences homologues entre les produits PCRs et le vecteur. De nombreux clones ont été obtenus après transformation bactérienne. Un plasmide sans mutation a été obtenu pour Ac-Llp, en revanche les constructions PuIA-Ac-Llp et opéron ont échoué et le choix a été fait de ne pas poursuivre les efforts de clonage sur ces deux constructions (plus d'information voir annexe 3).

Pour l'analyse de l'interaction de Ac-Llp avec FhuA *in vivo*, des souches sans polymérase T7 ont été utilisées. Pour permettre l'expression de Ac-Llp dans ces souches, Ac-Llp a été re-clonée dans le plasmide pASK7+ et le plasmide pASK-IBA3C contenant un promoteur reconnu par la polymérase de *E. coli*. Le plasmide pASK7+ (résistance ampicilline) a été utilisé dans les souches pathogènes (*Shigella* et *Salmonella*) et le plasmide pASK-IBA3C (résistance chloramphénicol) a été utilisée dans la souche AW740 (résistance tétracycline) contenant le plasmide pHX405-FhuA (résistance ampicilline).

## 6.2.2 Production de la protéine Ac-Llp

### 6.2.2.1 Culture bactérienne, expression et solubilisation

Dans un premier temps, la croissance des souches bactériennes a été contrôlée pour vérifier l'absence de toxicité de la protéine Ac-Llp. Nous avons décidé de tester deux souches d'expression. La souche BL21(DE3) possédant le gène de la polymérase T7 sous le contrôle de l'opéron lactose qui convient bien à l'expression de protéines non toxiques. La souche C43(DE3) dérive de la souche C41 (DE3), elle-même dérivant de la souche BL21(DE3). Ces souches ont été isolées pour tolérer l'expression de protéines membranaires ayant un effet toxique. Dans les deux souches testées, la croissance est restée en phase exponentielle après l'ajout de 1 mM d'IPTG (Figure 6-1). La surexpression de Ac-Llp n'est donc pas toxique et n'empêche pas la croissance bactérienne.

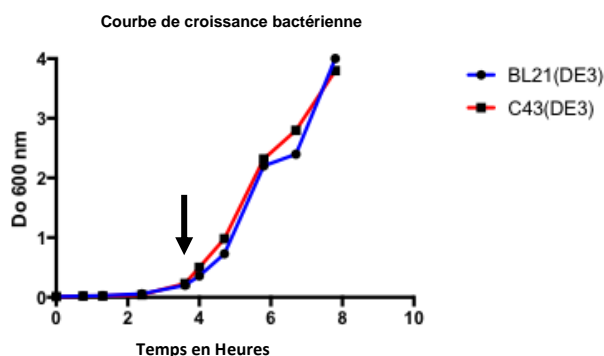
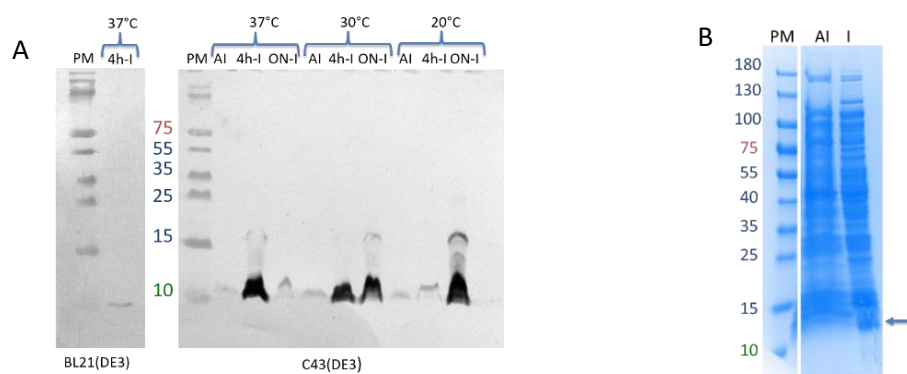


Figure 6-1 : courbe de croissance bactérienne de souches BL21(DE3) et C43(DE3) transformée avec le plasmide codant pour Ac-Llp. Courbe d'évolution de la densité optique à 600 nm, en fonction du temps. La flèche indique l'induction par 1 mM d'IPTG.

L'expression de Ac-Llp a été contrôlée par immuno-révélation sur bactéries totales avec des anticorps dirigés contre l'étiquette polyhistidine. Pour les souches BL21(DE3) et C43(DE3), une bande à la taille attendue est visible dans toutes les conditions testées. Le niveau d'expression après 4h d'induction à 37°C est bien plus élevé dans la souche C43(DE3) que dans la souche BL21 (DE3) (figure 6-2 (A)). En revanche, après une induction à 37°C sur une nuit entière dans C43(DE3), il y a une dégradation de la protéine. Ce phénomène n'est pas observé pour une induction à 30°C ou 20°C. La surexpression de Ac-Llp dans la souche C43(DE3) est confirmée par un gel d'acrylamide coloré au bleu de Coomassie (figure 6-2 (B)). Pour BL21(DE3), les gels d'acrylamide colorés au bleu de Coomassie n'ont pas permis d'identifier cette bande supplémentaire du fait de la faible expression. La condition retenue est alors une croissance de la C43(DE3) à 37°C avec une induction à l'IPTG pendant 4h (figure 6-2).



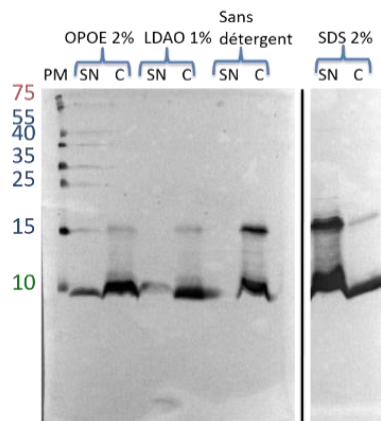
**Figure 6-2 : Test d'expression de la protéine Ac-Llp dans les souches BL21(DE3) et C43(DE3).**

*A – immuno-révélation réalisée avec un anticorps dirigé contre l'étiquette polyhistidine sur des bactéries totales exprimant Ac-Llp après sur gel SDS. PM : poids moléculaire, AI : avant induction, 4h-I : 4 heures et ON-I : une nuit d'induction avec 1mM d'IPTG. Gauche : test d'expression dans BL21(DE3) à 37°C. Droite : test d'expression dans C43(DE3) à 37, 30 ou 20°C. La taille attendue de Llp est de 7,8 kDa. En raison d'un pont disulfure intermoléculaire artefactuel lié à la préparation de l'échantillon pour le gel, une bande à environ 14 kDa est également visible. B – gel coloré au bleu de Coomassie des bactéries C43(DE3) totales avant induction (AI) ou après 4h d'induction à 37°C (I).*

Pour les protéines membranaires, la purification des membranes et la solubilisation de la protéine permet une prés purification. En particulier si le détergent solubilise préférentiellement la protéine ciblée la prés purification sera meilleure. Des essais de solubilisation ont donc été menés sur les membranes contenant la protéine Ac-Llp.

Ac-Llp est une protéine adressée à la membrane externe tout comme FhuA dont la solubilisation a été optimisée en 2 étapes. Le protocole de FhuA a été testé sur les membranes contenant Ac-Llp. Pour des membranes à 2 mg/mL de protéines totales, une première solubilisation permet d'éliminer les protéines de la membrane interne avec une solubilisation à l'OPOE (81,5 mM) puis une deuxième solubilisation avec le LDAO (43,6 mM) solubilise les protéines de la membrane externe. Comme attendu, Ac-Llp est principalement retrouvé dans le culot de solubilisation OPOE (Figure 6-3). En revanche, il est surprenant de noter qu'Ac-Llp est principalement dans le culot de la solubilisation au LDAO et donc pas solubilisée dans ces conditions. Le témoin de solubilisation avec 2 % de SDS ne solubilise que partiellement Ac-Llp ce qui n'est pas habituel (Figure 6-3).

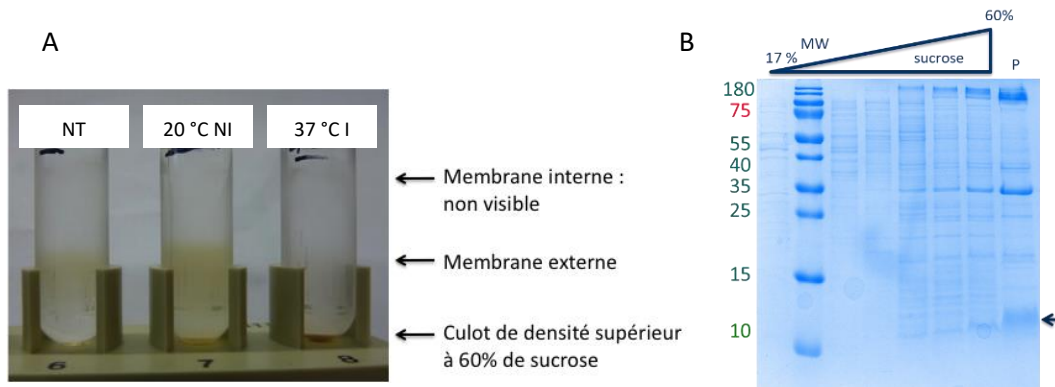




**Figure 6-3 : Test de solubilisation de la protéine Ac-Llp suivi par électrophorèse sur gel.** Immunorévélation avec un anticorps dirigé contre l'étiquette polyhistidine. Les membranes totales sont solubilisées par 81,5 mM d'OPOE. Le matériel non solubilisé est ensuite solubilisé par 43,6 mM LDAO. SN : surnageant et C : culot de solubilisation. PM : poids moléculaire. Témoin négatif sans détergent et témoin positif solubilisation au SDS sur les membranes totales.

Afin de trouver des conditions de solubilisation de la protéine, des tests complémentaires ont été effectués sur les membranes totales. Les membranes ont été incubées, entre 2 mg/mL et 0,1 mg/mL de protéines totales, avec divers détergents. Ont été testés le LDAO et l'OPOE jusqu'à 174,5 mM et 163 mM respectivement, un mélange LDAO/OPOE 174,5 mM / 163 mM, la FosCholine 113,8 mM, l'Octyl Glucoside (OG) 136,8 mM, l'Octyl Thio Glucoside (OTG) 1297 mM, le Triton X-100 30 mM et le Sodium lauroyl sarcosinate (sarkosyl) 136,3 mM. Le résultat le plus prometteur a été obtenu avec le Sarkosyl à 136,3 mM. Cependant, la majorité d'Ac-Llp est présent dans la fraction insoluble (non montré). Ceci indique qu'une grande partie de Llp est probablement agrégée dans les membranes.

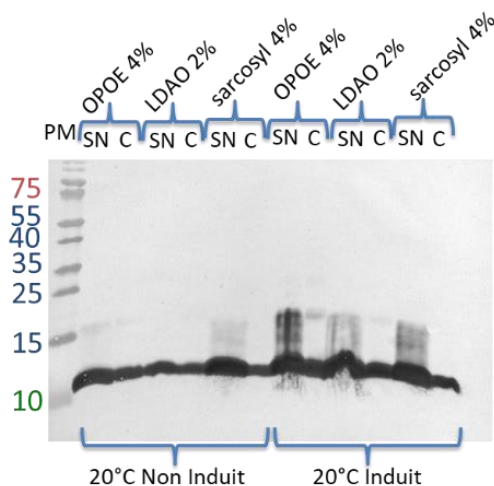
Une analyse des membranes issue de cultures à plus basse température avec ou sans induction a donc été réalisée. Les bactéries incubées à 37°C, 30 °C et 20 °C avec ou sans induction à 1 mM d'IPTG ont toutes surexprimé Ac-Llp de façon comparable (non montré). Les membranes de C43(DE3), transformées ou non, après une culture de 5 h à 37°C, ou après une nuit à 20°C en absence ou en présence d'IPTG, ont été centrifugées sur des gradients de saccharose afin de séparer les membranes externes et internes. Quel que soit l'échantillon, aucune bande correspondant à la membrane interne, attendue à environ 37-42% du gradient de saccharose n'a pu être observée à l'œil nu (Figure 6-4 (A)). Cependant, l'analyse sur un gel d'acrylamide coloré au bleu de Coomassie de cette zone révèle bien les protéines de la membrane interne (figure 6-4 (B)). Une bande à environ 54-57% du gradient, correspondant à la membrane externe, est observée avec les membranes issues de cellules non transformées et avec les cellules dans lesquelles Ac-Llp a été produite à 20 °C sans induction (Figure 6-4 (A)). L'analyse sur gel d'acrylamide de cette partie du gradient montre que la protéine Ac-Llp co-migre avec la membrane externe (figure 6-4 (B)). En revanche, la membrane externe contenant Ac-Llp induite à 37°C forme un culot au fond du gradient de saccharose, dans lequel se retrouve Ac-Llp également. Cela indique que ces membranes sont plus denses que des membranes externes habituelles, confirmant qu'Ac-Llp est probablement agrégée dans la membrane externe quand elle est produite à 37°C avec induction (figure 6-4 (A)).



**Figure 6-4 : Séparation des membranes de C43(DE3), transformées ou non, 37°C ou 20 °C, en présence ou en absence d'IPTG sur gradient de saccharose et analyse sur gel.**

A : gradient linéaire de saccharose de 17 à 60% - centrifugés à 268000 g pendant 6h à 4°C, membrane déposée en haut du gradient. NT : membranes issues de cellules non transformées ; 20°C NI : sans induction, croissance à 20°C ; 37°C I : induction 1 mM IPTG, 5h à 37°C. B. Gel des fractions du gradient de saccharose pour les membranes non induites et incubées à 20°C. Dépôt de gauche à droite de 20 à 60% de saccharose. P : culot formé au fond du gradient. La flèche indique la position de migration de Ac-Llp.

Les membranes contenant Ac-Llp produite à 20°C et 30°C avec ou sans induction ont été solubilisées à une concentration de 1 mg/mL de protéine totale avec de l'OPOE 163 mM, du LDAO 87,2 mM ou du sarcosyl 163,3 mM, par une incubation de 1h à 37°C. Les résultats ont montré que la majorité de la protéine se trouve dans la fraction soluble pour les détergents OPOE, LDAO, et sarcosyl. Une partie d'Ac-Llp reste néanmoins dans la fraction insoluble et le rendement de solubilisation est légèrement meilleur avec le sarcosyl pour les membranes à 20°C non induits (figure 6-5 / résultats de solubilisation à 30°C non montrés). Les conditions retenues pour la production d'Ac-Llp sont donc une culture à 20°C sans induction sur une nuit puis une solubilisation sur des membranes à 1 mg/mL de protéines totales avec du sarcosyl 136,3 mM.



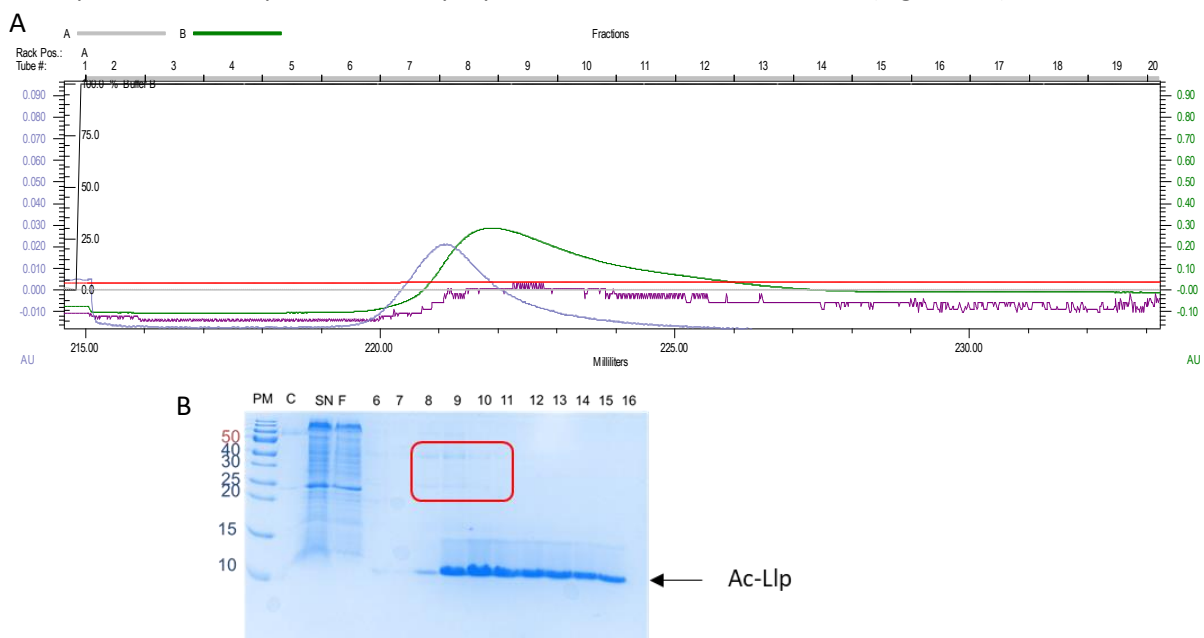
**Figure 6-5 : Solubilisation des membranes issues d'une culture à 20°C avec ou sans induction**

Immunorévélation avec un anticorps anti polyhistidine. Surnageant (SN) et culot (C) de solubilisation. PM : poids moléculaire. Solubilisation de membrane totale avec OPOE 163 mM, LDAO 87,2 mM, sarcosyl 136,3 mM.

### 6.2.2.2 Mise au point de la purification

Les premiers tests de purification ont été réalisés sur 250 ml de culture. La présence d'une étiquette polyhistidine sur Ac-Llp, permet de réaliser une chromatographie d'affinité sur ions métalliques (Nickel) immobilisés (IMAC) après l'étape de solubilisation. Lorsque Ac-Llp est immobilisé sur la colonne IMAC, le sarcosyl est remplacé par le DDM (détergent plus doux). Afin de déterminer la quantité d'imidazole à utiliser pour éluer la protéine. Nous avons réalisé un gradient d'imidazole de 0 à 400 mM en présence de 300 mM de NaCl sur 50 ml. Un contaminant connu de FhuA de la membrane interne d'*E. coli*, la quinone oxydoréductase, absorbe à 412 nm. Ac-Llp étant également extraite des membranes d'*E. coli* les longueurs d'onde à 280 nm et 412 nm ont été suivies pour les chromatogrammes de purification. Le profil d'élution de la chromatographie montre de l'absorbance à 280 nm à partir de 40 mM d'imidazole et jusqu'à la fin du gradient à 400 mM. De 40 à 200 mM d'imidazole, il y a aussi de l'absorbance à 412 nm mais plus après 200mM. Ac-Llp est présente dans les fractions de 200 à 400 mM d'imidazole (non montré).

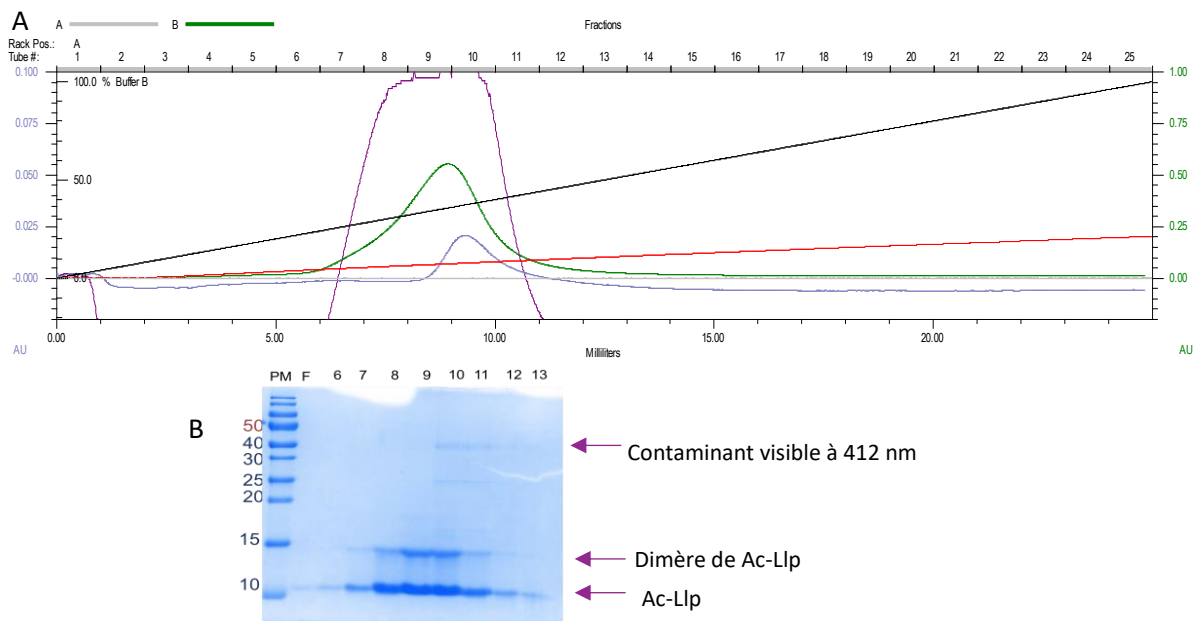
Après le premier essai de purification, l'élution d'Ac-Llp a été faite avec une concentration d'imidazole fixée à 250 mM. Pour limiter la dilution à réaliser avant la chromatographie d'échange d'ions, la concentration en NaCl a été baissée à 100 mM dans le tampon d'élution de l'IMAC. Cependant, contrairement à ce qui était attendu, Ac-Llp n'a pas élué en un pic bien défini, mais sur environ 7 ml dans les fractions 8 à 15 (Figure 6-6). En revanche, le contaminant principal suivi à 412 nm est relativement bien séparé à cette étape et faible en proportion dans les fractions 7 à 10 (Figure 6-6).



**Figure 6-6 : Vérification de la purification de la protéine Ac-Llp sur colonne d'affinité au Nickel, élue à 250 mM Imidazol.**

A. Chromatogramme d'élution de la colonne NiNTA. La courbe en vert est l'absorbance à 280 nm, la courbe en bleu l'absorbance à 412 nm, la courbe rouge la conductimétrie et la courbe violette l'absorbance à 212 nm. B. gel d'acrylamide, PM : poids moléculaire, C : culot de solubilisation, SN : surnageant de solubilisation, F : non retenu, 6 à 16 numéros de fraction correspondant aux fractions du chromatogramme A. Le contaminant visible à 412 nm est encadré en rouge.

Les fractions 10 à 17 ont été rassemblées et diluées 5 fois avant d'être déposées sur une colonne échangeuse d'anion. Le profil d'élution de la deuxième colonne montre une mauvaise séparation d'Ac-Llp et du contaminant suivi à 412 nm (Figure 6-7). En considérant que l'ensemble de la protéine déposée sur la colonne est de la Llp, le rendement de purification de cette colonne a été évalué à 58% ce qui est peu au vu de la pureté déjà obtenue avec la colonne au nickel (figure 6-7 (B)). De plus, l'analyse sur gel a confirmé qu'une partie d'Ac-Llp ne s'était pas accrochée à la colonne échangeuse d'anions (présence dans le non-retenu, la protéine déposée ayant été diluée (Figure 6-7 (B))). 2,9 mg de protéine ont été obtenus (sur l'ensemble des fractions calculé avec  $\epsilon = 13075 \text{ M}^{-1} \cdot \text{cm}^{-1}$ ) dont 1,4 mg de protéine totalement pure (Figure 6-7). Comme pour Sol-Llp (section 5.2.2.1), en présence de réducteur Ac-Llp on note la présence d'un dimère artificiel. La purification étant réalisée sur 250 ml de culture, le rendement de cette purification est de 5,6 mg de protéine pure par litre de culture. Cette première purification a permis d'obtenir Ac-Llp avec une grande proportion de monomères. Cependant, contrairement à Sol-Llp, on note la présence de dimère, trimère, et tétramère en absence de réducteur peu ou pas visible en bleu de Coomassie mais bien visible en immunorévélation (voir ci-dessous et figure 6-8). La quantité de protéine obtenue a été suffisante pour son analyse biochimique et les premiers tests d'interaction par SPR.



**Figure 6-7 : Vérification de la purification de la protéine Ac-Llp sur colonne échangeuse d'ions.**

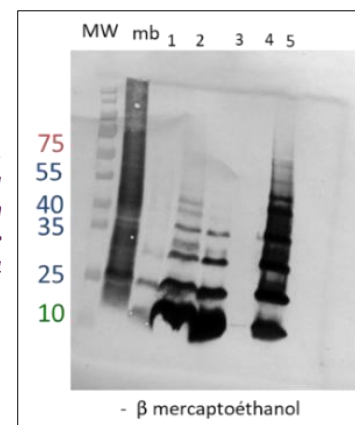
A. Chromatogramme d'élution de la protéine Ac-Llp sur colonne échangeuse d'ions par un gradient de NaCl de 0 à 1 M. La courbe en vert est l'absorbance à 280 nm, la courbe en bleu l'absorbance à 412 nm, la courbe rouge la conductimétrie et la courbe violette l'absorbance à 212 nm et courbe noire le pourcentage de tampon d'élution. B. Contrôle sur gel de la pureté de Ac-Llp en présence de  $\beta$ -mercaptoéthanol. PM : poids moléculaire, F : non retenue, 6 à 13 numéros des fractions correspondant aux fractions du chromatogramme.

Malgré les résultats très prometteurs de cette première purification, la purification de Ac-Llp a posé beaucoup de problèmes notamment de reproductibilité. En effet, deux autres purifications de Ac-Llp ont été réalisées sur 1L de culture. Le rendement final a été beaucoup plus faible, soit 0,25 mg de protéine pour 1 L de culture bactérienne, ce qui peut être expliqué par une faible production de protéine pendant la culture (environ 50 mg de protéine totale contre plus de 200 mg de protéine totale habituellement). De plus, la proportion d'oligomères en absence de réducteur par rapport au monomère était bien plus importante dans les deux dernières purifications (puits 5 Figure 6-8) que dans la première purification (puits 2 et 3 Figure 6-8). L'hypothèse posée a été que la culture de 1L ne permettait pas une oxygénation suffisante pour la formation correcte de ponts disulfures intramoléculaires, conduisant à la formation de ponts disulfures intermoléculaires. Deux conditions de culture bactérienne ont été comparées pour essayer de résoudre ce problème. Une culture « classique » de 1 L dans un Erlen de 3 L et une culture mieux oxygénée de 0,5 L dans un Erlen à ailettes de 2 L. Dans les deux conditions, des oligomères ont été retrouvés dans les membranes totales. Cependant, à la fin de la purification, dans les deux cas une majorité de monomères est observée. Le monomère serait donc préférentiellement solubilisé. En fin de purification, les proportions des différentes formes sont les mêmes, quelle que soit l'oxygénation de la culture. Une purification à partir de bactéries cultivées avec une bonne oxygénation, mais purifiée avec des tampons à pH 9 avait également montré des oligomères avec une faible proportion de monomères. Les groupements thiol -SH ont typiquement un pKa autour de 8,3 et la formation des ponts disulfure est donc favorisée à partir de pH 8. Une fois qu'un pont disulfure est formé il ne peut être réduit que par un agent réducteur. Pendant la production de Llp *in vivo* le système de formation des ponts disulfure est probablement saturé empêchant la formation de certains ponts disulfure. Quand la purification est effectuée à pH 9 les ponts disulfures sont favorisés. De fait, comme la Llp est petite, la formation des ponts disulfure *in vitro* peut être intermoléculaire.

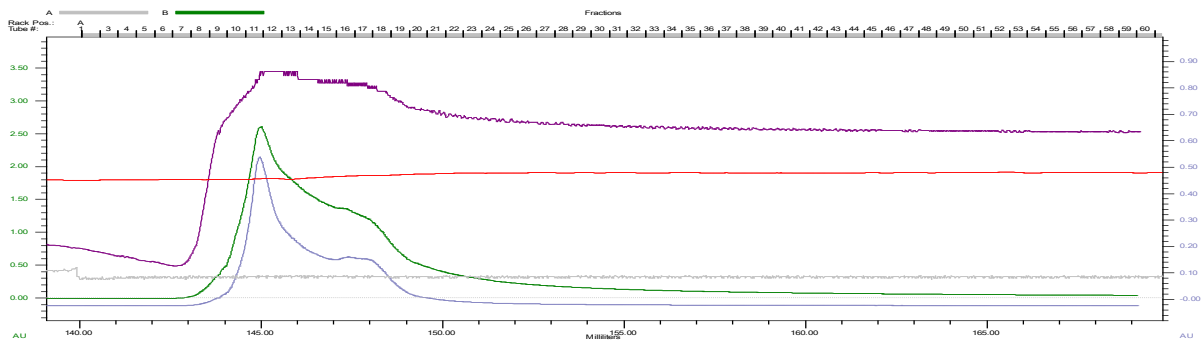
La proportion de monomères étant satisfaisante dans les purifications réalisées à pH 8,5 ; le protocole a été fixé avec un tampon à ce pH et la deuxième étape de purification n'est plus réalisée, la pureté des échantillons étant satisfaisante pour les études biochimiques. En revanche, la question se pose sur l'homogénéité de la protéine pour les études structurales, car même à pH 8,5 des oligomères sont présents.

**Figure 6-8 : Vérification de l'état d'oligomérisation de la protéine Ac-Llp.**

Immunorévélation avec un anticorps anti polyhistidine de différents lots de Ac-Llp, déposés sans réducteur. MW : poids moléculaires ; mb : membranes totales conservées à 4 °C pendant 1 an ; 1 : Ac-Llp purifiée en DDM, conservée 1 an à 4 °C ; 2 : Ac-Llp purifiée en LMGN, conservée 1 an à 4 °C ; 3 : Ac-Llp purifiée en DDM dans un tampon NaF, conservée 7 mois à 4 °C ; 4 : Sol-Llp purifiée sans étiquette histidine ayant été conservée 6 mois à 4 °C ; 5 : Ac-Llp purifiée en DDM, conservée 1 semaine à 4 °C

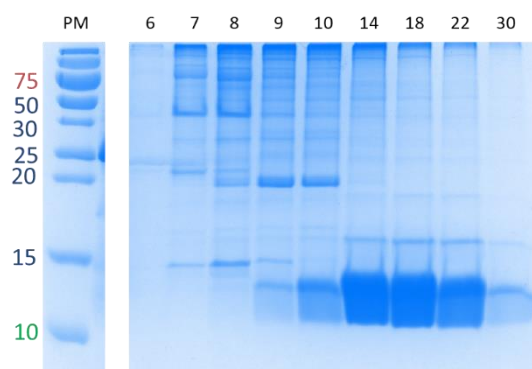


Une troisième série de purification pour l'étude structurale de Ac-Llp a posé un autre problème majeur. Des cultures de 1L avec un bon rendement de protéine totale (~ 300 mg de protéine totale) ont été utilisées pour une purification à plus grande échelle pour les études structurales de Ac-Llp avec FhuA. Afin d'éviter que le pic d'éluion de l'IMAC ne traîne sur plusieurs millilitres, les concentrations d'imidazole et de sel ont été augmentées respectivement à 500 mM et 150 mM. Ceci a produit un pic plus fin d'environ 3 ml avec un épaulement sur 8 ml et une DO maximum à 280 nm élevés (~2,5 AU Figures 6-9). Cependant, la protéine a précipité après l'éluion même si une dilution par 2 est réalisée.



**Figure 6-9 : Chromatogramme d'éluion de la protéine Ac-Llp sur colonne d'affinité au nickel.** Elution par une étape d'imidazole à 500 mM. La courbe en vert est l'absorbance à 280 nm, la courbe en bleu l'absorbance à 412 nm, le courbe rouge la conductimétrie et la courbe violette l'absorbance à 212 nm.

Le problème a été résolu en éluant la protéine avec une solution d'EDTA à 50 mM sans NaCl. Le contrôle des fractions sur un gel permet de voir que de nombreuses protéines sont éluées dès la fraction 7 et principalement jusqu'à la fraction 10 alors que Llp commence à être éluée en fraction 9 et devient majoritaire qu'après la fraction 10 et jusqu'à la fraction 30 (Figure 6-10). Cela suggère une interaction aspécifique entre Ac-Llp et la matrice de la colonne de Nickel. Cette purification a permis d'obtenir 2 mg de protéine raisonnablement pure. Cependant, l'effet de l'EDTA sur la protéine sera discuté dans le paragraphe 6.3.2.

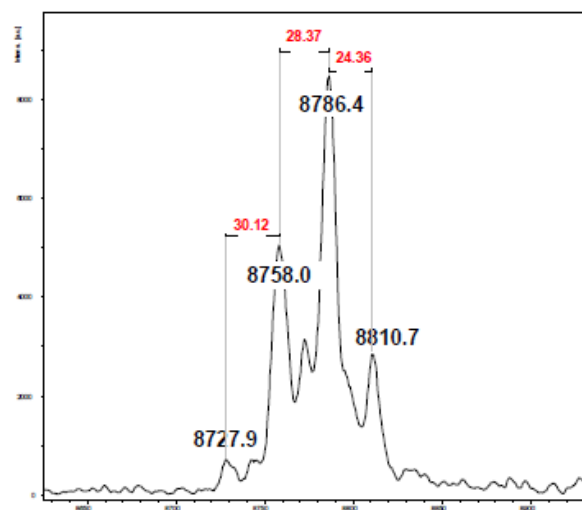


**Figure 6-10 : Contrôle de la pureté de Ac-Llp après purification sur une colonne d'affinité au nickel avec une éluion à l'EDTA.** PM : poids moléculaire, 6 à 30 numéros de fractions correspondant aux fractions du chromatogramme de la figure 6-9.

## 6.3 Caractérisation biochimique de Ac-Llp

### 6.3.1 Spectrométrie de masse

La spectrométrie de masse est un outil puissant pour analyser les protéines, car elle permet d'accéder à la masse précise des protéines. Ac-Llp a été contrôlée en spectrométrie de masse après purification. Les masses trouvées sont 8727,9 ; 8758,0 ; 8786,4 ; 8810,7 Da. On attend une masse de 7971,99 Da pour la protéine, plus 38 Da pour le glycérol (92 Da moins 3 H<sub>2</sub>O pour la formation des liaisons à la protéine et aux acides gras), plus 3 acides gras inconnus, mais issus de la membrane d'*E. coli* dont la masse moyenne est d'environ 250 Da, soit un total de masse attendu autour de 8759,99 Da. Le spectre présente plusieurs pics et témoigne d'une acylation avec des acides gras de différentes longueurs et insaturations (Figure 6-11). Cette hétérogénéité ne permet pas de conclure quant à la formation du pont disulfure dans la protéine. Un écart de masse de 2 carbones est de l'ordre de 24 Da comme observé entre le dernier et l'avant-dernier pic (Figure 6-11). Entre les pics 1 et 2 puis entre les pics 2 et 3 l'écart de masse correspond à 2 carbones ou 2 carbones et 2 ou 6 double liaison de différence. En revanche, les informations fournies par cette analyse ne permettent pas d'identifier précisément les acides gras présents, aucune investigation n'a été menée en ce sens. Les résultats de spectrométrie de masse ont été réalisés sur une protéine qui a été conservée environ 2 mois à 4°C. Ceci indique que la protéine est stable dans le temps et n'est pas dégradée contrairement à Sol-Llp qui perd rapidement ses 4 premiers acides aminés. Nous suggérons que la micelle stabilise ces acides aminés dans la forme acylée. En effet dans la forme soluble les acides aminés qui sont tronqués sont : une Glycine en position 1 qui n'existe pas dans la forme acylée et qui remplace une cystéine sur laquelle est greffée l'acylation, puis les acides aminés STF (position 2-4) qui sont probablement protégés par leur proximité de la micelle de détergent.



**Figure 6-11 : Caractérisation par spectrométrie de masse de la protéine Ac-Llp après purification.**

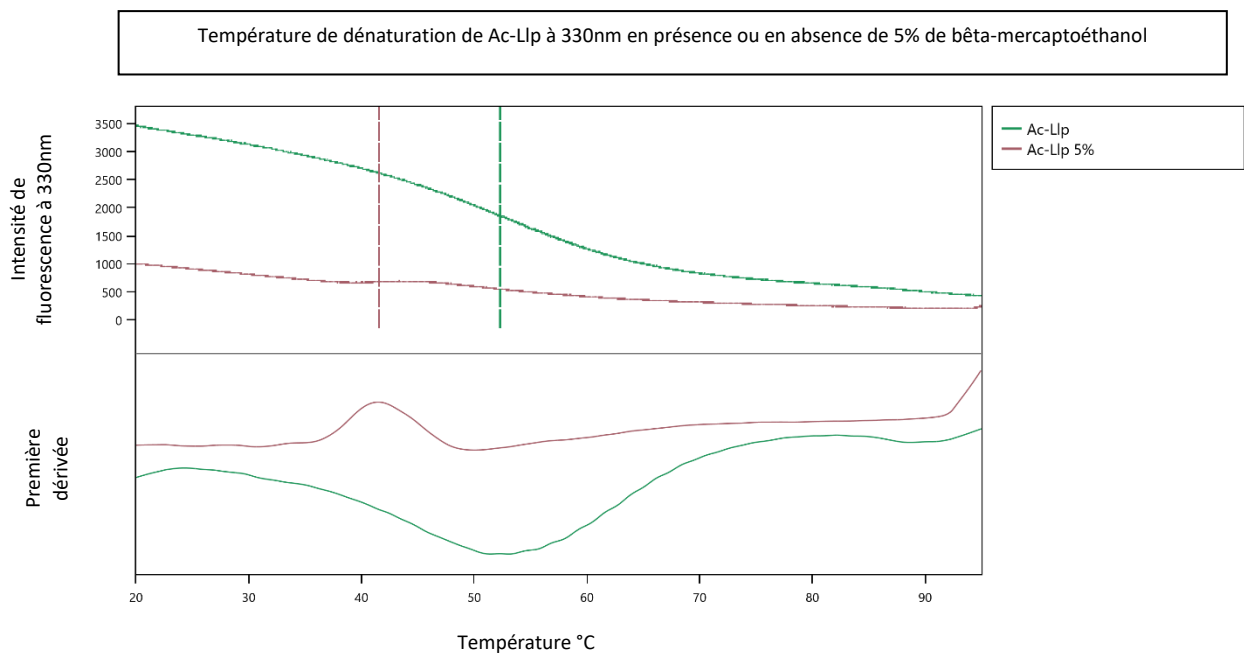
Abcisses : m/z, ordonnées : intensité en A.U., chaque pic présente la masse en Da, et les différences de masse entre les pics sont indiquées en rouge.



### 6.3.2 Stabilité thermique

La stabilité thermique de Ac-Llp a été déterminée par nano-DSF. Les résultats ont montré que la protéine est stable avec une  $T_m$  supérieure à 52°C en DDM. La stabilité dans différents détergents a été évaluée, Ac-Llp est plus stable en OG et en LMNG (58°C) qu'en DDM. Cependant, l'OG a une CMC très élevée, donc ce détergent a été écarté pour la manipulation de Ac-Llp. Le LMNG apporte aussi une stabilité plus importante, mais sa CMC très basse ne permet plus d'échanger le détergent et son prix est beaucoup plus élevé que le DDM. Au vu de nos besoins et de la stabilité de Ac-Llp, le DDM a été jugé suffisant pour les tests biochimiques.

Les  $T_m$  des protéines sont mesurés avec 5% ou sans  $\beta$ -mercaptoéthanol. Comme pour la protéine sous forme soluble, la présence de  $\beta$ -mercaptoéthanol a un fort impact sur la température de dénaturation puisqu'elle entraîne une chute de 10,7 °C (Figure 6-12). Ces résultats confirment la présence du pont disulfure dans la protéine. En outre, ils montrent que le pont disulfure est important pour la stabilité de la protéine.



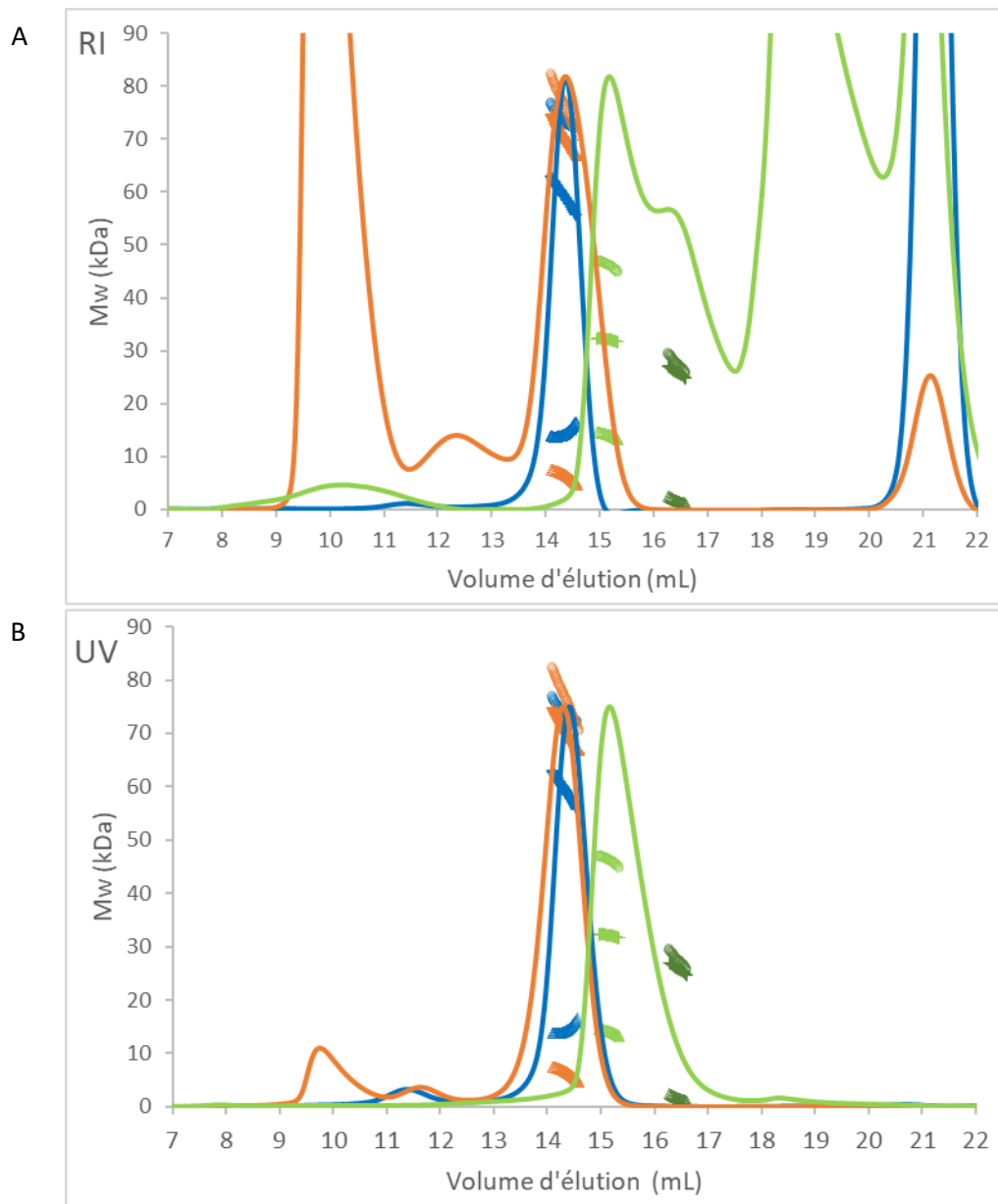
**Figure 6-12 : Courbe de suivi de dénaturation de la protéine Ac-Llp avec 5% ou sans  $\beta$ -mercaptoéthanol par nano-DSF.**  
Abscisse : température en °C, ordonnées bloc du haut : intensité de fluorescence en nombre d'évènements par minute, bloc du bas, dérivée première du signal.

Dans<sup>51</sup>, des expériences de dichroïsme circulaire ont montré que la présence EDTA changeait le profil de structures secondaires. Cela suggère la liaison d'un cation divalent. L'impact de l'EDTA sur la stabilité de Ac-Llp a donc été testé. Les mesures ont montré qu'en présence de 200 mM EDTA le  $T_m$  était diminué de 6,2 °C. Cependant ces résultats ne sont pas reproductibles et d'autres investigations ont été menées pour savoir si la protéine lie un cation divalent, sans résultat pour le moment.

### 6.3.3 Étude de l'état d'oligomérisation par SEC-MALS

Pour continuer à caractériser Ac-Llp, une étude de l'état d'oligomérisation en solution a été menée par une analyse en chromatographie d'exclusion de taille couplée à plusieurs détecteur (diffusion de lumière, indice de réfraction, UV), permettant de déduire la masse des particules éluées (SEC-MALS). Cette technique permet également de caractériser des complexes à plusieurs partenaires.

Dans le cas de Ac-Llp, le SEC-MALS a été réalisé pour caractériser à la fois l'état d'oligomérisation de la protéine, mais aussi la quantité de détergent lié (Bdet en g de détergent / g de protéine). Les échantillons en DDM sont issus du chromatographie échangeuse d'ion et n'ont pas été concentrés, le signal des micelles libres et donc a priori invisible. En DDM, le chromatogramme montre un seul pic avec un volume d'élution correspondant à la micelle libre. L'analyse des signaux donne une masse totale de 83 kDa pour le complexe protéine-détergent élué, dont 23 kDa de protéine (masse théorique du monomère 7,9 kDa) et 60 kDa de DDM correspondant à la masse de la micelle libre. Ces résultats, obtenus à une concentration de DDM de 0,4 mM (CMC + 0,2 mM), suggèrent donc qu'il y aurait 3 Ac-Llp par micelle. Ces résultats étant surprenants, d'autres détergents ont été testés pour vérifier le nombre de Ac-Llp par micelle. Dans un premier temps, une concentration plus importante de DDM a été testée : 2,2 mM (CMC + 2 mM). La masse totale du complexe observée est plus faible, 74 kDa avec 14 kDa attribués à la protéine et 60 kDa au détergent. À cette concentration, on aurait donc 2 Ac-Llp par micelle avec Bdet de 4,1 g/g (Figure 6-13). Des résultats similaires ont été obtenus avec une concentration de DDM à 5,2 mM (CMC + 5mM). Ac-Llp a été dilué sous la CMC du DDM avec un tampon contenant de l'OG ou du LMNG, puis reconcentré pour être analysé. Un résultat similaire au DDM à 2,2 mM a été observé en OG 21 mM (CMC + 1 mM) où les micelles sont séparées du complexe Ac-Llp-détergent, avec un Bdet de 2,3 g/g (Figure 6-13). En revanche, en LMNG à 1 mM (CMC + 1 mM), il n'y aurait qu'une Ac-Llp par micelle avec un Bdet de 11,1 g/g (Figure 6-13). Le Bdet très élevé du LMNG est probablement dû au fait que ce détergent forme des micelles hétérogènes<sup>105</sup> qui ne seraient pas entièrement séparées du complexe Ac-Llp-détergent. Dans ce cas, la stœchiométrie apparente de 1 Ac-Llp par micelle serait sous-estimée. Cette étude suggère que, sur un même lot de protéines, le nombre de Ac-Llp par micelle varie en fonction de la concentration et du type de détergent (Figure 6-13 (C)). Chaque détergent est caractérisé par un nombre d'agrégation et des formes différentes de micelles. La forme et la taille des micelles peut influencer la distribution de Llp dans les micelles. La Figure 6-13 montre, pour les pics principaux, des masses homogènes. Ceci est plutôt surprenant sachant qu'il existe plusieurs états d'oligomérisation liés aux ponts disulfures visibles sur gel (Figure 6-8). Une explication viendrait que la masse de la micelle, importante par rapport à la masse de Ac-Llp, pourrait masquer la faible proportion des différents états d'oligomérisation liés aux ponts disulfures intermoléculaires et conduire à une masse calculée qui serait en fait un masse moyen apparent.



C

Détergent CMC + 2 mM	Masse totale (kDa)	Masse détergent (kDa)	Masse protéine (kDa)	Nombre de Llp par micelle	Bdet apparent (en g de détergent/g de protéine)
DDM	73,7	59,3	14,4	2	4,1
OG	46	32,1	13,9	2	2.3
LMNG	76,4	70,1	6,3	1	11.1

**Figure 6-13 : Résultats du SEC-MALS sur la protéine Ac-Llp dans différents détergents.**

Expérience de SEC-MALS réalisée sur une SD 200 avec un tampon 20 mM tris, 150 mM NaCl et 2,2 mM DDM ou 22 mM OG ou 1 mM LMNG.

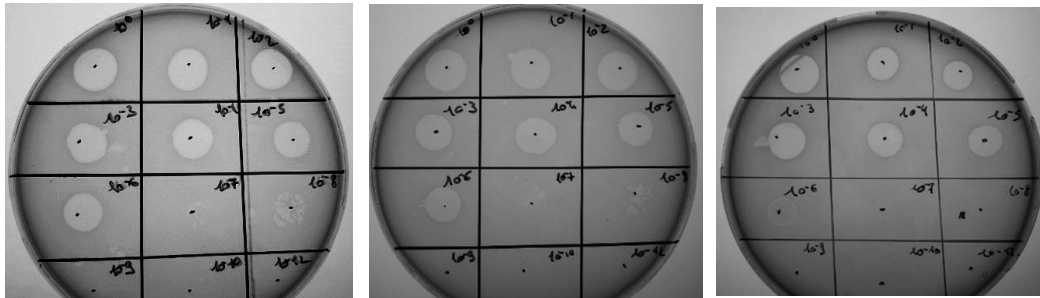
A. Chromatogramme du SEC-MALS détection IR normalisée sur le pic de protéine, détecte le complexe protéine-détergent, en sus des micelle libre OG (16,5 ml) et LMNG (10 ml). Il n'y a pas d'excès de DDM libre car l'échantillon n'a pas été concentré. B. Chromatogramme du SEC-MALS détection UV normalisée sur le pic de protéine, ne détecte que le complexe protéine-détergent. Analyse de masse : o : masse totale, Δ : masse protéine, + : masse détergent reporté sur le chromatogramme et le tableau C.

## 6.4 Interactions FhuA et Llp

FhuA et Ac-Llp sont deux protéines membranaires. L'étude de l'interaction entre ces deux molécules est un défi, car la présence des micelles de détergent impacte la quasi-totalité des techniques permettant d'explorer les interactions protéine-protéine. Néanmoins, l'interaction entre les deux protéines et sa caractérisation ont été menées par différentes méthodes.

### 6.4.1 Interactions *in vivo*

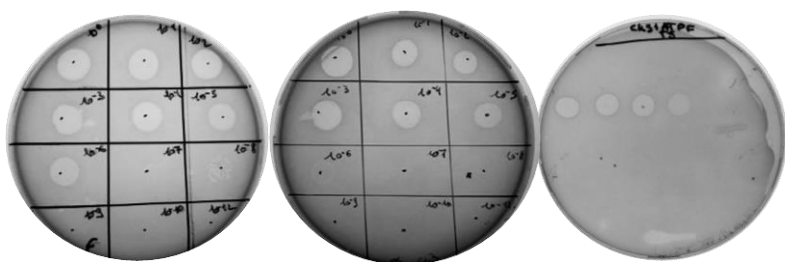
Pour le test phénotypique, le titre du phage T5 a été vérifié sur la souche F. Le titre de T5 purifiée est de  $2 \times 10^{12}$  UFP (unité formant des plages de lyse). La souche BL21(DE3) est aussi sensible au phage T5 que la souche F, avec un titre similaire. La souche C43(DE3) est également sensible au bactériophage T5, mais le titre est moindre, à  $2 \times 10^{10}$  UFP. Le bipyridyl est un chélateur du fer qui permet d'induire la production de FhuA ; en effet la production de cette protéine est sous contrôle d'un promoteur inductible en condition de carence en fer. Nous avons surexprimé la protéine FhuA dans la souche C43(DE3) en présence de bipyridyl et contrôlé la sensibilité de la souche au bactériophage T5. La sensibilité de C43(DE3) surexprimant FhuA n'a pas changé indiquant que la baisse de sensibilité de la souche C43(DE3) n'est pas liée à une expression plus faible de FhuA (Figure 6-14).



**Figure 6-14 : Test phénotypique de sensibilité au phage T5 des souches F, BL21(DE3) et C43(DE3).**

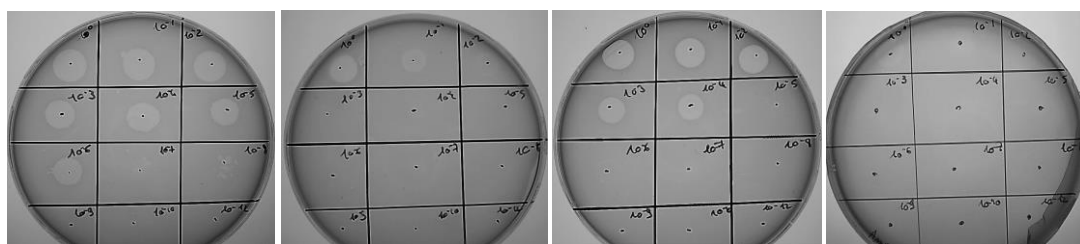
De gauche à droite : souches F, BL21 et C43. La sensibilité des bactéries au phage T5 est visible grâce à la lyse des bactéries formant des plages de lyse. Le titre est calculé approximativement à partir de la dernière dilution montrant des plages de lyses.

Nous avons cherché à comprendre pourquoi la souche C43(DE3) est moins sensible au bactériophage T5. La souche C43(DE3) présente une quantité importante de membranes internes quand la sous-unité b de l'ATP (appeler aussi AtpF) est surexprimée<sup>106</sup>. Cependant avant la surexpression d'AtpF une petite partie de la population bactérienne présente déjà une prolifération de membranes internes (Francesca Zito, communication personnelle). En revanche quand AtpF est surexprimé une grande quantité de la population bactérienne présente ce phénotype. Nous avons émis l'hypothèse que, dans le cas où le bactériophage infecte une cellule dans laquelle FhuA est proche d'une invagination de la membrane interne, le canal permettant d'injecter l'ADN à l'intérieur de la bactérie ne pourrait pas traverser les deux membranes ce qui aboutirait à une infection abortive, se traduisant par un titre moins élevé au niveau macroscopique. Pour valider cette hypothèse, nous avons testé la sensibilité de la souche C43(DE3) surexprimant AtpF, pour vérifier si le titre était affecté par l'augmentation du réseau de membrane interne dans une plus grande proportion de la population bactérienne. Quand la souche C43(DE3) exprime AtpF le titre est encore réduit de  $10^2$  ce qui tendrait à confirmer notre hypothèse.



**Figure 6-15 : Test phénotypique de sensibilité au phage T5 de la souche C43 contenant ou non le plasmide d'expression codant pour AtpF.**  
De gauche à droite : souches F (référence), C43 vide, C43 avec le plasmide AtpF.

Le test phénotypique consiste à vérifier si l'expression de Ac-Llp est capable de remplir son rôle physiologique, c'est-à-dire protéger la cellule d'une sur-infection au bactériophage T5. Les souches d'expression portant le plasmides Ac-Llp, ont été testées pour leur sensibilité au phage T5. Pour la souche BL21(DE3) portant le plasmide codant pour Ac-Llp, le titre est de  $2 \times 10^4$  UFP, soit un décalage de  $10^8$  par rapport à la souche sans plasmide. Il y a donc bien une protection, au moins partielle, de la bactérie au phage T5. Quand la souche C43(DE3) porte le plasmide codant pour Ac-Llp, alors les bactéries ne sont plus sensibles au bactériophage T5, le décalage est de  $10^{10}$  (Figure 6-16). Il y a donc expression fonctionnelle de la protéine Ac-Llp dans les deux souches, permettant de conclure que la protéine Llp codée par le plasmide est bien adressée au périplasm, acylée, ancrée dans la membrane externe et fonctionnelle.



**Figure 6-16 : test phénotypique par sensibilité au phage T5 des souches BL21 (DE3) et C43 (DE3) contenant ou non le plasmide d'expression codant pour Ac-Llp.**  
De gauche à droite : souches BL21(DE3) non transformée, BL21(DE3) transformée avec le plasmide codant pour Ac-Llp, C43(DE3) non transformée, C43(DE3) transformée avec le plasmide codant pour Ac-Llp.

Le niveau d'expression de Ac-Llp en BL21(DE3) ou en C43(DE3) est très différent (Figure 6-2 dans 6.2.2.1). De même, le niveau de protection des souches BL21(DE3) et de C43(DE3) exprimant Ac-Llp est différent. Nous avons supposé que le ratio FhuA / Llp pouvait être important pour la protection de la cellule. Pour vérifier notre hypothèse, nous avons joué sur les niveaux de sur-expression des protéines en induisant ou non FhuA (avec du bipyridyl) et Ac-Llp (avec ou sans IPTG). Pour C43(DE3) quelle que soit la condition, la souche est complètement protégée indiquant que toutes les FhuA sont en interaction avec Ac-Llp. Pour BL21(DE3) l'expression de Ac-Llp est beaucoup plus faible. Lorsque FhuA et Llp sont toutes les deux non induites, ou toutes les deux surexprimées, le titre apparent est d'environ  $2 \times 10^5$  phage/ml. Lorsque seul FhuA est surexprimée le titre apparent est de  $8 \times 10^8$  phage /ml et quand seul Ac-Llp est surexprimée le titre apparent est de  $2 \times 10^4$  phage/ml. Ceci indique que le ratio FhuA / Llp dans la cellule joue sur la protection de la souche. Cependant, ces résultats montrent également qu'une faible expression de Ac-Llp est suffisante pour protéger la cellule avec un décalage de  $10^4$  de sensibilité, ce qui est suffisant *in vivo* pour protéger de la surinfection, car les concentrations de phages sont plus faibles qu'en laboratoire.

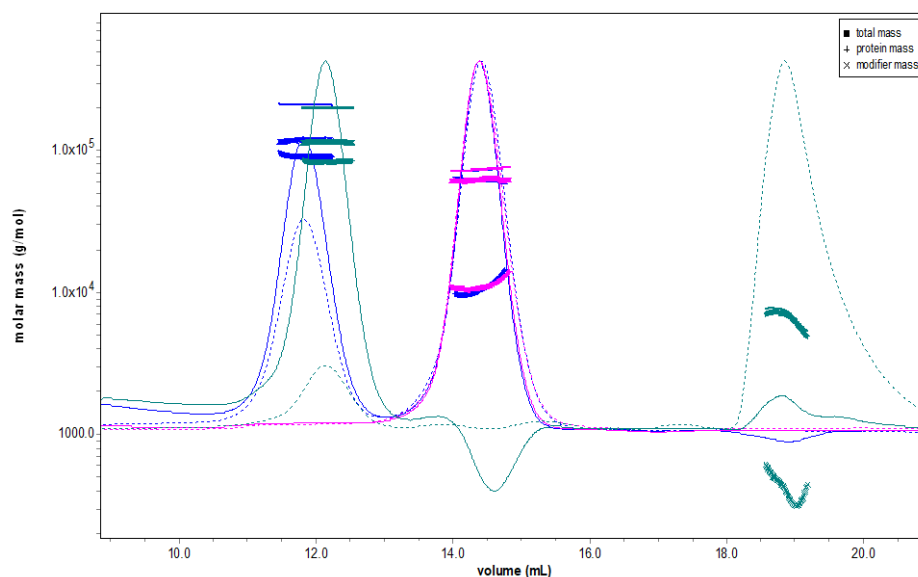
## 6.4.2 Interaction *in vitro*

L'objectif étant d'explorer la structure du complexe FhuA : Ac-Llp pour comprendre l'inhibition de FhuA par Llp, la formation et la fonctionnalité du complexe ont également été explorées *in vitro*.

### 6.4.2.1 SEC-MALS

L'ensemble des échantillons ont été analysés en DDM. Les échantillons contenant FhuA seule, Ac-Llp seule ou un mélange de FhuA : Ac-Llp à différents ratios molaires (1 : 8, 1 : 20 et 1 : 40) sont analysés en SEC-MALS. Pour tous les échantillons, un pic est élué à 15 ml correspondant à la micelle libre (FhuA et FhuA-Ac-Llp) ou liant Ac-Llp (Ac-Llp seule). Le pic de FhuA monomère est élué à 12,3 mL avec une masse totale de 193,3 kDa (Figure 6-17), dont 78,4 kDa attribué à FhuA (Masse théorique 80 kDa, figure 6-17) et 115,9 kDa au DDM (Figure 6-17), donnant un Bdet de 1,5 g de DDM/g de protéine (valeur raisonnable). Un pic additionnel est élué à 12,2 mL avec une masse totale de 210 kDa dont 120 kDa de DDM et une masse de protéine de 91 kDa pour le ratio 1 : 8 - pouvant correspondre à 1 FhuA pour 2 Ac-Llp. Pour le ratio 1 : 40 la masse calculée est de 87,8 kDa pouvant correspondre à 1 FhuA pour 1 Ac-Llp. Dans les deux cas le complexe a un Bdet de 1,3 g/g (Figure 6-17). Ces expériences sont reproductibles et montrent la formation du complexe FhuA-Llp. Les différentes stœchiométries entre les différents ratios de FhuA et Ac-Llp, peut traduire un équilibre d'association- dissociation rapide ou être due à la présence de 2 Ac-Llp par micelle (section 6.3.3).

En revanche, les résultats n'ont pas montré de décalage du pic de FhuA en présence de Sol-Llp aux concentrations testées, et un pic de l'absorbance dans le volume total pouvant correspondre à Sol-Llp. Cette technique ne met pas en évidence la formation du complexe pour la forme soluble. Ceci peut être dû à un équilibre rapide et au faible changement de masse qui pourrait masquer un pic intermédiaire.

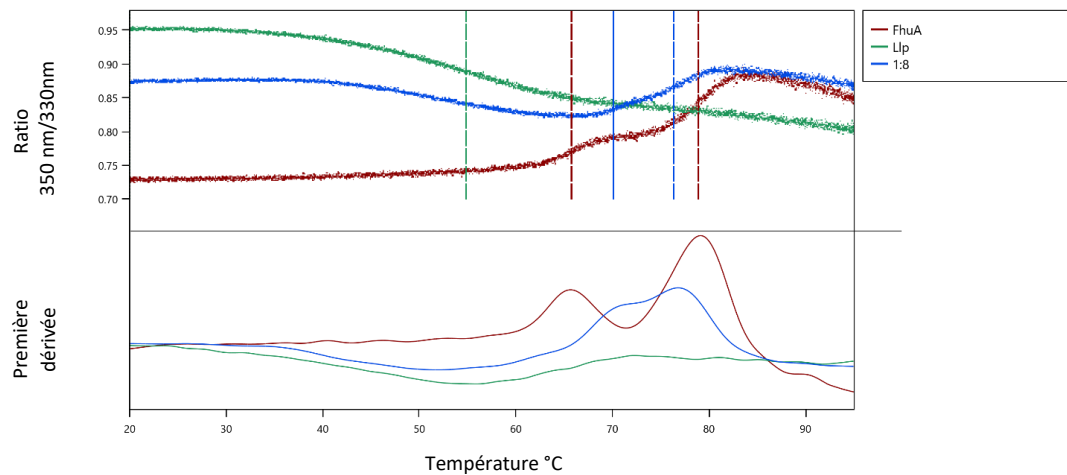


**Figure 6-17 : Expérience de SEC-MALS réalisées sur des mélanges FhuA-Ac-Llp et FhuA-Sol-Llp à un ratio 1/8.**

La SEC est réalisée avec une colonne SD 200 équilibrée dans un tampon 25 mM tris, 150 mM NaCl, 2,2 mM DDM. En vert FhuA-Sol-Llp, en bleu FhuA-Ac-Llp et en rose Ac-Llp, signal de diffusion de lumière (LS, light scattering), carré ■ : masse totale, ligne pleine - : indice de réfraction, ligne pointillée ... 280 nm.

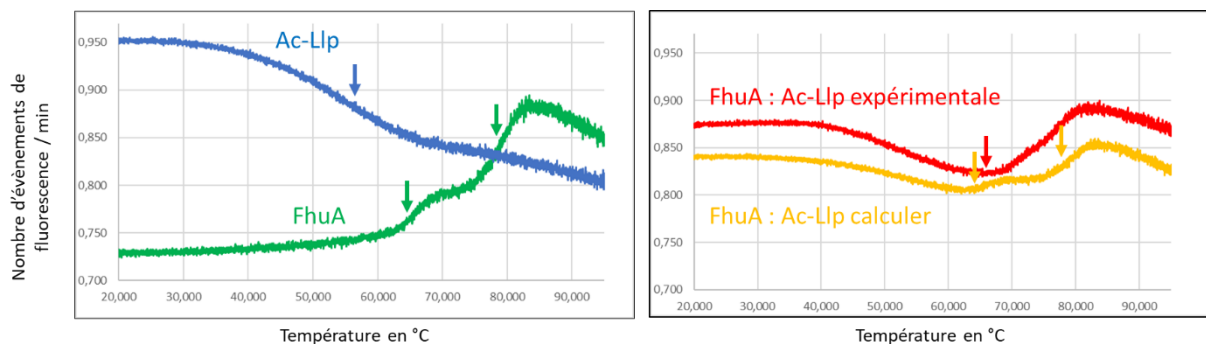
#### 6.4.2.2 Nano-DSF

Le Nano-DSF n'est pas la méthode optimale pour étudier la formation de complexe de protéines. Cependant, dans certains cas, cette technique peut mettre en évidence la formation d'un complexe, notamment quand le complexe est plus stable que les deux protéines séparées. FhuA a la particularité d'avoir deux transitions de dénaturation : la dénaturation du bouchon puis la dénaturation du tonneau bêta (Figure 6-18). Dans un mélange molaire, 1 :8 FhuA : Ac-Llp, la première transition de dénaturation est décalée vers des températures plus élevées, le  $T_m$  du bouchon augmente de 65,8°C à 70,2°C (Figure 6-18). Cela pourrait indiquer que l'interaction stabilise le bouchon de FhuA.



**Figure 6-18 : Courbe de suivi de dénaturation de la protéine Ac-Llp, FhuA et du complexe FhuA : Ac-Llp.**  
Abscisse : température en °C, ordonnées du panneau du haut : ratio de fluorescence 350nm/330nm. Panneau du bas : première dérivée de fluorescence, sans unité.

Cependant, les fluorescences de la protéine Ac-Llp et FhuA évoluent en sens inverse au cours de leurs dénaturations respectives. Pour s'assurer que le décalage du  $T_m$  du boudon de FhuA est dû à la liaison et non à une moyenne des deux signaux, les signaux de FhuA et Ac-Llp seules, à la même concentration, ont été moyennés et comparés aux signaux enregistrés. Les deux courbes ne sont pas identiques et les deux transitions de FhuA sont toujours visibles dans la courbe calculée, mais pas dans la courbe expérimentale (Figure 6-19). À un ratio molaire de 1 FhuA pour 8 Ac-Llp, la moyenne des signaux de FhuA et Ac-Llp n'a donc pas d'impact sur le signal de fluorescence de FhuA. Le nano-DSF peut être utilisé pour vérifier la formation du complexe FhuA : Ac-Llp.

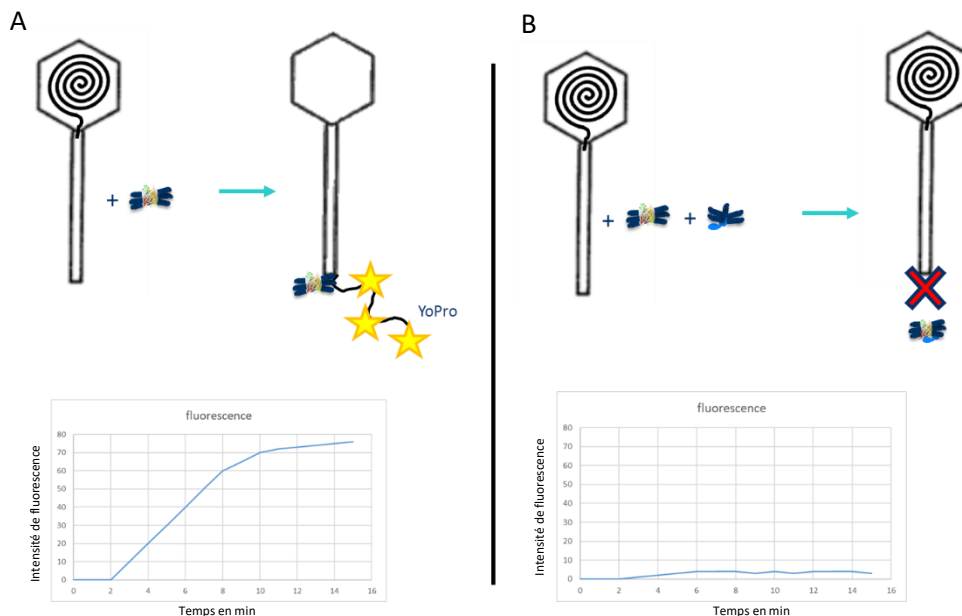


**Figure 6-19 : Courbe de suivi de dénaturation de la protéine Ac-Llp, FhuA et du complexe FhuA : Ac-Llp-expérimental ou calculée.**  
Abscisse : température en °C, ordonnées ratio intensité de fluorescence 350nm/330nm. Les flèches indiquent les évènements de dénaturation.



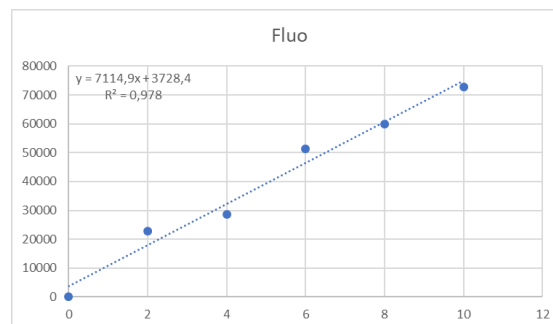
### 6.4.2.3 Protection du phage contre FhuA par Llp, détection de l'éjection d'ADN

Quand T5 est incubé avec FhuA solubilisée en détergent, il éjecte son ADN dans le tampon. La quantité d'ADN peut alors être mesurée avec une molécule fluorescente s'intercalant dans l'ADN (YoPro). Si T5 est incubé avec FhuA, la fluorescence du milieu doit donc augmenter. En revanche, si Ac-Llp est également présente et interagit avec FhuA, la fluorescence doit être stable, car les phages T5 seront protégés de l'interaction avec FhuA (Figure 6-20).



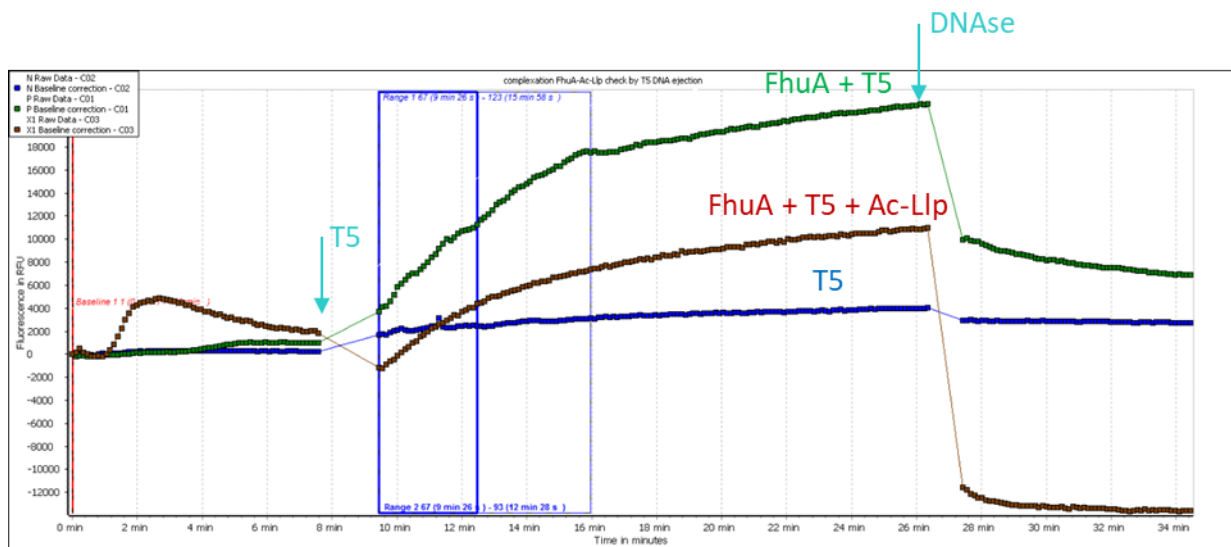
**Figure 6-20 :** Principe du test d'interaction *in vitro* entre FhuA et Ac-Llp avec T5 comme rapporteur de l'interaction. Résultats attendus de l'interaction entre FhuA et T5 (A) et T5, FhuA et Ac-Llp. La courbe présentée est une courbe théorique.

Dans un premier temps, une gamme d'ADN a été réalisée pour vérifier la proportionnalité entre la quantité d'ADN et la fluorescence. La régression linéaire sur la courbe d'étalonnage de la fluorescence donne une droite avec un  $R^2$  supérieur à 0,96. La fluorescence est donc proportionnelle à la quantité d'ADN présente pour une gamme allant de 2 à 10 ng d'ADN (Figure 6-21). Cette technique a l'avantage de permettre de voir la cinétique d'éjection de l'ADN du phage T5 et de voir si Ac-Llp peut influencer cette cinétique.



**Figure 6-21 :** Courbe d'étalonnage de l'intensité de fluorescence en fonction de la quantité d'ADN. Intensité de fluorescence (sans unité) en fonction de la quantité d'ADN en ng. Vérification de la proportionnalité par une droite de régression linéaire.

Le phage T5 a été incubé avec du Yo-pro et du détergent, la fluorescence reste inchangée au cours du temps et ne diminue pas avec l'ajout de DNase. En revanche si FhuA est ajoutée, à une concentration de 2  $\mu\text{M}$  au phage T5, la fluorescence augmente avec une pente de 2137 mAU/min et l'ajout de DNase fait baisser l'intensité de fluorescence. Si Ac-Llp est ajoutée à FhuA à une concentration de 100 nM finale dans le volume réactionnel (soit un ratio de 1 : 50), l'éjection d'ADN du T5 est plus lente avec une pente de 1302 mAU/min et la fluorescence diminue avec l'ajout de DNase (Figure 6-22). Ceci indique qu'à cette concentration de Ac-Llp une partie du FhuA est neutralisée par Ac-Llp, mais pas l'intégralité.



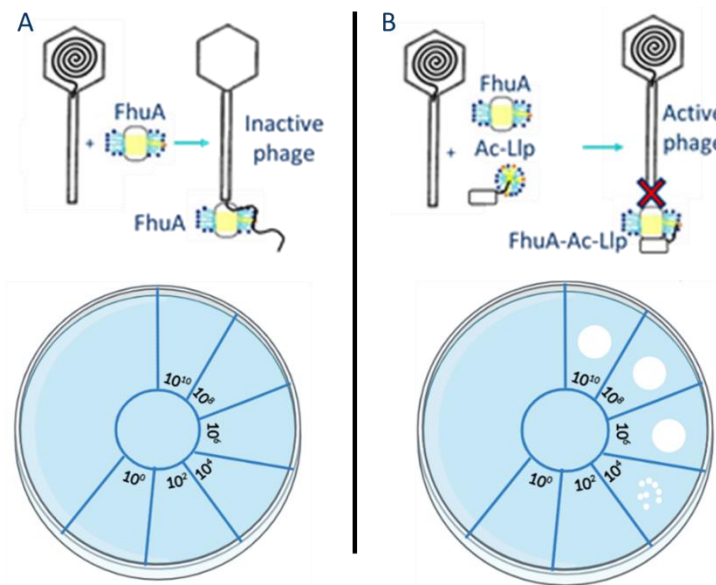
**Figure 6-22 : Suivi de l'éjection de l'ADN du phage T5 en présence de FhuA avec ou sans Ac-Llp.**

Courbe bleue témoin négatif vérification de l'intégrité du phage T5 en présence de détergent, courbe verte témoin positif suivi de l'éjection d'ADN en présence de 110 nM de FhuA, courbe rouge test d'interaction entre FhuA et Ac-Llp en présence de 100  $\mu\text{M}$  de Ac-Llp. Les courbes sont normalisées à une fluorescence de 0 à T0 min. La courbe FhuA-Ac-Llp avait un niveau initial de fluorescence élevé est instable ce qui fait penser qu'une bulle était présente au départ et a été enlevé lors de l'ajout de T5 et/ou de DNase ce qui explique le l'intensité de fluorescence finale soit plus basse que le niveau initial.

Cette expérience permet de conclure que Ac-Llp à 100  $\mu\text{M}$  bloque une partie de FhuA à une concentration de 2  $\mu\text{M}$ . Cependant les mesures ont été réalisées sur un lecteur de fluorescence en plaque 96 puits et les résultats sont fortement perturbés par la présence de bulles dues au détergent. Les résultats étant peu reproductibles, les écarts types n'ont pas pu être calculés. Au vu des problèmes techniques rencontrés, une seule concentration de Ac-Llp a été testée et l'interaction entre Sol-Llp et FhuA n'a pas été essayée par cette méthode, car nous supposons que l'absence de l'acylation fragilise la formation du complexe. Nous avons donc choisi de confirmer les résultats obtenus par une technique plus robuste.

#### 6.4.2.4 Protection du phage contre FhuA par Llp, détection par inactivation du phage

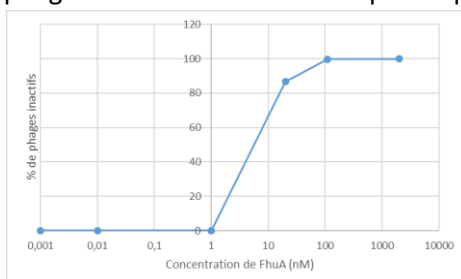
Pour contrôler l'intégrité/l'infectivité de T5, il suffit de le déposer sur un tapis bactérien d'une souche sensible. L'interaction entre FhuA et Ac-Llp a donc été explorée dans le même système que précédemment, mais l'infectivité de T5, qui est le rapporteur de cette interaction, a été testée directement sur la souche F : si FhuA est présente, elle interagit avec T5 qui est alors inactivé et ne forme pas de plages de lyse quand mis en présence d'un tapis bactérien de souche F. En revanche, si Llp bloque l'interaction entre FhuA et pb5, alors le phage reste intègre et peut infecter les bactéries de la souche F (Figure 6-23).



**Figure 6-23 : Principe du test in vitro d'interaction entre FhuA et Ac-Llp avec T5 comme reporteur de l'interaction.**

A. FhuA seule est incubée avec le phage T5 permettant l'interaction et l'éjection d'ADN dans le milieu réactionnel. Résultats attendus : T5 n'est plus actif : pas de plages de lyse sur la souche F. B. FhuA et Ac-Llp sont incubées ensemble avec le phage T5. Si Ac-Llp interagit avec FhuA, le phage T5 est protégé et peut infecter son hôte. Résultats attendus : plages de lyse sur la souche F.

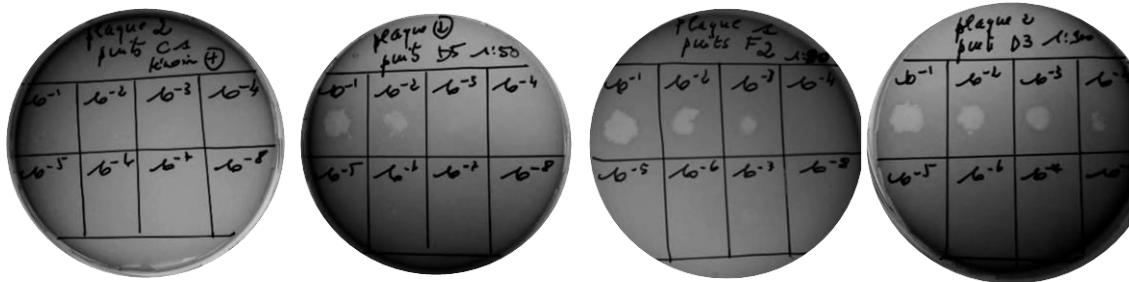
Dans un premier temps, une titration des bactériophages fonctionnels en fonction de la quantité de FhuA a été réalisée pour pouvoir interpréter la quantité de FhuA encore disponibles en fonction de la titration obtenue lors du test d'interaction. À des concentrations inférieures ou égales à 1 nM de FhuA, T5, à un titre de  $1,5 \times 10^5$  n'est pas inactivé, et ce même lors d'une incubation prolongée (3 semaines). Le phénomène d'inactivation du bactériophage T5 est quasiment un phénomène de tout ou rien. En effet, lorsque 150 000 phages sont incubés avec 1 nM FhuA tous les phages sont intacts alors que, 98,7 % des phages sont inactivés de présence de 20 nM de FhuA. En présence de 110 nM de FhuA, 99,9 % des phages sont inactivés et 100% pour 2  $\mu$ M de FhuA (Figure 6-24).



**Figure 6-24 : Pourcentage de phages T5 inactivés en fonction de la concentration en FhuA.**

Noter que la concentration de FhuA est sur une échelle logarithmique.

Les conditions testées pour la formation du complexe FhuA-Llp sont les mêmes que pour le test avec le Yo-pro. L'intégrité des phages au contact de FhuA ou du complexe FhuA-Llp a été contrôlé sur la souche F. Les ratios molaires FhuA : Ac-Llp 1/50, 1/130 et 1/200 avec une concentration de FhuA à 110 nM ont montré de la lyse jusqu'en dilution  $10^{-2}$  ce qui correspond à une concentration de FhuA libre entre 1 et 20 nM. Autrement dit, Ac-Llp a neutralisé au moins 82 % de FhuA. Pour un ratio de 1/900, la lyse est observée jusqu'à  $10^{-4}$ . Il reste donc moins de 1 nM de FhuA, ce qui veut dire que Ac-Llp a inactivé près de 100% de FhuA. Cependant la précision de cette manipulation n'est pas optimale, car elle fait intervenir de nombreux partenaires et de nombreux pipetages qui peuvent induire des erreurs. Par exemple un test avec le ratio 1/90 a montré de la lyse jusqu'en dilution  $10^{-3}$  indiquant entre 20 et 110 nM de FhuA libre, soit plus que dans le ratio 1/50 (Figure 6-25). Néanmoins cette expérience permet de conclure que Ac-Llp interagit avec FhuA de façon fonctionnelle *in vitro*. Pour Sol-Llp aucune protection n'a pu être observée jusqu'à une concentration de Sol-Llp de 3 mM. Ceci indique que l'acylation est importante pour stabiliser l'interaction entre FhuA et Llp.



**Figure 6-25 : Test fonctionnel d'interaction entre FhuA et Ac-Llp contrôlé par la protection du phage T5 contre l'inactivation par FhuA.**

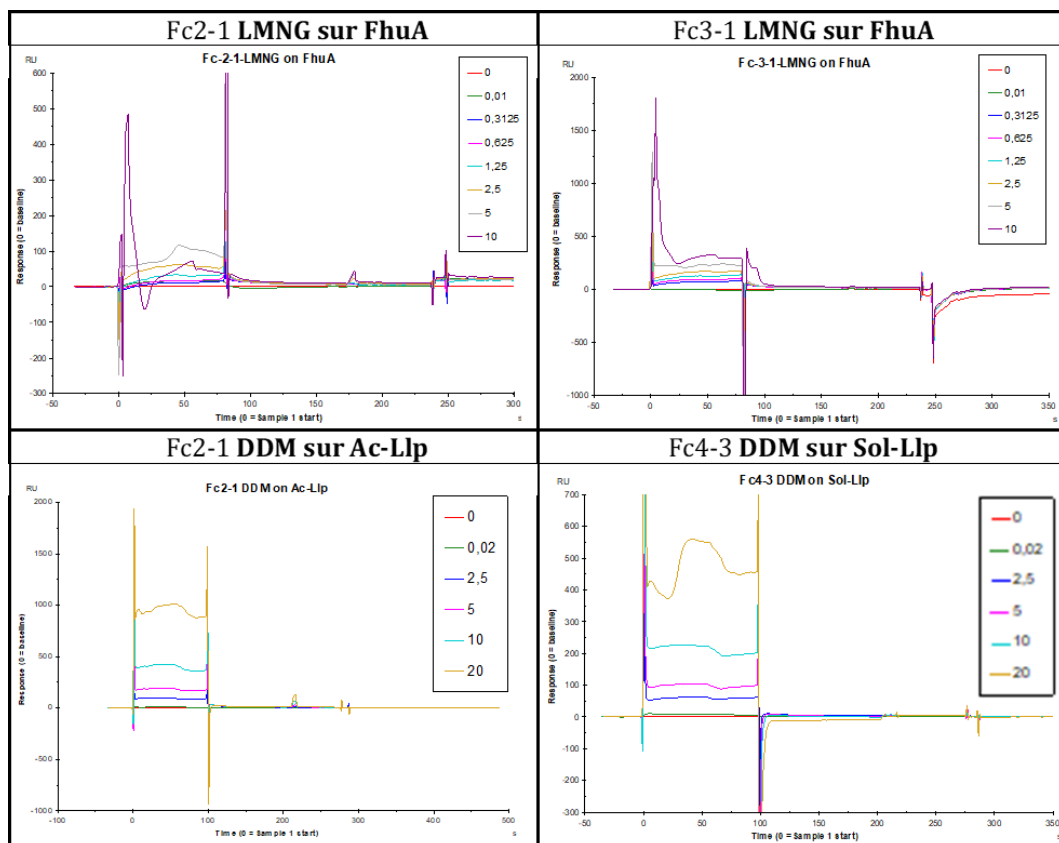
Nombre de phages initial 15000. Mélange de T5 + FhuA avec ou sans Ac-Llp à différent ratio molaire. De gauche à droite dilution des mélanges T5 complexe FhuA-Llp : témoin sans Ac-Llp, dilution pour le ratio 1/50, dilution pour le ratio 1/90, dilution pour le ratio 1/900.

### 6.4.3 Détermination du Kd de l'interaction entre FhuA et Ac-Llp

#### 6.4.3.1 Résonance de plasmonique de Surface, SPR

L'interaction entre FhuA et Ac-Llp ou Sol-Llp a été étudiée par SPR. La SPR permet de déterminer le Kd, mais aussi la vitesse d'association et de dissociation d'un complexe. Un des partenaires, le ligand, est fixé à une surface, et le second, l'analyte, est passé en flux. Si l'analyte interagit avec le ligand alors l'indice de réflexion à la surface est changé et le signal enregistré en unité de résonance (RU) est proportionnel à la masse d'analyte fixé. FhuA étant une grosse protéine qu'il est difficile de concentrer au-delà 125  $\mu$ M, nous avons choisi de la fixer sur la surface et de faire passer Ac-Llp et Sol-Llp en flux. FhuA a été fixée sur la puce *via* ses lysines (par traitement EDC/NHS) à une densité de 1910 RU sur la piste FC2 et de 3400 sur la piste FC4 de la puce. Sachant que 1000 RU équivalent à 1 ng/mm<sup>2</sup> soit  $\sim$ 10 mg/mL, FhuA a donc été fixé à une concentration de  $\sim$ 240  $\mu$ M et 425  $\mu$ M respectivement. La technique de SPR est très sensible aux différences entre le tampon de course et le tampon de l'analyte. La concentration en détergent étant difficile à contrôler et quantifier dans les échantillons, l'effet du

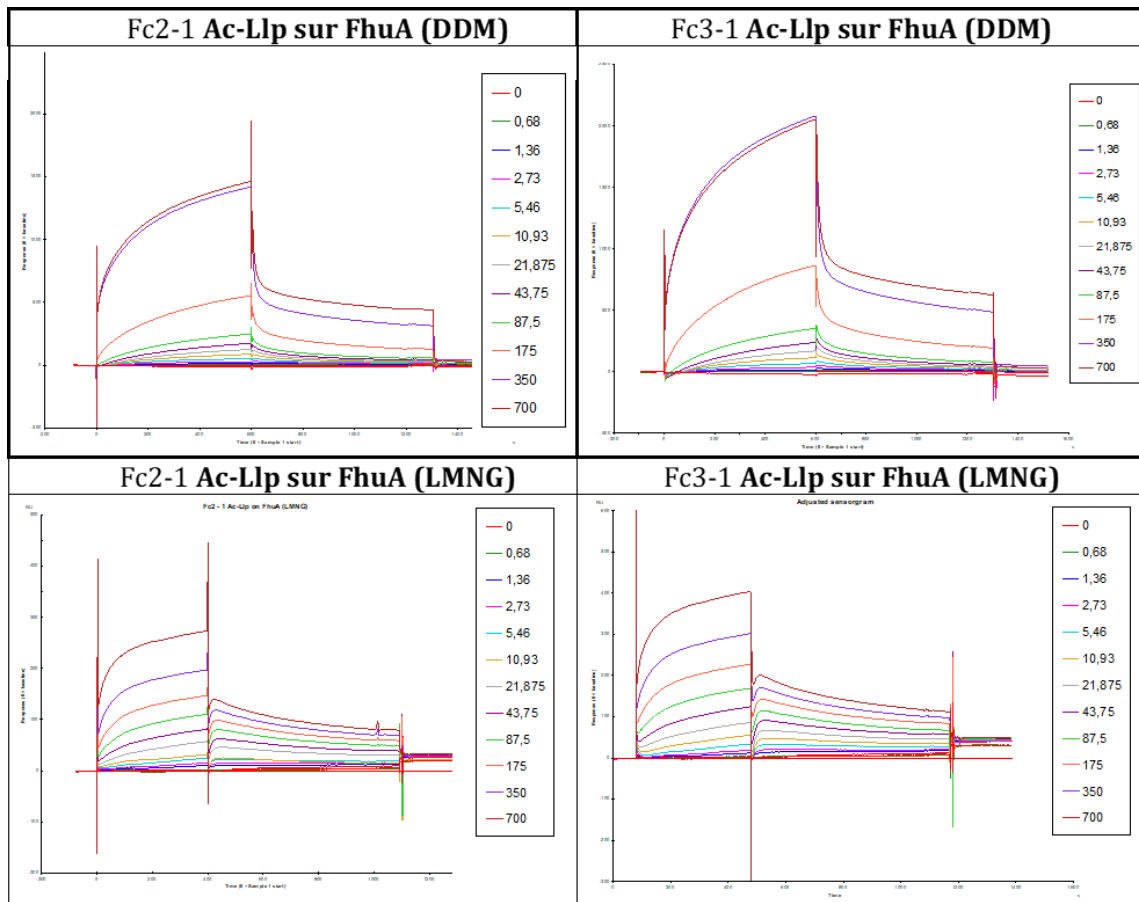
détergent a été testé sur la surface. Les deux détergents utilisés (DDM et LMNG) ont montré qu'ils affectent le signal de SPR en provoquant des sauts importants du signal, en affectant probablement l'épaisseur à la surface. Néanmoins, il semble que l'intensité du signal augmente globalement de façon proportionnelle à la quantité de détergent (Figure 6-26). Le LMNG a moins d'impact sur le signal que le DDM, mais son effet est plus irrégulier, notamment après l'arrêt de l'injection correspondant à la phase de dissociation. Une équation de régression linéaire a été obtenue pour les gammes de concentration en détergent en reportant les valeurs à l'équilibre en RU des pistes présentant les ligands en fonction des valeurs de la piste référence. À partir de cette courbe, les valeurs obtenues avec les gammes d'analytes (Llp ou FhuA) ont subi une correction permettant de retrancher la contribution liée au détergent sur le signal. L'allure des courbes en association et en dissociation ont aussi été examinées pour conclure sur l'interaction entre Llp et FhuA.



**Figure 6-26 : Interaction du LMNG ou du DDM avec FhuA, Ac-Llp ou Sol-Llp immobilisées.**

Chaque sensorgramme représente la réponse enregistrée sur la piste 2 (contenant la protéine d'intérêt) à laquelle a été soustraite la réponse de la piste 1 (greffée avec un témoin négatif, la BSA). Pour chacun des détergents, les concentrations, indiquées en % dans le panneau, sont données en % masse/volume.

Lorsque Ac-Llp en DDM est passée sur la piste de FhuA, les courbes présentent des sauts en début et fin d'injection dûs aux variations de tampon, mais une association et une dissociation progressives et proportionnelles à la concentration de Ac-Llp sont mesurables. En revanche, malgré un temps d'association relativement long de 600 secondes, le signal n'atteint pas de plateau (témoin d'un d'équilibre), mais une fin de courbe linéaire probablement due à des phénomènes d'agrégation de Ac-Llp. De plus pour les concentrations de Ac-Llp les plus élevées (175  $\mu\text{M}$  à 700  $\mu\text{M}$ ), la réponse n'est plus proportionnelle à la concentration contrairement aux plus faibles concentrations (0,68 à 87,5  $\mu\text{M}$ ), ce qui pourrait confirmer ce phénomène (Figure 6-27). L'agrégation serait liée à la présence du détergent qui pourrait participer à l'interaction en stabilisant l'interaction FhuA-Ac-Llp par des mécanismes de fusion de micelles. Cependant, cet effet semble dépendant du détergent utilisé. En effet quand l'interaction de FhuA et Ac-Llp est testée en LMNG, cet effet d'agrégation ne s'observe qu'à partir de 700  $\mu\text{M}$  (figure 6-27). Malgré les effets de tampons, un  $K_d$  approximatif est estimé (voir ci-dessous) pour guider le choix des gammes de concentration de protéines pour d'autres techniques de détermination du  $K_d$ . Pour la protéine soluble, aucune interaction n'a pu être mesurée avec FhuA et ce malgré des concentrations très élevées de Sol-Llp (2 mM).



**Figure 6-27 : Interaction de Ac-Llp avec FhuA immobilisée, en DDM ou en LMNG.**

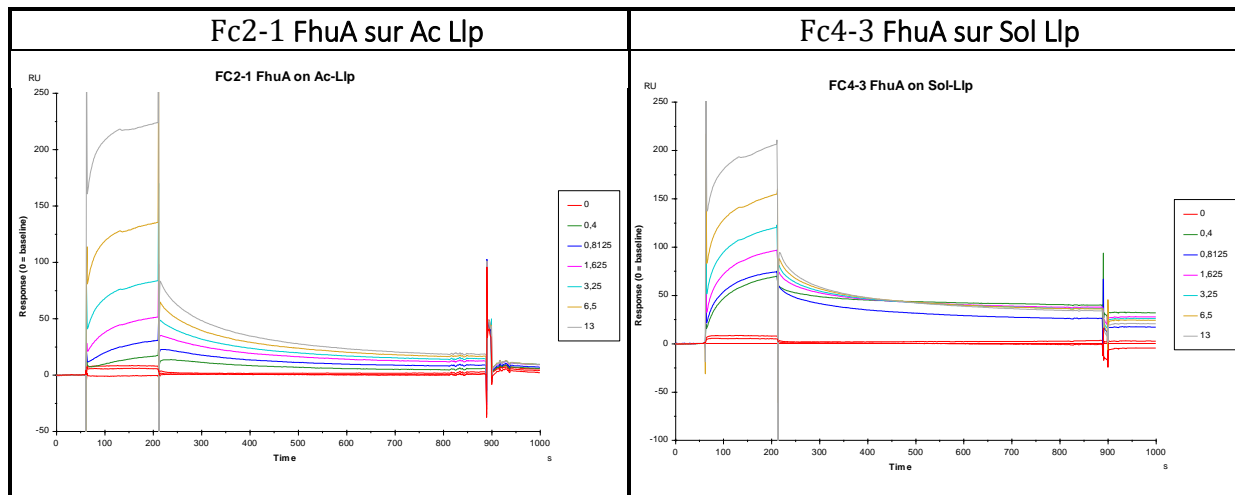
Chaque sensorgramme représente la réponse enregistrée sur la piste 2 (contenant FhuA) à laquelle a été soustraite la réponse de la piste 1 (greffée avec un témoin négatif, la BSA). Pour chaque condition, les gammes de concentrations, sont indiquées en  $\mu\text{M}$  dans le panneau.

Malgré le fait que l'équilibre ne soit pas atteint, l'analyse est tout de même réalisée en « point final » (analyse par le modèle « Steady state affinity »). Grâce aux courbes obtenues en DDM, on peut évaluer l'effet du tampon en traçant la courbe des réponses FC2-FC1 en fonction FC1. Pour la piste 2 on obtient une régression linéaire avec un  $R^2 = 0,9816$  et une équation de droite  $FC2-1 = 0,0256 Fc1 + 12,451$ . Pour la piste 3, la régression linéaire donne un  $R^2 = 0,8750$ , ce qui n'est pas suffisamment fiable pour utiliser l'équation pour corriger les données d'interaction. La valeur corrigée de Fc2-1 des analytes est calculée en appliquant un facteur correcteur calculé grâce à l'intensité de FC1 pour chaque réponse de la gamme à l'aide de l'équation ci-dessus. Ceci permet d'enlever la contribution du DDM aux réponses. Les trois concentrations les plus élevées ont été retirées de l'analyse, car les réponses ne sont plus proportionnelles à l'augmentation de concentration. La nouvelle courbe est alors analysée dans le logiciel kalaidagraphe. Pour Ac-Llp les Kd calculés avec ou sans correction sont très proches (21.3  $\mu\text{M}$  et 22.5 $\mu\text{M}$  respectivement). Pour Ac-Llp en LMNG, le même type de traitement a été appliqué et un Kd de 46,8  $\mu\text{M}$  a été obtenu après correction de l'effet du LMNG. Les expériences n'ont été réalisées qu'une fois au vu des artéfacts induits par la présence de détergent et aucune étude statistique n'a été réalisée. Nous n'avons donc pas conclu quant à l'impact du détergent sur le Kd, mais pensons que les Kd obtenus dans les différents détergents ne sont pas significativement différents et sont du même ordre de grandeur.

Au vu des difficultés rencontrées, nous avons tenté d'estimer le Kd en inversant la fixation de protéine sur la puce. Nous maîtrisons mieux la quantité de micelles libres dans l'échantillon de FhuA du fait du protocole d'échange de détergent par précipitation/ re-solubilisation de la protéine. Nous avons donc fixé sur une nouvelle puce et avec le même principe de fixation la protéine Ac-Llp sur la piste Fc2 à une intensité de 796 RU (soit  $\sim 960 \mu\text{M}$ ) et Sol-Llp sur la piste Fc4 à une intensité de 1196 RU (soit  $\sim 1700 \mu\text{M}$ ). Une gamme de FhuA de 0,4 à 13  $\mu\text{M}$  a été testée. Pour réduire la consommation de FhuA et aux vues des résultats précédents, le suivi d'association n'a été réalisé que sur 200 s. Comme dans l'expérience précédente, de grands sauts de signal ont été observés sur les courbes d'interaction, mais la forme et l'évolution en fonction de la concentration de FhuA nous permet de conclure qu'il y a bien interaction entre FhuA et Ac-Llp. Dans cette configuration FhuA interagit également avec Sol-Llp (Figure 6-28). Le Kd a été calculé comme précédemment pour les deux formes de la protéine et des valeurs de 3,3  $\mu\text{M}$  et 0,5  $\mu\text{M}$  ont été respectivement évaluées pour Ac-Llp et Sol-Llp. Ces Kd plus faibles ne peuvent pas être expliqués par la densité de protéine sur la surface, car le Kd est indépendant des concentrations. En revanche, des études ont montré que l'orientation des protéines sur la surface pouvait jouer un rôle important dans l'interaction par exposition préférentielle de la surface d'interaction. Nous accrochons les protéines de cette étude à la surface à l'aide des lysines libres présentes dans la protéine. Dans le cas de Llp seules 5 lysines sont présentes et notées en vert dans la

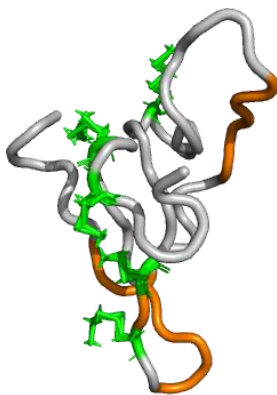


structure (Figure 6-29). De plus, ces lysines ne jouent *a priori* pas de rôle dans l'interaction avec FhuA (voir section 6.4.5) et sont toutes disposées à l'opposé de la surface d'interaction (Figure 6-29). La fixation de Llp sur la puce oriente donc la surface d'interaction déterminée expérimentalement vers le haut. Cette orientation doit faciliter l'interaction avec FhuA et donc diminuer le  $K_d$  apparent par rapport aux expériences précédentes.



**Figure 6-28 : Interaction de Ac-Llp avec FhuA immobilisée, en DDM.**

Chaque sensorgramme représente la réponse enregistrée sur la piste piste 2 (FhuA) à laquelle a été soustraite la réponse de la piste 1 (BSA).



**Figure 6-29 : Structure Sol-Llp.**

En orange la surface d'interaction déterminée expérimentalement et en vert la position des lysines dans la structure.

L'ensemble des analyses de SPR permet d'avoir une première approximation du  $K_d$ , mais sans analyse statistique et avec de gros biais de signal dûs à la présence du détergent. Encore une fois, les résultats suggèrent que les acylations jouent un rôle important dans la liaison entre FhuA et Llp. En effet, en présence des acylations, une interaction a été détectée quelle que soit la protéine fixée sur la surface. Pour Sol-Llp l'interaction n'est mesurable que lorsque Sol-Llp est fixée sur la surface, ce qui peut s'expliquer par une meilleure présentation de la surface d'interaction. L'acylation a donc probablement pour rôle de maintenir Ac-Llp dans une orientation propice à l'interaction avec FhuA au sein du feuillet interne de la membrane externe.

#### 6.4.3.2 Micro Scale Thermophoresis MST

La MST permet de mesurer les interactions entre 2 protéines, dont une marquée en fluorescence, en suivant le mouvement d'un des deux partenaires dans un gradient de chaleur. Le mouvement va être affecté par la taille de la particule, mais aussi son niveau d'hydratation. Le pouvoir MST allant de 20% à 80% permet d'optimiser le signal sur bruit des mesures. Pour marquer un des deux partenaires, plusieurs techniques existent, et nous avons choisi de marquer FhuA à l'aide de son étiquette histidine. En effet, l'étiquette histidine est insérée dans une boucle et ne peut être clivée. Sol-Llp ne possède pas d'étiquette histidine et a pu être testée directement. En revanche, Ac-Llp a une étiquette histidine que nous avons dû cliver (nous avons dû insérer un site de coupure à la TEV entre la séquence de Llp et l'étiquette histidine dans le plasmide pET20b-Ac-Llp), afin que ce partenaire ne se retrouve pas marqué à son tour. D'autre part, des tests préliminaires ont montré que la concentration de détergent pouvait influencer les résultats de MST. Nous avons donc optimisé la préparation d'Ac-Llp pour obtenir une concentration en sortie de colonne suffisante, sans avoir à concentrer la protéine (les micelles sont alors également concentrées). Ainsi, un témoin contenant exactement la même quantité de détergent libre a pu être réalisé (le tampon d'éluion de la Ac-Llp).

Dans un premier temps, l'affinité du marqueur pour l'étiquette histidine a été évaluée. La concentration du fluorophore « RED-tris-NTA dye » a été fixée à 25 nM puis des concentrations de FhuA de 0.122 à 200  $\mu$ M ont été testées selon les recommandations du kit de marquage. Le signal de MST n'a pas été probant pour cette expérience. En revanche, l'intensité de fluorescence initiale a changé de façon proportionnelle à la concentration. Le  $K_d$  du fluorophore a été calculé à partir de ce changement de fluorescence initial et a donné une affinité de 402 nM. Ce  $K_d$  est suffisant pour marquer la protéine et les recommandations de marquage ont été suivies pour cette affinité.

L'affinité de Sol-Llp pour FhuA a été testée avec une gamme de concentration allant 1,6 mM à 97,7 nM de façon indépendante avec un pouvoir de MST de 20% ou de 40%. Le  $K_d$  calculé est de 0.8 +/- 0.2 mM et 1.5 +/- 0.3 mM. Ces deux affinités ne sont pas significativement différentes (Figure 6-30). Le  $K_d$  est beaucoup plus élevé que celui défini en SPR, mais ici nous sommes dans un système à 3 dimensions où la surface n'est pas orientée. Cette expérience est la deuxième preuve d'interaction entre FhuA et Sol-Llp.

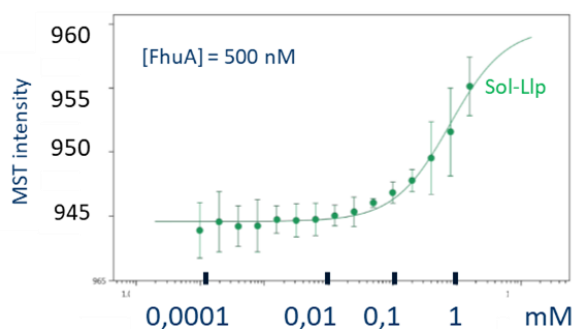


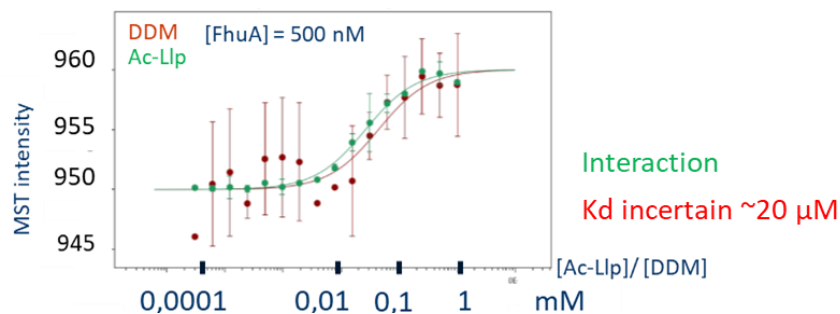
Figure 6-30 : Courbe d'affinité de Sol-Llp pour FhuA. Ajustement de la fluorescence initiale par le logiciel permettant de calculer le  $K_d$ .

Interaction

$K_d = 1.5$  mM

[Sol-Llp]

Pour Ac-Llp, la gamme de concentration testée a été de 0.30  $\mu\text{M}$  à 969  $\mu\text{M}$ . La courbe de liaison montre une interaction avec un  $K_d$  à 26.1  $\mu\text{M}$  +/- 3.3  $\mu\text{M}$ . Cependant, la concentration de DDM a également été testée pour des concentrations de 3.9  $\mu\text{M}$  à 144.4 mM. Cette gamme a été choisie en tenant compte de l'ensemble du détergent présent dans l'échantillon, sachant que Ac-Llp lie 4.1 g de détergent par g de protéine. La courbe du détergent montre un signal relativement proche de la courbe de la protéine. On peut tout de même noter que l'erreur de mesure est beaucoup plus grande pour la courbe avec le DDM seul et donne un  $K_d$  de 43  $\mu\text{M}$   $\pm$  28 $\mu\text{M}$  alors que la courbe avec Ac-Llp donne un  $K_d$  de 26  $\mu\text{M}$   $\pm$  1  $\mu\text{M}$  (Figure 6-31). Cette technique ne permet cependant pas de conclure sur le  $K_d$  de Ac-Llp.

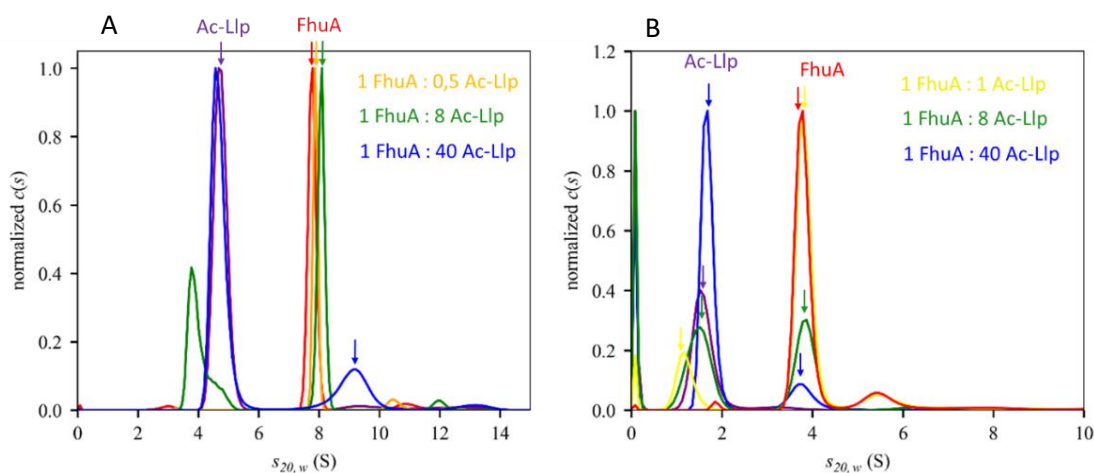


**Figure 6-31 : Courbes d'affinité de Ac-Llp et du DDM pour FhuA.**

Ajustement de la fluorescence initiale par le logiciel permettant de calculer le  $K_d$ . Courbe verte : Ac-Llp, courbe rouge : DDM.

#### 6.4.3.3 AUC

L'AUC permet également d'explorer le  $K_d$  des complexes ainsi que leur stœchiométrie. Si le complexe se forme entre FhuA et Llp, on s'attend à ce que le coefficient de sédimentation de FhuA augmente. Quand on mélange FhuA avec des concentrations croissantes de Sol-Llp, le coefficient de sédimentation de FhuA reste constant, ce qui indique que le complexe ne se forme pas aux concentrations testées (entre 5  $\mu\text{M}$  et 400  $\mu\text{M}$  de Sol-Llp pour 10  $\mu\text{M}$  de FhuA) (Figure 6-32 B). En revanche, en présence de Ac-Llp aux mêmes concentrations, les résultats montrent une augmentation du coefficient de sédimentation de FhuA (Figure 6-32 A). Cependant, la gamme de concentrations était trop faible pour obtenir un  $K_d$ . Pour compléter ces résultats, une nouvelle série d'expériences a été menée, mais les résultats n'ont pas été reproduits. Ceci est probablement dû à la protéine Ac-Llp. En effet, la purification de cette protéine a été modifiée plusieurs fois et toutes les expériences d'AUC ont été réalisées avec des purifications différentes. Les échantillons de la dernière analyse d'AUC, ne donnant pas de résultats exploitables, ont été passés en SEC-MALS : les résultats ont montré que le complexe était bien formé. Une étude récente<sup>107</sup>, suggère que la formation du complexe FhuA-Llp peut être assez longue (48H). Nous avons donc programmé de vérifier l'influence du temps sur la formation du complexe et le signal AUC, car les complexes utilisés ont été formés à différents temps avant les expériences d'AUC.



**Figure 6-32 : Coefficient de sédimentation de Ac-Llp, Sol-Llp, FhuA et de complexe FhuA : Llp à différents ratios molaires.**  
 A. Echantillons de FhuA solubilisés en DDM. B Echantillons de FhuA solubilisés en LDAO.

#### 6.4.3.4 Autres méthodes testées

La technique de BLI (Interférométrie de bicouches) permet également de détecter des interactions entre protéines. Comme la SPR elle permet d'enregistrer des sensorgrammes sur des gammes de concentration des deux partenaires et de calculer des  $K_d$ . Cependant, dans ce système, le flux de protéine n'est pas assuré par de la microfluidique, mais par le mouvement du liquide dans des puits sous agitation. Nous avons donc testé cette technique, car le système de détection et de flux étant différent nous aurions pu obtenir de meilleurs résultats. Comme dans le cas de la SPR nous avons fixé FhuA sur les puces puis incubé avec Ac-Llp. Nous notons une interaction entre FhuA et Ac-Llp, mais avec des sauts de signaux pour les plus fortes concentrations (voir annexe 4). Le  $K_d$  n'a pas été calculé. Dans le cas de Sol-Llp aucune interaction n'a pu être détectée avec FhuA (non montré).

La méthode Fidabio est basée sur l'analyse de la dispersion des protéines dans un capillaire soumis à un flux. Le profil de diffusion hydrodynamique est analysé pour définir la taille des particules. Selon ce principe, la formation du complexe doit être visible par l'augmentation de la taille des particules. Le complexe FhuA-détergent apparaît comme une particule de Rh (rayon hydrodynamique) de 4,41 nm, ce qui est cohérent avec les données d'AUC et de SEC-MALS. En revanche, aucune différence de Rh significative n'a pu être observée pour une gamme de Ac-Llp de 0,1 à 50  $\mu$ M. Ac-Llp est une petite protéine, et la formation du complexe peut engendrer une réorganisation de la micelle autour de FhuA, ce qui pourrait expliquer qu'on ne voit pas de changement de taille au cours de la titration. Pour tester notre hypothèse, nous avons marqué Ac-Llp, malheureusement nous n'avons pas réussi à retirer le fluorophore libre, ce qui n'a pas permis d'enregistrer une courbe de titration (voir annexe 4).

#### 6.4.3.5 Synthèse des différentes méthodes utilisées :

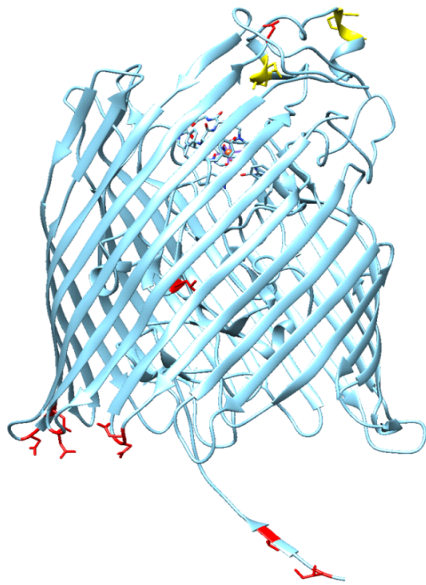
Comme la montré les résultats chaque méthode biophysique utilisé pour la détermination du kd a montré des limites dans l'interprétation des résultats. Cela est due au fait que nous étudions un système complexe ou interviennent plusieurs partenaires dont des micelles de détergent qui sont souvent responsable de fort bruit de fond dans les systèmes d'analyse biophysique. Mais nous avons aussi un système qui pourrait s'équilibrer en plusieurs jours ce qui peut aussi fausser les résultats. Nous n'avons donc pas pu obtenir un Kd fiable pour Ac-Llp, mais l'ensemble des résultats converge vers un Kd de l'ordre de 0,5 à 80  $\mu$ M voir tableau récapitulatif ci dessous.

Méthode utilisé	Ac-Llp		Sol-Llp	
	Détection d'interaction	kd	Détection d'interaction	kd
AUC	Oui	0,5 à 50 $\mu$ M	Non	-
SPR	Oui	0,5 à 50 $\mu$ M	Non	-
MST	Oui	20 $\mu$ M	Oui	1,5 mM
Fidabio	Oui	Non calculé	NA	NA
BLI	Oui	20 à 80 $\mu$ M	Non	-
RMN	NA	NA	Oui	Non calculé / compatible avec 1,5 mM

#### 6.4.4 Exploration de la surface d'interaction côté FhuA

##### 6.4.4.1 Mutants de FhuA

Des mutations dans FhuA empêchent la protection par Llp<sup>66</sup>. Braun et al. décrivent des mutants distribués du côté périplasmique, dans le bouchon et du côté extracellulaire. Nous avons cherché à savoir si les mutants étaient impliqués dans la liaison à Llp ou dans la transmission de l'information entre le périplasme et le domaine extracellulaire. La majorité des mutants décrits dans l'étude de 1994 sont des mutations spontanées ou des délétions de 4 ou 16 acides aminés. Les mutations ponctuelles isolées dans le papier de Braun et al. ont été reprises et complétées par une série de nouveaux mutants ponctuels et de petites délétions de 3 acides aminés, répartis dans les boucles extracellulaires et dans le domaine périplasmique de FhuA (Figure 6-33)

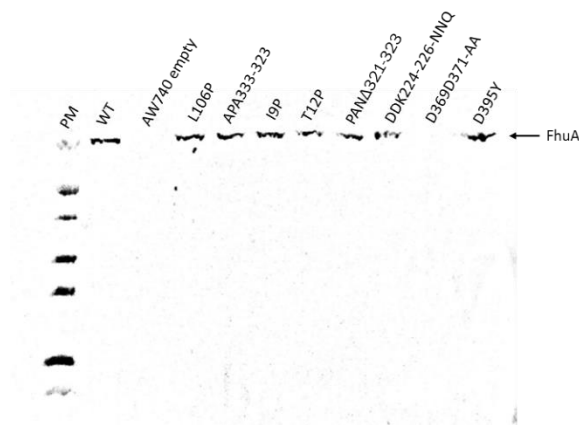


Périplasmique	Extracellulaire	Bouchon
<ul style="list-style-type: none"> <li>• T12P</li> <li>• D292A</li> <li>• D369D371-AA</li> <li>• DDK224-226-NNQ</li> <li>• I9P*</li> </ul>	<ul style="list-style-type: none"> <li>• PANΔ321-323</li> <li>• APAΔ333-335</li> <li>• D395Y*</li> </ul>	<ul style="list-style-type: none"> <li>• L106P*</li> </ul>

**Figure 6-33 : Localisation des mutants de FhuA étudiés.**

En rouge les mutations ponctuelles, en jaune les délétions, les mutants repris de l'étude 64 sont marqués par un astérisque.

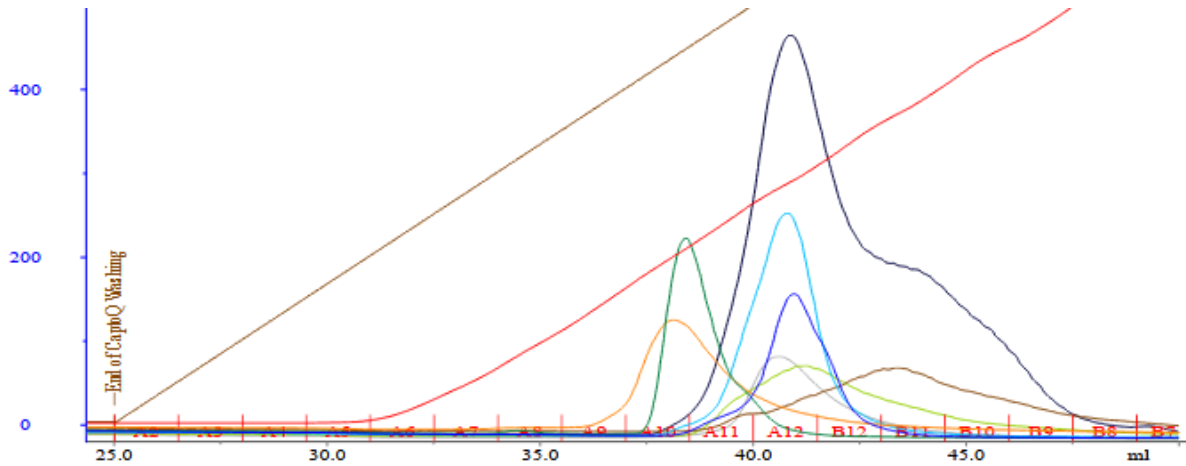
L'impact des mutations dans FhuA sur la capacité de Llp à protéger les cellules est testé *in vivo*. La souche AW740 est délétée de FhuA endogène, ce qui permet de l'utiliser pour tester l'activité des mutants de FhuA. La souche AW740 a été transformée avec les plasmides codant pour les mutants de FhuA (pH405-His-FhuA avec les différentes mutations inductibles en absence de fer). Comme attendu, la souche AW740 vide est insensible au phage T5 (non montré). La fonctionnalité des mutants de FhuA a été contrôlée par sensibilité au phage T5 de la souche AW740 transformée avec les plasmides contenant les mutations. Lorsque cultivés en LB, à l'exception du mutant D369D371-AA, tous les mutants sont sensibles à T5, et l'expression des protéines a pu être détectée par immuno-révélation sur migration sur gel de cellules entières (Figure 6-34). Cependant, en présence de bipyridyl, le mutant D369D371-AA a permis l'infection de la souche AW740 (non montré) montrant que ce mutant est peu exprimé.



**Figure 6-34 : Vérification de l'expression des gènes de FhuA mutés dans la souche AW740 en absence de bipyridyl.**

Afin de tester la protection de la souche à l'infection par T5, une co transformation du plasmide pH405-His des différents mutants et le plasmide codant pour Ac-Llp (plasmide pASK-IBA3C inducible à l'anhydrotetracycline (AHT) a été réalisée. Pour s'assurer de l'expression de FhuA et Ac-Llp, les tests de protection ont été réalisés en présence de bipyridyl et de AHT. La souche AW740 exprimant FhuA WT n'est que très partiellement protégée par Ac-Llp. Deux hypothèses permettent d'expliquer ce résultat. Il est possible que le nombre de copies des plasmides codant pour FhuA et Llp soit différent, que les niveaux d'expression des protéines soient ainsi inégaux et qu'il n'y ait donc pas assez de Llp pour protéger les cellules. Néanmoins, Ac-Llp réduit quand même l'infection par T5 d'environ  $10^2$  (le titre de T5 de  $3 \times 10^{13}$  est réduit à  $3 \times 10^{11}$ ), ce qui montre une interaction efficace entre FhuA et Llp *in vivo*. Par ailleurs, Arguijo-Hernandez *et al.* 2018, montrent que les protéines Cor ne peuvent pas protéger les cellules qui surexpriment FhuA mais qui n'expriment pas certaines OMPs. Hors, la souche AW740 est déletée pour les principales OMPs. Cependant, nous n'avons aucune indication pour proposer un rôle des OMPs dans l'infection de T5. Concernant les mutants de FhuA, seuls I9P (périplasmique) et D995Y (extracellulaire) ne montrent pas de protection contre le phage T5 en présence de Ac-Llp (Tableau 6-2). La faible réponse de Ac-Llp sur FhuA WT ne permet pas de conclure sur les protections partielles : le mutant L106P pourrait être légèrement protégé de  $10^1$ , comme dans Braun *et al.* 1994, où il est noté comme levant partiellement la protection conférée par Llp (tableau 6-2). Afin de caractériser plus avant ces mutations, nous avons donc purifié les protéines mutantes et effectué une analyse en AUC et en nano-DSF (Figure 6-35 et Tableau 6-2).





**Figure 6-35 : Comparaison des profils d'éluion des mutants de FhuA sur NiNTA, purifiés sur la plateforme M3P de l'IBS.**  
 Bleu foncé FhuA WT, vert clair T12P, vert foncé APAΔ333-335, orange DDK224-226 – NNQ, bleu clair PANΔ321-323, marron L106P, gris D995Y, bleu marine I9P.

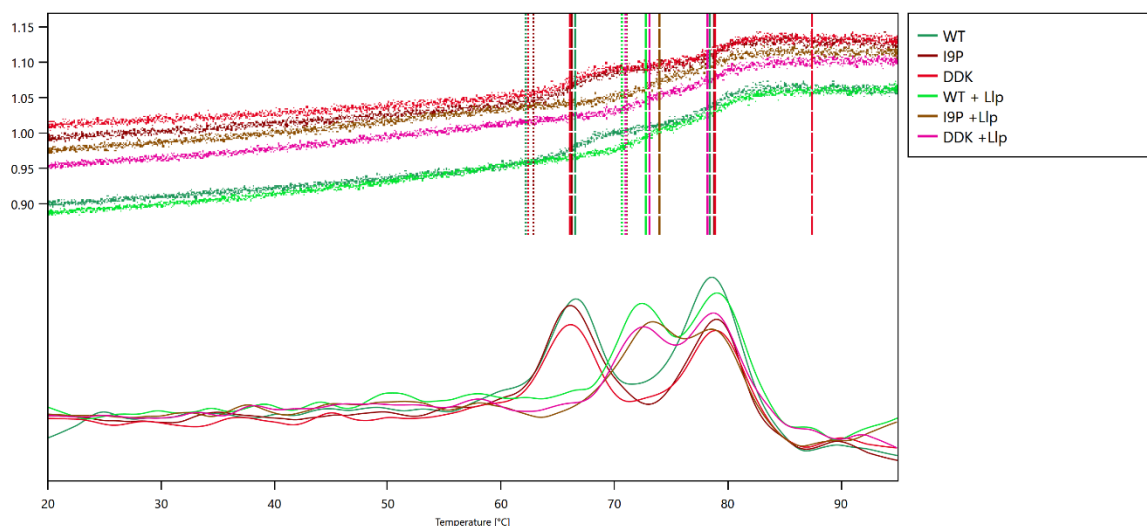
Mutant de FhuA	Clonage	Expression	Purification	Rendement de purification en mg/L	TM en °C	Protection
L106P	+	+	+	1.45	61.3/73	+/-
I9P	+	+	+	1.34	66.2/78.7	-
D995Y	+	+	+	0.92	~60	-
T12P	+	+	+	1.2	66.1/78	+
D292A	-	N.A	N.A	N.A	N.A	N.A
D369D371-AA	+	+ si bipyridyl	-	N.A	N.A	N.A
DDK224-226 - NNQ	+	+	+	2.4	66.1/78.9	+
PANΔ321-323	+	+	+	2.94	59,1/78,6	+
APAΔ333-335	+	+	+	1.14	58.2/79.1	+
WT	Ref	Ref	Ref	6.66	66,5/78,5	+

**Tableau 6-2 : Récapitulatif des données des mutants de FhuA.**

Le nom de la mutation donne l'acide aminé initial, sa position puis l'acide aminé de remplacement. Les mutations en bleu ont déjà été décrites, les mutations en noir ont été réalisées pour cette étude. Les valeurs en rouge soulignent des différences importantes par rapport au résultat de FhuA wt.

La stabilité de tous les mutants de FhuA purifiés a été contrôlée par nano-DSF. Les mutants I9P, T12P, DDK224-226-NNQ présentent un profil de dénaturation identique à FhuA WT comprenant deux évènements de dénaturation. Le mutant L106P montre le même profil mais décalé vers les basses températures, 61,3°C et 73°C respectivement (Tableau 6-2). Pour les mutants PAN $\Delta$ 321-323 et APA $\Delta$ 333-335, la dénaturation du bouchon est plus basse (~58°C), mais la dénaturation du tonneau bêta est inchangée (Figure 6-36 et Tableau 6-2). Le mutant D995Y a un fort impact sur la stabilité de FhuA, car il présente un seul évènement de dénaturation à une température de 60°C.

Les différences de rendement de purification observé ne sont pas directement liées à la stabilité de FhuA sauf pour le mutant D995Y, qui montre à la fois le rendement et la stabilité les plus faibles. Tous les mutants ont pu être purifiés, sauf D369D371-AA, en raison de son faible niveau d'expression et probablement d'une instabilité (Tableau 6-2).



**Figure 6-36 : Courbe de suivi de dénaturation de FhuA mutant et des complexes FhuA : Ac-Llp.**

*Abscisse : température en °C, ordonnées du panneau du haut : ratio de fluorescence 350nm/330nm. Panneau du bas : première dérivée de fluorescence, sans unité.*

Les complexes FhuA-Llp de l'ensemble des mutants de FhuA purifiés ont été analysés par nano-DSF. Comme pour FhuA WT, la température de dénaturation du bouchon est plus élevée en présence de Ac-Llp pour les mutants I9P, T12P, DDK224-226-NNQ et L106P, suggérant qu'il y a formation du complexe FhuA-Ac-Llp (Figure 6-36). Pour les mutants PAN $\Delta$ 321-323 et APA $\Delta$ 333-335, la stabilité du bouchon semble affectée mais est moins claire que pour les autres mutants. Enfin, pour le mutant D395Y, l'absence du premier Tm ne permet pas de conclure sur l'association. L'ensemble des mutants semble interagir avec FhuA et stabiliser le bouchon y compris les mutants ne permettant pas la protection par Llp (I9P, D395Y et peut-être L160P – Braun et al). Il semble donc que ces trois acides aminés ne soient pas directement impliqués dans l'interaction avec Llp mais sont en revanche

importants pour transmettre la résistance le long de FhuA pour bloquer l'interaction avec pb5. Les expériences d'AUC ont permis de valider l'interaction pour certain mutant voir tableau récapitulatif ci-dessous.

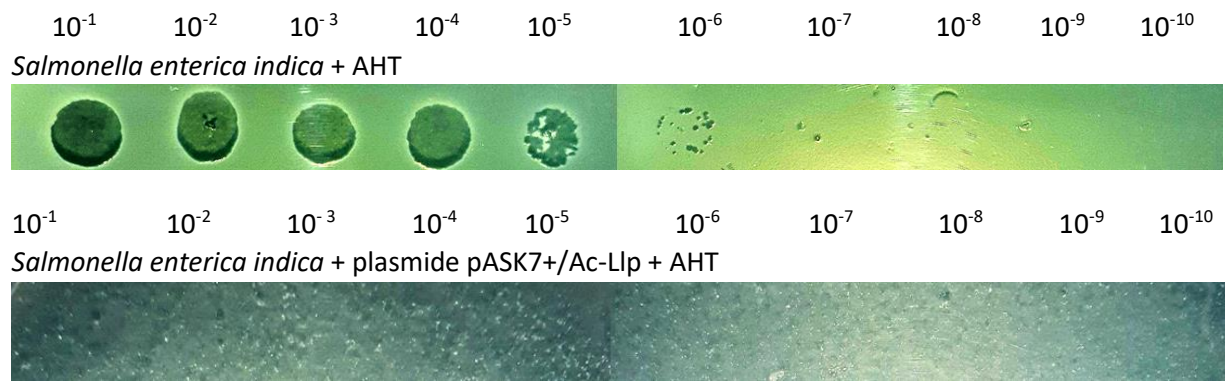
Mutant de FhuA	interaction	Protection par Llp
WT	Oui (Nano-DSF + AUC)	Oui
I9P	Oui (Nano-DSF + AUC)	Non
T12P	Oui (Nano-DSF) / Kd plus faible ? (AUC)	Oui
DDK	Oui (Nano-DSF) / Kd plus faible ? (AUC)	Oui
L106P	Incertain (Nano-DSF + AUC)	Partielle
PAN	Oui (AUC) / Incertain (Nano-DSF )	Oui
APA	Incertain (Nano-DSF + AUC)	Oui
D395Y	Oui / Incertain (Nano-DSF )	Non

#### 6.4.4.2 Variations naturelles de FhuA dans des souches de *Shigella* et de *Salmonella*

Les FhuA des souches *Shigella sonnei*, *Shigella flexneri*, *Salmonella enterica indica*, et *Salmonella enterica bongori* présentent des variations d'acides aminés par rapport à la souche de *E. coli* (supplementary figure 2 section 4.3). Les souches ont été testées pour la sensibilité au bactériophage T5 dans le but d'étudier l'impact des variations d'acides aminés sur la liaison de pb5, et seule *Salmonella enterica indica* a été sensible (section 4.Chapitre 1).

Dans le même esprit, la protection de *Salmonella enterica indica* par Ac-Llp a été contrôlée. *Salmonella enterica indica* a été rendue chimio-compétant et a été transformée par choc thermique par le plasmide pASK7+ contenant le gène Ac-Llp.

La présence de l'inducteur, AHT n'a pas d'incidence sur la croissance de *Salmonella enterica indica* ni sur la capacité du bactériophage T5 à l'infecter. En revanche, la surexpression de Ac-Llp affecte la croissance de *Salmonella enterica indica*. Néanmoins Ac-Llp est capable de protéger complètement cette souche contre l'infection du phage T5 (Figure 6-37). Ce résultat n'est pas surprenant car les variations de séquence sont présentes dans les boucles extracellulaires alors que le tonneau beta, le bouchon et les boucles périsplasmiques sont conservés (supplementary figure 2 section 4.3). La liaison de Ac-Llp du côté périsplasmique et sur le bouchon est préservée. Ce résultat indique également que certaines variations de séquence dans les boucles extracellulaires n'empêchent pas la protection par Ac-Llp.



**Figure 6-37 : Titration du phage T5 sur les souches *Salmonella enterica indica* en absence ou en présence de Ac-Llp.**

Les titrations sont effectuées avec le phage T5 purifié à un titre de  $1.4 \times 10^{10}$  phage/ml contrôlé sur la souche F. 5  $\mu$ l de différentes dilutions de phage ont été déposés sur *Salmonella enterica indica* transformée ou non avec le plasmide Ac-Llp, poussant dans de la gélose molle contenant de l'AHT

#### 6.4.5 Exploration de la surface d'interaction coté Llp

##### 6.4.5.1 RMN

Ce travail a été fait en collaboration avec Loïc Salmon et Emeline Mestdach au (CRMN, Lyon).

Les pics présents sur les spectres HSQC d'une protéine marquée au  $^{15}\text{N}$  reflètent l'environnement chimique dans lequel se trouvent les atomes. Une perturbation du déplacement chimique ou de l'intensité des pics traduit un changement d'environnement. Sol-Llp a donc été titrée avec des quantités croissantes de FhuA non marquée et les perturbations de chaque pic de Sol-Llp sont suivies pour identifier les résidus impliqués dans l'interaction. De plus, la perturbation des pics peut permettre de calculer un  $K_d$  qui pourrait confirmer celui obtenu par MST.

Pour commencer, l'effet du détergent sur le signal RMN de Sol-Llp a été examiné. L'observation directe du spectre et l'analyse individuelle de chaque pic montre que globalement le signal de la protéine est légèrement affecté par le détergent (Figure 6-38 et 6-39). Sont affectés par le DDM les acides aminés Y22 à D28 et M16. Les spectres de titration avec FhuA seront donc comparés au spectre avec DDM pour retirer les effets du détergent dans l'analyse de l'interaction.

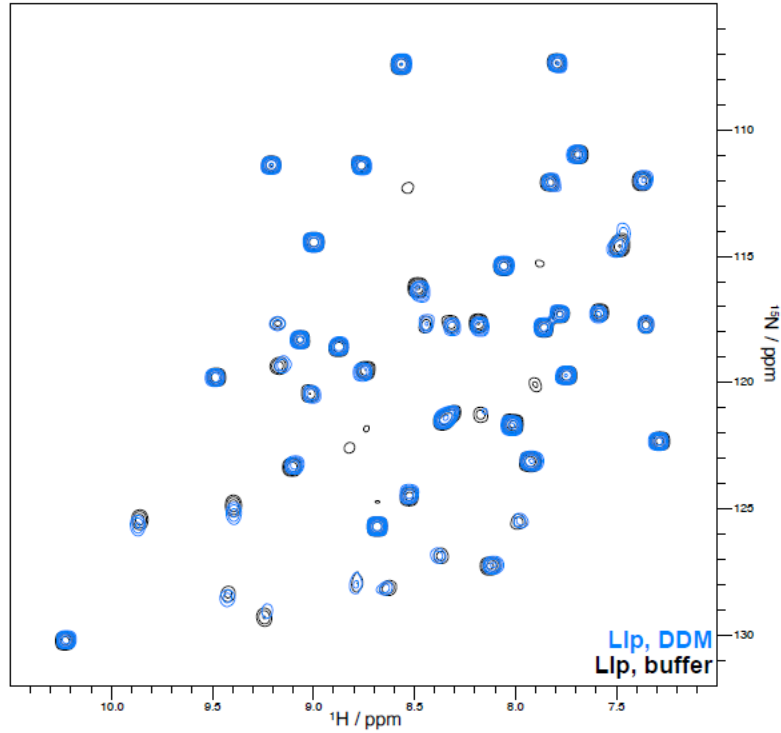


Figure 6-38 :  $^1\text{H},^{15}\text{N}$ -SOFAST-HSQC sur Llp.  
Llp en tampon (noir) ou en présence de 0,4 mM de DDM (bleu)

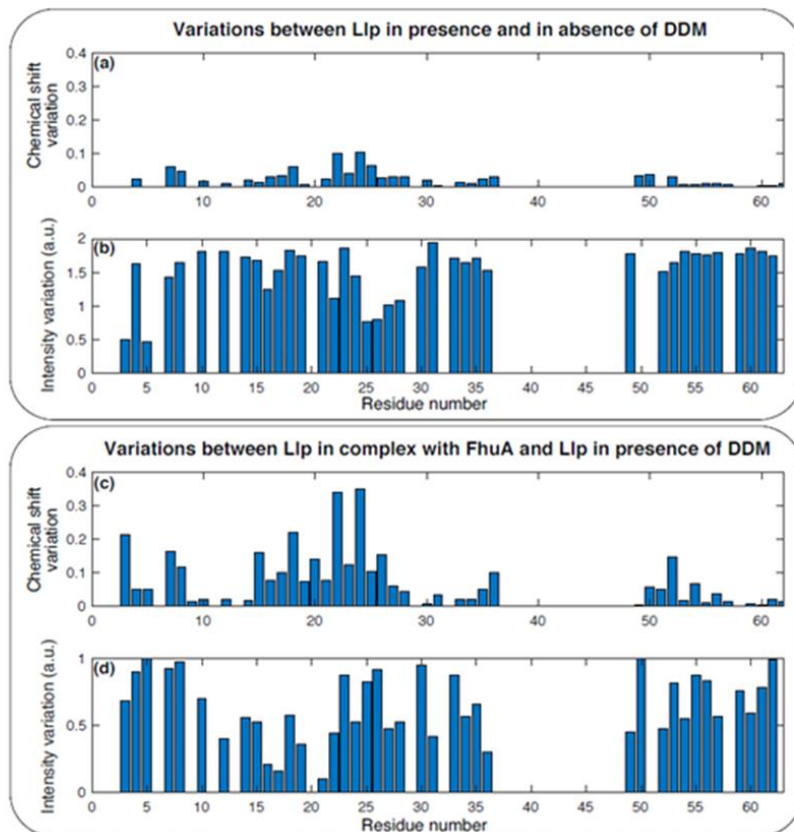


Figure 6-39 : Investigation de l'influence du DDM (a & b) ou de FhuA (c & d) sur Sol-Llp par l'analyse des pics du spectre  $^1\text{H},^{15}\text{N}$ -SOFAST-HSQC.

Variation des déplacements chimiques (a, c) et de l'intensité (b, d) entre les spectres de Llp en présence ou en absence de DDM et en présence ou absence de FhuA solubilisée en DDM.

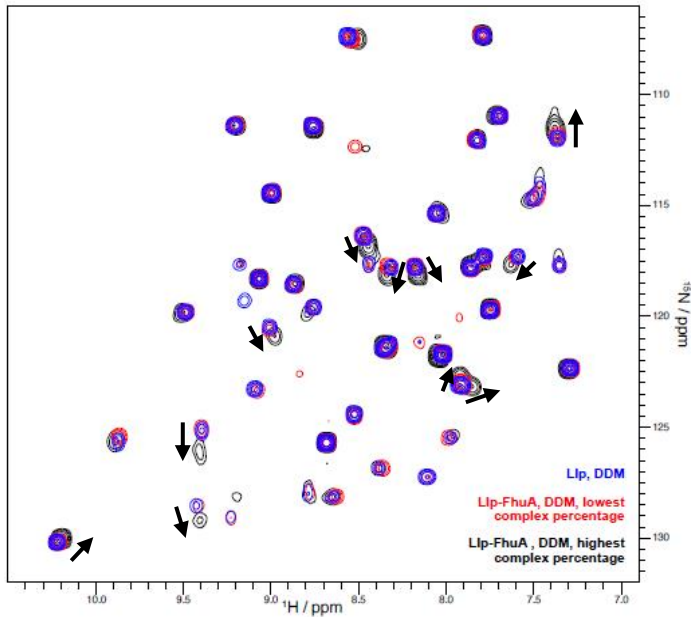
Au cours de la titration, la quantité de FhuA est fixe, et l'ajout de Llp fait augmenter la concentration de Llp et diminuer la concentration de FhuA, ce qui permet d'avoir une plus grande proportion de complexe au cours de la titration (Tableau 6-3). Pour éviter de diluer le détergent, 0,4 mM de DDM ont été ajoutés à Sol-Llp. Pour chaque point de titration, un spectre  $^1\text{H}$ - $^{15}\text{N}$ -SOFAS- $^2\text{D}$ -HSQC a été acquis. Plus le pourcentage de complexe est important, plus l'impact sur le spectre est visible. Malheureusement, les rapports de titration sont restés assez bas (7.4% de complexe au maximum pour un  $K_D$  de 1,5 mM), car FhuA est une grosse protéine membranaire et la concentration utilisée de 150  $\mu\text{M}$  correspond déjà à 12 mg/ml. Les données sont insuffisantes pour déterminer un  $K_D$  précis. En revanche, les variations enregistrées sur les spectres pour l'interaction avec FhuA sont relativement faibles ce qui est cohérent avec un  $K_D$  élevé, de l'ordre du mM, et une faible proportion de complexe final, cohérent avec le  $K_D$  déterminé par MST.

$C_f$ Llp ( $\mu\text{M}$ )	$C_f$ FhuA ( $\mu\text{M}$ )	% de complexe
58	127	0,4
131	120	0,9
207	114	1,5
290	106	2,2
288	75	3,1
350	71	3,9
414	67	4,9
484	63	6,2
545	59	7,4

**Tableau 6-3 : Concentration des différents partenaires pour chaque conditions mesurées en RMN.**  
*Le pourcentage de complexe formé est calculé pour un  $K_D$  de 1,5 mM (valeur obtenue par MST).*

Les variations des déplacements chimiques ont été analysées pour chaque ajout de Llp pendant la titration. En considérant les pics bien isolés et la variation lente et graduelle observable, l'attribution du déplacement chimique a été plutôt simple, sans avoir recours à la réattribution des parties affectées de la protéine (Figure 6-40).

Comme attendu, le spectre global de Sol-Llp (Figure 6-38 et 6-40) subit une diminution d'intensité au contact de FhuA. En effet Llp fait 7.06 kDa, mais FhuA a une masse apparente de 200 kDa (80 kDa de protéine et 120 kDa de détergent). Le mouvement Brownien du complexe est donc plus lent, ce qui réduit la relaxation transversale et élargit les pics sur les spectres HSQC. L'influence du DDM est toujours observable, mais certaines régions présentent une réduction d'intensité et des variations de déplacements chimiques encore plus importantes, comme les résidus E12-Y22 ainsi que les résidus Y51-E54 situés près de la région C-terminale (Figure 6-39 (c) et (d)). Pour rappel, les résidus Y14-A24 forment une boucle désordonnée, disponible pour interagir avec FhuA.



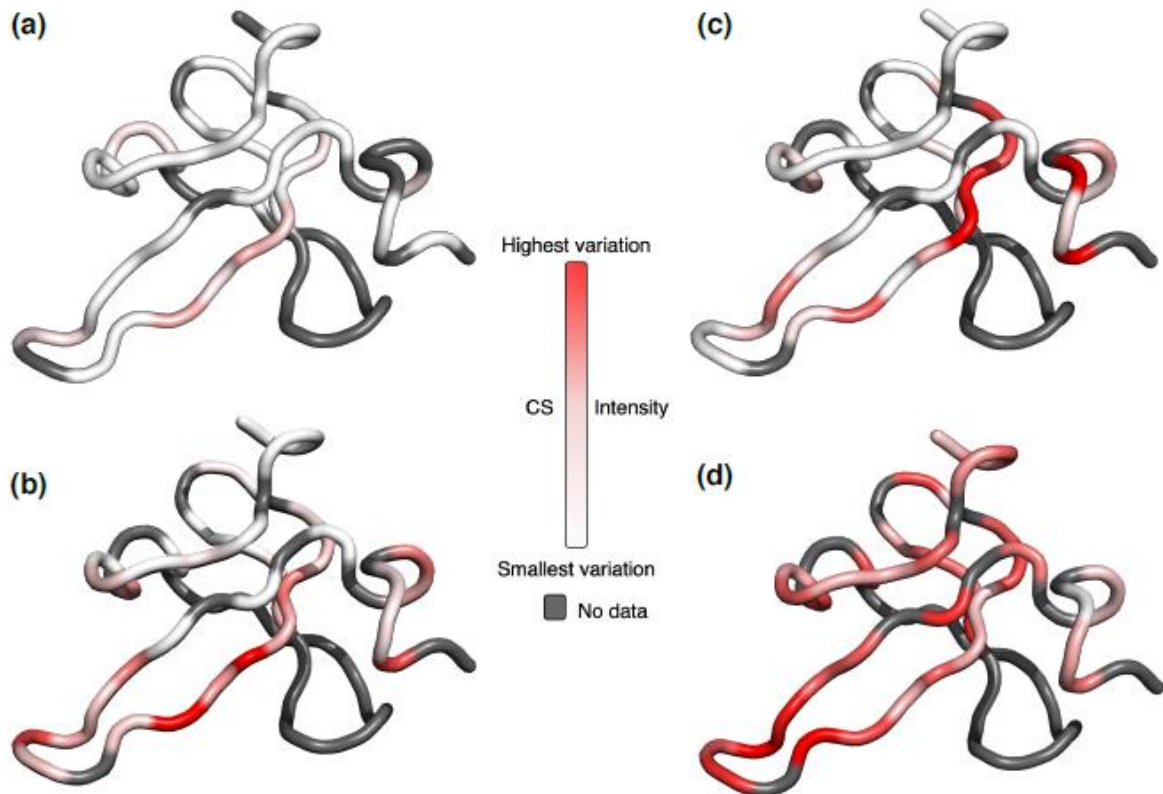
**Figure 6-40: 1H,15N-SOFAST-HSQC de Sol-Llp.**

En présence de 0,4 mM DDM (bleu), 127  $\mu$ M de FhuA (rouge) et 59  $\mu$ M de FhuA (noir) solubilisée en 0,4 mM de DDM. Les Flèches indiquent le sens de déplacement des pics affectés.

Pour résumer, les résidus Y22-D28 ainsi que les résidus T3 et G5 présentent une réduction d'intensité significative lors de l'ajout de DDM à Sol-Llp. Les résidus E12-Y22 subissent une diminution d'intensité observable en comparant les spectres de Sol-Llp dans le DDM et de Sol-Llp interagissant avec FhuA en présence de DDM, diminution qui devient de plus en plus évidente avec l'augmentation de complexe, ce qui suggère que cette variation n'est pas liée uniquement à l'ajout de détergent. Les acides aminés T3 et G5 ne semblent pas ou peu affectés par l'ajout de FhuA. Les acides aminés Y51-E54 semblent être affectés uniquement en présence de FhuA. Dans tous les cas, les variations d'intensité sont plus importantes que les variations de déplacements chimiques sauf pour les résidus Y22 et A24 (Figure 6-41).



A la lumière de notre analyse, il semble que les résidus interagissant avec le DDM et avec FhuA sont proches, mais pas identiques. Compte tenu de leur proximité, l'impact que ces zones pourraient avoir l'une sur l'autre n'est pas clair. Les acides aminés proches du N-terminal (S2-P6) sont plutôt hydrophobes et ils font le lien entre l'ancre acylée et le cœur de la protéine. Il est donc normal qu'ils interagissent avec le DDM. Llp étant une protéine ancrée à la membrane, elle doit interagir avec celle-ci avant sa liaison à FhuA, il se peut que la même région de la protéine puisse interagir avec la membrane puis avec FhuA.



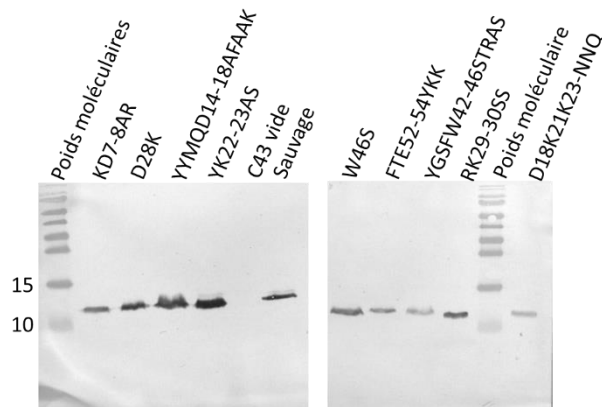
**Figure 6-41 : Variations de déplacements chimiques et d'intensités reportées sur la structure de Llp.**

(a) et (c) Sol-Llp en présence de 0,4 mM DDM ; (b) et d) Sol-Llp en présence de 59 μM FhuA; (a) et (b) variation de déplacements chimiques ; (c) et (d) variation d'intensités.

#### 6.4.5.2 Exploration de la surface d'interaction de Sol-Llp par mutagenèse

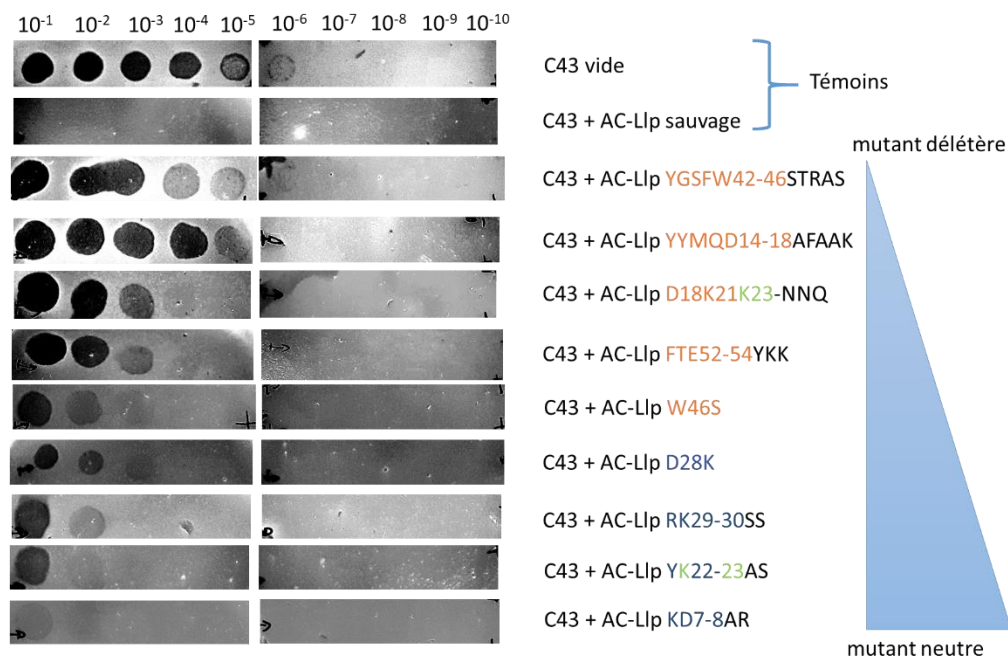
Les données de RMN ont permis de déterminer des zones d'intérêts pour définir une surface d'interaction entre FhuA et Llp. Cependant, d'une part, les enregistrements ont montré que certains acides aminés étaient affectés par le détergent, et d'autre part aucune information n'est disponible pour les acides aminés I39-A47 non attribués. Pour consolider et compléter les résultats, un éventail de mutants Llp a été préparé pour observer l'impact des résidus modifiés sur la protection à T5. Les mutants ont été réalisés par mutagenèse dirigée sur le plasmide de Ac-Llp pour permettre le contrôle *in vivo* de l'interaction entre FhuA et Llp.

La souche C43 a été transformée avec chacun des plasmides portant les mutations de Llp et une immuno-révélation d'un gel de cellules totales a été réalisée pour tous les mutants afin de s'assurer que les mutations n'empêchent pas l'expression de Ac-Llp. Tous les mutants sont exprimés dans la souche C43 avec un niveau d'expression comparable (Figure 6-42).



**Figure 6-42 : Vérification de l'expression des mutants de Ac-Llp dans la souche C43.**  
Immunorévélation avec un anticorps anti poly-histidine de cellules totales migrées sur gel.

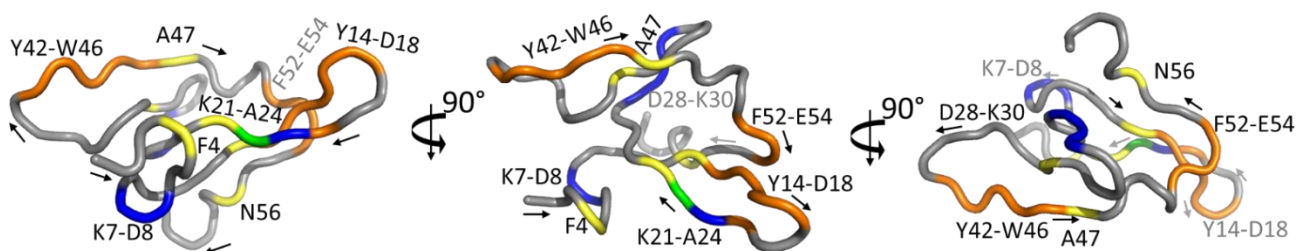
La sensibilité de C43 portant les plasmides de Ac-Llp mutée a été contrôlée de la même façon que pour Ac-Llp WT (voir section 6.4.1). Le mutant KD7-8AR est capable de protéger la souche C43 (Figure 6-43). Les mutants YGSFW42-46STRAS et YYMQD14-18AFAAK ne sont plus capables de protéger la cellule : ils sont donc importants pour la liaison à FhuA. Les autres mutants protègent partiellement la souche C43. La mutation ponctuelle W46S a presque le même impact que celles des résidus FTE52-53YKK. De plus, la chaîne latérale du W46 est orientée vers l'extérieur, montrant que cette chaîne latérale doit être importante pour la liaison et qu'un certain nombre d'acides aminés hydrophobes stabilise l'interaction entre FhuA et Ac-Llp. Les mutants D28K, RK29-30SS et YK22-23AS semblent importants pour stabiliser l'interaction FhuA-Llp mais pas essentiels à la protection puisque dans des conditions de faibles concentrations de phage, ils sont capables de protéger la cellule. Les résultats confirment et complètent ceux obtenus en RMN. Les acides aminés qui semblaient le plus impactés lors de la titration semblent aussi ceux les plus importants pour la fonction de Ac-Llp.



**Figure 6-43 : Influence des mutations de Ac-Llp sur la protection des souches C43 portant les plasmides des mutants de Llp face au phage T5.**

Les acides aminés ont été colorés en orange quand ils sont impliqués dans l'interaction avec FhuA et en bleu quand ils ne le sont pas. Ces couleurs sont reportées sur la structure de Sol-Llp voir figure 6-44.

Les résidus peuvent être classés en deux catégories : Les mutations délétères ou les mutations ayant peu ou pas d'impact sur la fonction de Ac-Llp. On peut voir sur la figure 6-44 que les résidus Y14-D18, K21-A24, Y42-A47 et F52-E54 semblent former une surface d'interaction plutôt convaincante. En revanche les résidus K7-D8 et D28-K30, qui semblent moins importants pour l'interaction, semblent être situés sous la surface d'interaction. Ces acides aminés et les résidus F4 et N56 ne semblent pas en interaction directe avec FhuA. Leurs signaux RMN étant impactés lors de l'interaction avec FhuA, on peut supposer que la structure de Llp peut être affectée par la liaison de façon locale.



**Figure 6-44 : Synthèse des acides aminés pouvant être impliqués dans l'interaction avec FhuA sur la structure de Sol-Llp.**

Trois vues, à 90°, de Llp. Acide aminé en orange : impact observé sur l'interaction, en bleu : impact nul / léger sur l'interaction, acides aminés affectés en signal RMN mais non testés en mutation : jaune. Les numéros des acides aminés colorés ont été notés à côté de chaque zone. Les flèches noires indiquent le sens de la chaîne principale.

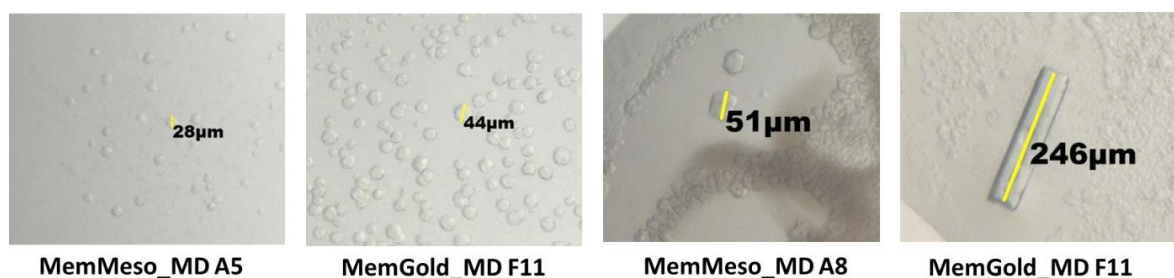
## 6.5 Études structurales du complexe FhuA-Llp

Un des objectifs de ma thèse est d'obtenir la structure de FhuA-Llp pour comprendre le mécanisme d'action de Llp pour bloquer l'interaction entre FhuA et pb5 et ainsi protéger la cellule de l'infection par un nouveau phage. Le temps imparti pour ma thèse ne m'a pas permis d'obtenir la structure finale du complexe. Néanmoins, plusieurs approches ont été testées pour obtenir cette structure (détaillées dans cette section) et poser les bases pour un travail futur pour la résolution structurale de ce complexe.

### 6.5.1 Cristallogenèse

#### 6.5.1.1 Macro cristaux

Des essais de cristallogenèse ont été menés sur la plateforme du HTXlab de l'EMBL avec succès. De multiples cristaux de petite taille (entre 20 et 240  $\mu\text{m}$ ), présents dans les plaques de cristallogenèse (FhuA:Ac-Llp 1:3 en DDM) (figure 6-45), ont été pêchés pour être testés à l'ESRF. Cependant, une erreur d'identification du carrousel n'a pas permis de tester la diffraction de ces cristaux. Certaines conditions ont été reproduite à la main (voir section 6.5.1.2), mais les cristaux obtenus dans ces conditions n'ont pas diffracté.



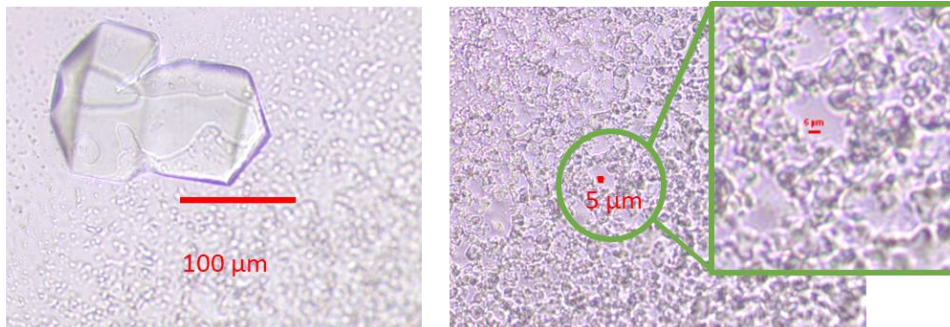
**Figure 6-45 : Exemple de cristaux obtenus sur la plateforme du HTX-lab.**

*Cristallisation du complexe FhuA : Ac-Llp en DDM avec un ratio 1 : 4 à une concentration de 8 mg/ml de protéine totale en présence du cristallogène TbXo4 (ref) (F11) ou de leur dérivé TbXo4-OH (A5) ou TbXo4 – phoPegNme (A8). Les plaques sont conservées à 4°C et le suivi de cristallisation est effectué par les photos fournies sur la plateforme CRIMS.*

#### 6.5.1.2 Microcristaux

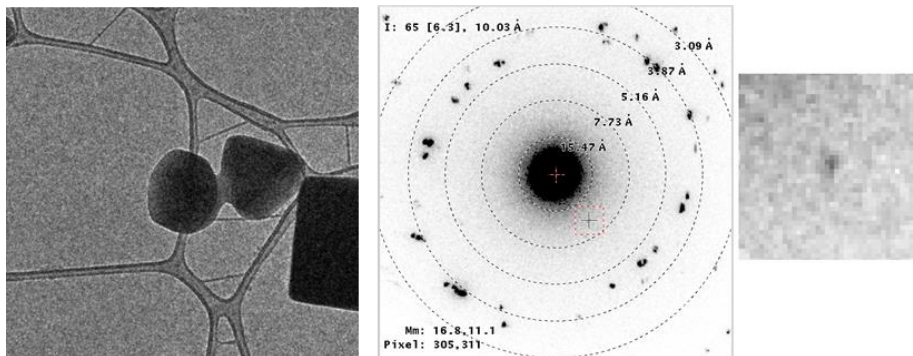
Ce travail a été mené en collaboration avec Dominique Housset. La diffraction électronique sur micro cristaux a été testée en parallèle, la condition F11 de la plaque de cristallographie MemGold\_MD (0.2M ammonium sulfate, 0.02M sodium acetate pH 4, 33% (v/v) PEG 200, 0.02M sodium chloride) ayant servi de base exploratoire pour former des microcristaux à la main. La reproduction à la main a permis d'obtenir des microcristaux de quelque  $\mu\text{m}$  et cristaux de l'ordre de la centaine de micromètre de pH3,5 à pH6 pour des concentrations de PEG de 30 à 45 % (Figure 6-46). Les microcristaux obtenus ont pu être observés au microscope électronique. En diffraction électronique, seulement une tâche de diffraction pouvant correspondre à un cristal de protéine a pu être observée (Figure 6-47).





**Figure 6-46 : Exemple de cristaux obtenus à la main.**

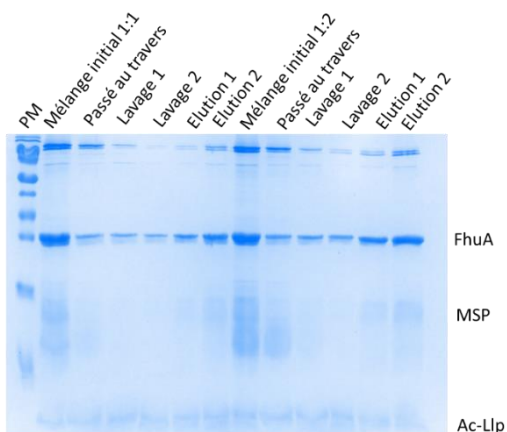
Cristallisation du complexe FhuA-Ac-Llp en DDM avec un ratio 1 : 4 à une concentration de 8 mg/ml de protéine totale en présence du cristallophore TbXo4. Cristaux obtenus autour de la condition F11 MemGold\_MD avec des variations de pH de 3,5 à 6 et des concentrations en PEG 200 de 30% à 45%.



**Figure 6-47 : Observation des microcristaux au microscope électronique : Cristaux du complexe FhuA : Ac-Llp en DDM avec un ratio 1 : 4 à une concentration de 8 mg. A gauche l'image d'un microcristal observé au microscope électronique à droite diffraction électronique de ce cristal.**

### 6.5.2 Formation de nanodisques contenant FhuA-Ac-Llp pour la microscopie électronique

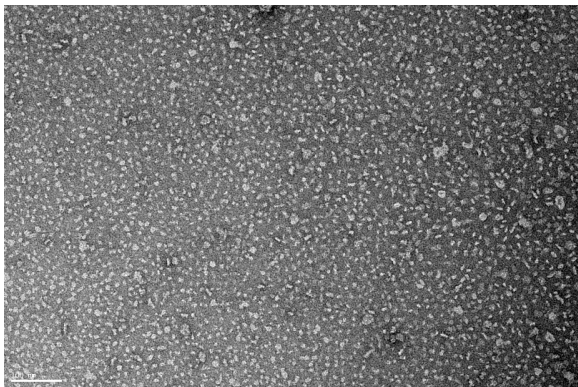
Afin d'obtenir la structure du complexe FhuA-Llp nous avons envisagé de stabiliser le complexe dans des nanodisques. Des nanodisques ont été reconstitués autour du complexe FhuA -Ac-Llp à des ratios FhuA : Llp de 1 : 1 ou 1 : 2. Après incubation de FhuA et Ac-Llp solubilisées en DDM avec la MSP et les lipides, le détergent a été éliminé par incubation avec des BioBeads et les nanodisques contenant FhuA et/ou Llp ont été purifiés sur une colonne de nickel. Les nanodisques ne se sont pas bien accrochés sur la colonne puisque l'on retrouve beaucoup de protéines dans le non-retenu de la colonne et dans les lavages (Figure 6-48). Au moment de l'éluion on retrouve bien FhuA, la MSP et Llp. Les échantillons sont donc observés en coloration négative pour évaluer la formation des nanodisques.



**Figure 6-48 : Suivi de la purification des nanodisques contenant FhuA et Ac-Llp sur colonne de NiNTA.**

Gel d'acrylamide en gradient coloré au bleu de Coomassie mélange initial 1 : 1 FhuA 3,75 nmol + Llp 3,75 nmol + MSP 22,5 nmol + lipide 1350 nmol, mélange initial 1 : 2 FhuA 3,75 nmol + Llp 7,5 nmol + MSP 22,5 nmol + lipide 1350 nmol, lavage 1 et 2 à une concentration de 50 mM d'imidazole. Elution 1 et 2 à une concentration de 300 mM imidazole.

La coloration négative montre que les nanodisques sont bien formés. On peut voir plusieurs orientations dans l'espace et une répartition homogène sur la grille (Figure 6-49). Avant d'aller plus loin dans la démarche de résolution des structures, on veut s'assurer que la majorité des nanodisques contiennent bien FhuA et Ac-Llp. Pour cela, les échantillons sont analysés en SEC-MALS et en tests phénotypiques. Malheureusement le SEC-MALS n'a pas pu déterminer avec précision les masses de protéines et de lipides car les signaux étaient faibles et le pic d'éluion très large, montrant une grande hétérogénéité dans l'échantillon (non montré). De plus, lors du test phénotypique, tous les phages ont été désactivés ce qui signifie que le complexe FhuA-Ac-Llp n'est pas majoritairement formé dans les nanodisques (non montré).



**Figure 6-49 : Vérification de la formation des nanodisques de FhuA : Llp ratio 1 : 1 par microscopie électronique.**  
*limage des nanodisques colorés avec 1% de SST (Sodium Silico Tungstate).*

Pour améliorer la formation du complexe dans les nanodisques, des ratios avec plus de Ac-Llp peuvent être testés. En outre, la sélection des nanodisques contenant à la fois FhuA et Ac-Llp peut être réalisée à l'aide d'une double purification avec une étiquette histidine pour FhuA et une autre étiquette pour Ac-Llp, comme une étiquette streptavidine. Nous n'avons pas mené plus d'investigations sur la formation de nanodisques, car nous avons voulu tester une autre approche, la formation du complexe par une protéine de fusion.

### 6.5.3 Formation de fusion FhuA-Llp pour la résolution de structure

Afin de stabiliser l'interaction entre FhuA et Llp, des protéines de fusion FhuA-Llp, sans ou avec lien entre les deux protéines (5 et 10 résidus), ont été pensées. La séquence codant pour la Llp, sans la séquence signal ni la cystéine terminale, a été insérée dans le plasmide codant pour FhuA en 3' du gène de FhuA. La protéine de fusion FhuA-Llp sans lien a été produite avec le même protocole que FhuA de même que la première étape de purification. Ceci a permis d'obtenir environ 1 mg de protéine fusion. Cette dernière a été incubée avec des bactériophages et leur intégrité a été contrôlée sur la souche F. Malheureusement, la fusion n'a pas montré une protection de T5 vis-à-vis de FhuA. Des protéines de fusion FhuA-Llp avec différentes longueurs de liens sont actuellement en train d'être clonées.

## 6.6 Discussion

La production et la purification de Ac-Llp a été difficile à reproduire. Cependant nous avons réussi à augmenter le rendement de purification décrit précédemment (1,8 mg de Llp acylée pour 30 g de culot bactérien<sup>51</sup> avec un rendement moyen de 5 mg pour 10 g de culot bactérien. Au cours de l'optimisation de la purification nous avons également testé une élution avec de l'EDTA. Or dans pedruzzi et al.<sup>51</sup> les auteurs montrent que le spectre de dichroïsme circulaire est différent entre Ac-Llp traité ou non à l'EDTA. Nous avons émis l'hypothèse que Ac-Llp pouvait contenir un cation divalent qui pourrait être important pour l'interaction. Les expériences complémentaires que nous avons menées n'ont pas permis de trancher la question dans un sens ou dans l'autre pour le moment.

Quand Ac-Llp est exprimée dans les souches BL21(DE3) ou C43(DE3), elle est capable de protéger la cellule avec une grande efficacité ( $10^8$  et  $10^{12}$  de protection respectivement). Cependant, dans la souche AW740, Llp n'est capable de protéger la souche que de  $10^2$ . La souche AW740, a la particularité d'être déletée pour la majorité des OMPs. Une étude<sup>36</sup> suggère que la présence de ces protéines est importante pour la protection des souches bactériennes par les protéines Cor, qui sont très semblables à la protéine Llp (voir introduction, section 6.4.4.1 et section 5.5 discussion). Cependant, Ac-Llp est capable d'inactiver FhuA *in vitro*. Le mécanisme par lequel les OMPs permettraient une protection des souches par les protéines Cor n'est pas connue. Le rôle des OMPs dans la protection de l'infection à T5 pourrait être confirmé par nos résultats. Le niveau d'expression de Ac-Llp pourrait être comparé au niveau d'expression de FhuA dans la souche AW740 par une immuno-révélation.

Le Kd de l'interaction entre FhuA et Ac-Llp n'a pas pu être mesuré avec précision à cause de la présence des détergents qui influencent beaucoup les méthodes biophysiques. Cependant, l'ensemble des résultats converge vers un Kd approximatif de 0,5 à 80  $\mu$ M. Pour la forme soluble le Kd mesuré est de 1,5 mM. Ceci montre que les acylations sont importantes pour l'interaction entre FhuA et Llp. L'acylation permet en effet d'ancrer Llp à la membrane permettant de limiter les degrés de liberté de Llp et probablement aussi d'orienter Llp pour l'interaction avec FhuA.

Les acides aminés Y14-18D, K21-A24, Y42-A47 et F52-E54 sont importants pour fixer FhuA et bloquer la liaison de pb5 à FhuA. La surface d'interaction de Llp est située autour de ces acides aminés impliquant aussi les acides aminés A24 et A47 qui sont affectés en titration par RMN. La surface d'interaction semble impliquer une vingtaine d'acides aminés avec quelques acides aminés chargés capables de former des ponts salins et des liaisons hydrogènes et une majorité d'acides aminés hydrophobes.

Le profil de dénaturation thermique de FhuA en présence de Llp suggère que le bouchon de FhuA est stabilisé par l'interaction avec Ac-Llp. Llp se lie probablement, au moins en partie, sur le bouchon de FhuA. Le bouchon est peu impliqué dans la liaison à pb5 et seulement 7 acides aminés sont impliqués dans l'interaction. En revanche, quand le fer ferrichrome est présent, pb5 ne peut pas se lier à FhuA en



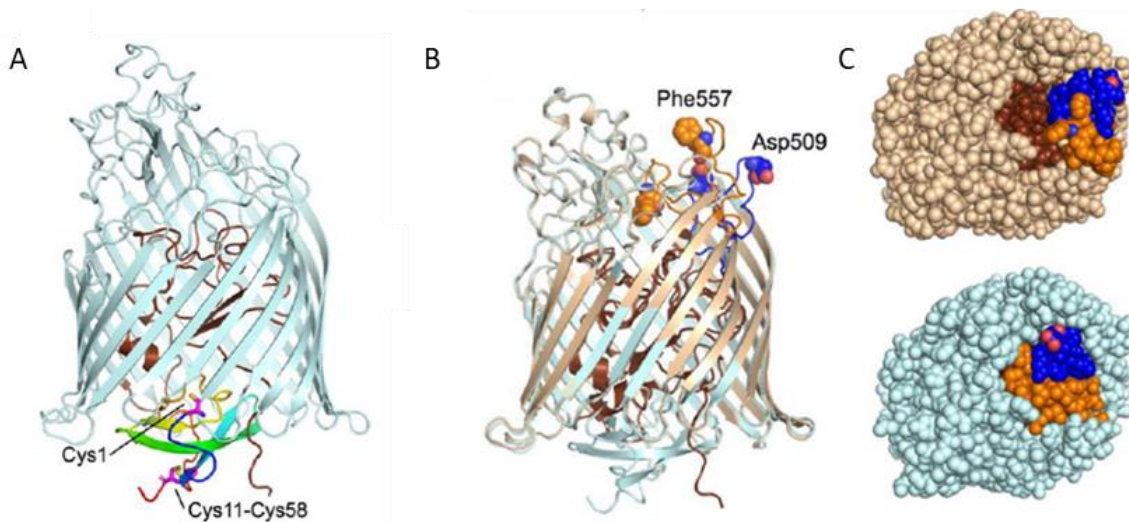
raison de l'encombrement stérique car les sites de liaison sont partiellement superposés <sup>60</sup>. On peut donc envisager que la liaison de Llp « pousserait » le bouchon plus haut entre les boucles externes entraînant des modifications de la surface d'interaction de FhuA à pb5.

La souche AW740 portant FhuA WT n'est protégée que de 10<sup>2</sup> par rapport au phage T5. Nous avons confirmé que les résidus I9 et D395 de FhuA sont importants pour l'inactivation de FhuA par Llp car la souche AW740 portant ces deux mutants ne sont pas protégée par la présence de Ac-Llp. Nous n'avons pas pu confirmer ce résultat pour le mutant L106P, mais les conditions du test *in vivo* ne permettent pas de conclure pour une réponse modérée à la non-protection de Llp. En revanche, nous avons programmé de compléter ces résultats en vérifiant la capacité de Ac-Llp à protéger les mutants de FhuA avec le test *in vitro* que nous avons mis au point. Les autres mutants ne semblent pas impliqués dans l'inactivation de FhuA par Llp, avec la réserve que leurs impacts soient modérés comme L106P. Les résultats de la dénaturation de FhuA montrent que l'ensemble des mutants de FhuA sont capables de lier Llp puisque le Tm du bouchon est augmenté. Il semble donc que les résidus I9, D395 et L106 soient importants pour l'inactivation de FhuA par Llp, mais pas pour sa liaison. Le mutant D395Y se trouve dans la boucle L5 et empêche Llp d'inactiver FhuA. En revanche, FhuA de *Salmonelle enterica bongori*, qui est insensible au phage T5 en présence de Llp, porte une variation D520N située dans la boucle 7. Les boucles externes semblent donc aussi impliquées dans l'inactivation de FhuA pour lier ces ligands.

Pris tous ensembles ces résultats suggèrent que le bouchon de FhuA est stabilisé ou poussé dans la structure de FhuA, permettant probablement le mouvement des boucles externes impliquées dans l'interaction avec pb5. Au vu de l'impact de la mutation D395Y dans la boucle 5, nous suggérons que celle-ci peut bouger pour empêcher l'interaction avec pb5.

Au cours de la rédaction de ce chapitre, un article a été proposé dans BiorXiv sur le complexe FhuA-pb5 et FhuA-Llp <sup>107</sup> (Figure 6-50 (A)). L'ensemble des données n'est pas encore accessible (notamment les fichiers PBD). Cependant, avec les éléments présentés dans l'article, nous pouvons valider certains de nos résultats. L'ensemble des acides aminés impliqués dans la surface d'interaction avec FhuA de Llp n'est pas décrit. Cependant, les acides aminés 15, 17, 18, 19, 23, 25, 39, 42, 44, et 46 sont décrits comme participant à des liaisons hydrogènes et/ou un pont salin. Nos résultats de RMN et l'analyse de nos mutants sont en parfait accord avec cette structure. En outre, dans la structure du complexe FhuA : Llp, le bouchon est bien impacté par la liaison de Llp. Llp se situe au niveau des acides aminés 20 à 37 de la structure de FhuA sous sa forme apo, et le haut du bouchon est décalé vers le haut de la structure, ce qui confirme notre hypothèse et explique parfaitement l'augmentation de Tm du bouchon de FhuA. En revanche, la boucle L5 n'est pas impliquée directement dans la protection de FhuA pour la liaison à pb5. En effet, ce sont les boucles L7 et L8 qui sont rabattues au niveau du site de fixation des ligands de FhuA, ce qui explique que l'ensemble des ligands transportés (ferrichrome, albomicine, Colicin M, ...) soit bloqué par Llp. Dans Degroux et al. (section 4.3), nous avons montré que pb5 se fixe

au même niveau que le ferrichrome et interagit avec les boucles 4 à 11. Le basculement des boucles L8 et L7 obstrue le site de fixation de pb5 par encombrement stérique, expliquant la protection des bactéries de l'infection par T5 (Figure 6-50 B). Il est difficile de déterminer avec exactitude comment la fixation de Llp provoque le basculement de deux boucles externes. Cependant, nous proposons que la transmission de l'information de liaison implique aussi la boucle L5 de façon transitoire au vu du phénotype du mutant D395Y. En outre Berg et al. 2022 propose que Llp se lie sur un état de transition de FhuA après le transport du fer ferrichrome. Nos observations ne sont pas en accord avec cette hypothèse. Leurs deux principaux arguments sont le fait que le mutant de FhuA I9P ne permet pas à la bactérie d'être protégée en présence de Llp, et qu'il forme le complexe FhuA : Llp avec un ratio 1 : 2 *in vitro* en 48 heures. Cependant, nous avons montré que nous pouvions former un complexe fonctionnel (ratio 1 : 50) en seulement 30 min. De plus, le mutant FhuA T12P empêchant la liaison de FhuA à TonB comme le mutant I9P peut être protégé par Llp *in vivo*. La liaison de FhuA à TonB et donc le transport du fer ferrichrome ne semble donc pas être un prérequis à la fixation de Llp. De plus, Llp est capable de protéger la cellule même en présence de bipyridyl (un chélateur de fer), réduisant la présence de « l'intermédiaire » nécessaire à la fixation de Llp selon Bert et al. <sup>107</sup>



**Figure 6-50 : Structure du complexe FhuA-Llp.**

A. structure de FhuA-Llp avec le tonneau beta de FhuA en bleu, le bouchon de FhuA en marron et Llp en rainbow. B. Superposition des structures de FhuA (crème et marron) et FhuA-Llp (bleu), Phe 557 et Asp 509 sont représentés en sphère dans les deux structures. C. représentation en sphère de FhuA (crème et marron) et de FhuA-Llp (bleu) avec Phe 557 en orange et Asp 509 en bleu montrant la fermeture du site de fixation des ligands de FhuA. Figure issue de 106.

## 7 Conclusion et perspective

Une des étapes clés du cycle d'infection par les bactériophages est la liaison au récepteur, qui initie l'infection et déclenche l'éjection de l'ADN dans le cytoplasme bactérien. Dans cette thèse, nous présentons la structure du complexe FhuA-pb5 par microscopie électronique à une résolution de 2,6 Å. Il s'agit de la première structure d'une RBP de phage liant un récepteur protéique. La structure de pb5 montre un repliement original partageant tout de même une homologie de structure en C-terminal avec un domaine de RBP de phage liant des sucres. La structure du complexe FhuA-pb5 montre que l'interaction entre FhuA et pb5 est principalement due à un réseau de liaisons hydrogènes et d'interactions hydrophobes, expliquant la liaison irréversible de pb5 sur FhuA. Cette structure, la prédiction de la structure de pb5 isolée par le nouveau logiciel AlphaFold2 ainsi que la structure de l'extrémité distale de la queue du phage après interaction avec FhuA reconstitué en nanodisques nous permet de proposer un mécanisme expliquant la transmission du signal d'interaction de pb5 à FhuA au reste de la queue de T5 (chapitre I-figure 4 de l'article). De plus, la structure du complexe FhuA-pb5 permet de connaître le site d'interaction de pb5 à FhuA. Cette information est essentielle pour comprendre l'inhibition de fixation de pb5 à FhuA lorsque Llp est fixée (voir ci-après).

La protéine Llp, une protéine Sie, codée par le phage T5, joue un rôle important dans la protection de l'usine virale au cours de la réplication mais aussi pour la protection des nouveaux virions de T5 produits. Nous avons cherché à décrypter le mécanisme d'action de cette protéine d'exclusion de surinfection. Nous avons résolu sa structure par RMN pour la forme soluble (Sol-Llp), révélant un repliement principalement en feuillets bêta avec une boucle flexible. Cette structure est une nouvelle structure de protéine de phage. Cependant, ce repliement est identique à celui d'une protéine codée par *Salmonella* en dehors d'une zone de prophage, et dont la fonction est inconnue. La comparaison de structure de Llp avec les prédictions de structure des protéines Cor permet de supposer que leur structure et donc leur mécanisme d'action sont identiques, notamment sur des récepteurs proches de FhuA comme ButB.

L'étude du complexe FhuA-Llp avec diverses méthodes biochimiques et biophysiques a permis de définir un  $K_D$  de 1,5 mM pour Sol-Llp et une estimation de  $K_D$  pour la protéine acylée de 0,5 à 80  $\mu$ M, ce qui suggère que l'acylation joue un rôle important dans la formation du complexe. La titration de Sol-Llp par FhuA en RMN montre des perturbations d'intensité et de déplacements chimiques pour les résidus 12 à 24, et 51 à 54, ce qui suggère qu'ils sont impliqués dans l'interaction avec FhuA. Des mutants de ces résidus ainsi que des résidus 39 à 45 - non visibles en RMN - confirment l'importance de ces résidus dans l'interaction avec FhuA, les protéines mutantes étant incapables de protéger la cellule d'une infection à T5 contrairement à la protéine sauvage. L'augmentation du  $T_m$  du bouchon de FhuA au cours de la dénaturation suggère que le bouchon de FhuA est impliqué dans la liaison de Llp. La structure de FhuA-Llp a été récemment publiée, permettant de confirmer une partie de nos résultats. La liaison de Llp permet de faire basculer les boucles 7 et 8 fermant ainsi le site de fixation de l'ensemble

des ligands de FhuA. L'analyse détaillée de cette structure permettra sûrement de proposer un mécanisme de transmission de l'information entre la surface périplasmique de FhuA et son domaine extracellulaire. En outre, des mutants de Llp protègent partiellement la cellule : la résolution du complexe FhuA-Llp mutant pourrait peut-être capturer des intermédiaires de FhuA entre la forme ouverte et fermée, permettant de mieux comprendre les changements de conformation du bouchon de FhuA, par exemple dans le mécanisme de transport de son ligand physiologique.



# Bibliographies

1. Ansaldi, M. *et al.* Un siècle de recherche sur les bactériophages. **24**, 15 (2020).
2. Clokie, M. R. J., Millard, A. D., Letarov, A. V. & Heaphy, S. Phages in nature. *Bacteriophage* **1**, 31–45 (2011).
3. Twort, F. W. Further Investigations on the Nature of Ultra-Microscopic Viruses and their Cultivation. *J. Hyg. (Lond.)* **36**, 204–235 (1936).
4. Summers, W. C. In the beginning..2. *Bacteriophage* **1**, 50–51 (2011).
5. Chanishvili, N. Phage Therapy—History from Twort and d’Herelle Through Soviet Experience to Current Approaches. in *Advances in Virus Research* vol. 83 3–40 (Elsevier, 2012).
6. Keen, E. C. A century of phage research: Bacteriophages and the shaping of modern biology: Cause to reflect. *BioEssays* **37**, 6–9 (2015).
7. Lamarck, J.-B. Histoire naturelle des animaux sans vertèbre. in *Histoire naturelle des animaux sans vertèbre* 49. (1815).
8. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
9. Chibani-Chennoufi, S., Bruttin, A., Dillmann, M.-L. & Brussow, H. Phage-Host Interaction: an Ecological Perspective. *J. Bacteriol.* **186**, 3677–3686 (2004).
10. Danovaro, R. *et al.* Marine viruses and global climate change. *FEMS Microbiol. Rev.* **35**, 993–1034 (2011).
11. Hambly, E. & Suttle, C. A. The virosphere, diversity, and genetic exchange within phage communities. *Curr. Opin. Microbiol.* **8**, 444–450 (2005).
12. Pujato, S. A., Quiberoni, A. & Mercanti, D. J. Bacteriophages on dairy foods. *J. Appl. Microbiol.* **126**, 14–30 (2019).
13. Adriaenssens, E. M. *et al.* Taxonomy of prokaryotic viruses: 2018-2019 update from the ICTV Bacterial and Archaeal Viruses Subcommittee. *Arch. Virol.* **165**, 1253–1260 (2020).
14. Novik, G., Ladutska, A. & Rakhuba, D. Bacteriophage taxonomy and classification. 9.



15. Ackermann, H.-W. Phage Classification and Characterization. in *Bacteriophages* (eds. Clokie, M. R. J. & Kropinski, A. M.) vol. 501 127–140 (Humana Press, 2009).
16. Harada, L. K. *et al.* Biotechnological applications of bacteriophages: State of the art. *Microbiol. Res.* **212–213**, 38–58 (2018).
17. Abedon, S. T. Chapter 1 Phage Evolution and Ecology. in *Advances in Applied Microbiology* vol. 67 1–45 (Elsevier, 2009).
18. Al-Shayeb, B. *et al.* Clades of huge phages from across Earth’s ecosystems. *Nature* **578**, 425–431 (2020).
19. Bacteriophages: Biology and Applications. *CRC Press*  
<https://www.routledge.com/Bacteriophages-Biology-and-Applications/Kutter-Sulakvelidze/p/book/9780849313363>.
20. Larsson, D. G. J. & Flach, C.-F. Antibiotic resistance in the environment. *Nat. Rev. Microbiol.* (2021) doi:10.1038/s41579-021-00649-x.
21. Kuek, M., McLean, S. K. & Palombo, E. A. Application of bacteriophages in food production and their potential as biocontrol agents in the organic farming industry. *Biol. Control* **165**, 104817 (2022).
22. Carvalho, C. M. *et al.* The in vivo efficacy of two administration routes of a phage cocktail to reduce numbers of *Campylobacter coli* and *Campylobacter jejuni* in chickens. *BMC Microbiol.* **10**, 232 (2010).
23. Jault, P. *et al.* Efficacy and tolerability of a cocktail of bacteriophages to treat burn wounds infected by *Pseudomonas aeruginosa* (PhagoBurn): a randomised, controlled, double-blind phase 1/2 trial. *Lancet Infect. Dis.* **19**, 35–45 (2019).
24. HISTOIRE DE LA PHAGOTHERAPIE A LYON. <https://www.crioac-lyon.fr/document/histoire-de-la-phagotherapie-a-lyon/>.
25. Adhya, S., Merrill, C. R. & Biswas, B. Therapeutic and Prophylactic Applications of Bacteriophage Components in Modern Medicine. *Cold Spring Harb. Perspect. Med.* **4**, a012518–a012518 (2014).

26. Dufour, N. & Debarbieux, L. La phagothérapie: Une arme crédible face à l'antibiorésistance. *médecine/sciences* **33**, 410–416 (2017).
27. Rami, A., Behdani, M., Yardehnavi, N., Habibi-Anbouhi, M. & Kazemi-Lomedasht, F. An overview on application of phage display technique in immunological studies. *Asian Pac. J. Trop. Biomed.* **7**, 599–602 (2017).
28. Mäntynen, S., Laanto, E., Oksanen, H. M., Poranen, M. M. & Díaz-Muñoz, S. L. Black box of phage–bacterium interactions: exploring alternative phage infection strategies. *Open Biol.* **11**, 210188 (2021).
29. Suttle, C. A. Marine viruses — major players in the global ecosystem. *Nat. Rev. Microbiol.* **5**, 801–812 (2007).
30. van Houte, S., Buckling, A. & Westra, E. R. Evolutionary Ecology of Prokaryotic Immune Mechanisms. *Microbiol. Mol. Biol. Rev.* **80**, 745–763 (2016).
31. Rostøl, J. T. & Marraffini, L. (Ph)ighting Phages: How Bacteria Resist Their Parasites. *Cell Host Microbe* **25**, 184–194 (2019).
32. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).
33. Cumby, N., Davidson, A. R. & Maxwell, K. L. The moron comes of age. *Bacteriophage* **2**, e23146 (2012).
34. Folimonova, S. Y. Superinfection Exclusion Is an Active Virus-Controlled Function That Requires a Specific Viral Protein. *J. Virol.* **86**, 5554–5561 (2012).
35. Cumby, N., Edwards, A. M., Davidson, A. R. & Maxwell, K. L. The Bacteriophage HK97 gp15 Moron Element Encodes a Novel Superinfection Exclusion Protein. *J. Bacteriol.* **194**, 5012–5019 (2012).
36. Arguijo-Hernández, E. S. *et al.* Cor interacts with outer membrane proteins to exclude FhuA-dependent phages. *Arch. Virol.* **163**, 2959–2969 (2018).

37. Domingo-Calap, P., Mora-Quilis, L. & Sanjuán, R. Social Bacteriophages. *Microorganisms* **8**, 533 (2020).
38. Hunter, M. & Fusco, D. Superinfection exclusion: A viral strategy with short-term benefits and long-term drawbacks. *PLOS Comput. Biol.* **18**, e1010125 (2022).
39. Braun, V. FhuA (TonA), the Career of a Protein. *J. Bacteriol.* **191**, 3431–3436 (2009).
40. Wang, J. *et al.* Complete genome sequence of bacteriophage T5. *Virology* **332**, 45–65 (2005).
41. Zivanovic, Y. *et al.* Insights into Bacteriophage T5 Structure from Analysis of Its Morphogenesis Genes and Protein Components. *J. Virol.* **88**, 1162–1174 (2014).
42. Decker, K., Krauel, V., Meesmann, A. & Heller, K. J. Lytic conversion of *Escherichia coli* by bacteriophage T5: blocking of the FhuA receptor protein by a lipoprotein expressed early during infection. *Mol. Microbiol.* **12**, 321–332 (1994).
43. Effantin, G., Boulanger, P., Neumann, E., Letellier, L. & Conway, J. F. Bacteriophage T5 Structure Reveals Similarities with HK97 and T4 Suggesting Evolutionary Relationships. *J. Mol. Biol.* **361**, 993–1002 (2006).
44. Böhm, J. *et al.* FhuA-mediated phage genome transfer into liposomes: A cryo-electron tomography study. *Curr. Biol.* **11**, 1168–1175 (2001).
45. Huet, A., Duda, R. L., Boulanger, P. & Conway, J. F. Expansion of the bacteriophage T5 capsid: a highly dynamic process resolved at high-resolution. 42.
46. Huet, A., Conway, J. F., Letellier, L. & Boulanger, P. *In Vitro* Assembly of the T=13 Procapsid of Bacteriophage T5 with Its Scaffolding Domain. *J. Virol.* **84**, 9350–9358 (2010).
47. Linares, R. *et al.* Structural basis of bacteriophage T5 infection trigger and *E. coli* cell wall perforation. <http://biorxiv.org/lookup/doi/10.1101/2022.09.20.507954> (2022)  
doi:10.1101/2022.09.20.507954.
48. Linares, R., Arnaud, C.-A., Degroux, S., Schoehn, G. & Breyton, C. Structure, function and assembly of the long, flexible tail of siphophages. *Curr. Opin. Virol.* **45**, 34–42 (2020).

49. Heller, K. & Braun, V. Polymannose O-antigens of *Escherichia coli*, the binding sites for the reversible adsorption of bacteriophage T5+ via the L-shaped tail fibers. *J. Virol.* **41**, 222–227 (1982).
50. Lanni, Y. T. First-step-transfer deoxyribonucleic acid of bacteriophage T5. *BACTERIOL REV* **32**, 16 (1968).
51. Pedruzzi, I., Rosenbusch, J. P. & Locher, K. P. Inactivation in vitro of the *Escherichia coli* outer membrane protein FhuA by a phage T5-encoded lipoprotein. *FEMS Microbiol. Lett.* **168**, 119–125 (1998).
52. Arnaud, C.-A. *et al.* Bacteriophage T5 tail tube structure suggests a trigger mechanism for Siphoviridae DNA ejection. *Nat. Commun.* **8**, 1953 (2017).
53. Noinaj, N., Guillier, M., Barnard, T. J. & Buchanan, S. K. TonB-Dependent Transporters: Regulation, Structure, and Function. *Annu. Rev. Microbiol.* **64**, 43–60 (2010).
54. Ferguson, A. D., Coulton, J. W., Diederichs, K. & Welte, W. The Ferric Hydroxamate Uptake Receptor FhuA and Related TonB-Dependent Transporters in the Outer Membrane of Gram-Negative Bacteria. in *Handbook of Metalloproteins* (eds. Messerschmidt, A. *et al.*) met160 (John Wiley & Sons, Ltd, 2006). doi:10.1002/0470028637.met160.
55. Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K. & Welte, W. Siderophore-Mediated Iron Transport: Crystal Structure of FhuA with Bound Lipopolysaccharide. *Science* **282**, 2215–2220 (1998).
56. Locher, K. P. *et al.* Transmembrane Signaling across the Ligand-Gated FhuA Receptor: Crystal Structures of Free and Ferrichrome-Bound States Reveal Allosteric Changes. **8**.
57. Bonhivers, M. *et al.* Stability Studies of FhuA, a Two-Domain Outer Membrane Protein from *Escherichia coli*. *Biochemistry* **40**, 2606–2613 (2001).
58. Krauel, V. & Heller, K. J. Cloning, sequencing, and recombinational analysis with bacteriophage BF23 of the bacteriophage T5 oad gene encoding the receptor-binding protein. *J. Bacteriol.* **173**, 1287–1297 (1991).

59. Mondigler, M., Holz, T. & Heller, K. J. Identification of the Receptor-Binding Regions of pb5 Proteins of Bacteriophages T5 and BF23. *Virology* **219**, 19–28 (1996).
60. Plançon, L. *et al.* Characterization of a High-affinity Complex Between the Bacterial Outer Membrane Protein FhuA and the Phage T5 Protein pb5. *J. Mol. Biol.* **318**, 557–569 (2002).
61. 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. *J. Biol. Chem.* **288**, 30763–30772 (2013).
62. Flayhan, A., Wien, F., Paternostre, M., Boulanger, P. & Breyton, Ce. New insights into pb5, the receptor binding protein of bacteriophage T5, and its interaction with its Escherichia coli receptor FhuA. *8* (2012).
63. Mills, A., Le, H.-T., Coulton, J. W. & Duong, F. FhuA interactions in a detergent-free nanodisc environment. *Biochim. Biophys. Acta BBA - Biomembr.* **1838**, 364–371 (2014).
64. Braun, M., Endriss, F., Killmann, H. & Braun, V. In Vivo Reconstitution of the FhuA Transport Protein of *Escherichia coli* K-12. *J. Bacteriol.* **185**, 5508–5518 (2003).
65. Endriß, F. & Braun, V. Loop Deletions Indicate Regions Important for FhuA Transport and Receptor Functions in *Escherichia coli*. *J. Bacteriol.* **186**, 4818–4823 (2004).
66. Braun, V., Killmann, H. & Herrmann, C. Inactivation of FhuA at the cell surface of *Escherichia coli* K-12 by a phage T5 lipoprotein at the periplasmic face of the outer membrane. *J. Bacteriol.* **176**, 4710–4717 (1994).
67. Zückert, W. R. Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **1843**, 1509–1516 (2014).
68. Kaplan, E., Greene, N. P., Crow, A. & Koronakis, V. Insights into bacterial lipoprotein trafficking from a structure of LolA bound to the LolC periplasmic domain. *Proc. Natl. Acad. Sci.* **115**, (2018).
69. Yamaguchi, K., Yu, F. & Inouye, M. A single amino acid determinant of the membrane localization of lipoproteins in *E. coli*. *Cell* **53**, 423–432 (1988).

70. Seydel, A., Gounon, P. & Pugsley, A. P. Testing the '+2 rule' for lipoprotein sorting in the Escherichia coli cell envelope with a new genetic selection. *Mol. Microbiol.* **34**, 810–821 (1999).
71. Robichon, C., Bonhivers, M. & Pugsley, A. P. An intramolecular disulphide bond reduces the efficacy of a lipoprotein plasma membrane sorting signal: Lipoprotein localization. *Mol. Microbiol.* **49**, 1145–1154 (2003).
72. Coulton, J. W., Mason, P. & DuBow, M. S. Molecular cloning of the ferrichrome-iron receptor of Escherichia coli K-12. *J. Bacteriol.* **156**, 1315–1321 (1983).
73. Ferguson, A. D., Diederichs, K., Welte, W. & Coulton, J. W. An internal affinity-tag for purification and crystallization of the siderophore receptor FhuA, integral outer membrane protein from Escherichia coli K-12. *J. Biol. Chem.* **273**, 1013–1018 (1998).
74. Dumon-Seignovert, L., Cariot, G. & Vuillard, L. The toxicity of recombinant proteins in Escherichia coli: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr. Purif.* **37**, 203–206 (2004).
75. Ingham, C., Buechner, M., Adler, J. & Bacteriol, J. Effect of outer membrane permeability on chemotaxis in Escherichia coli. *J BACTERIOL* **8** (1990).
76. Bayburt, T. H. & Sligar, S. G. Membrane protein assembly into Nanodiscs. *FEBS Lett.* **584**, 1721–1727 (2010).
77. Denisov, I. G., Baas, B. J., Grinkova, Y. V. & Sligar, S. G. Cooperativity in Cytochrome P450 3A4. *J. Biol. Chem.* **282**, 7066–7076 (2007).
78. Homola, J. Surface Plasmon Resonance Sensors for Detection of Chemical and Biological Species. *Chem. Rev.* **108**, 462–493 (2008).
79. Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **78**, 1606–1619 (2000).
80. Zhao, H. *et al.* A Multilaboratory Comparison of Calibration Accuracy and the Performance of External References in Analytical Ultracentrifugation. *PLOS ONE* **10**, e0126420 (2015).

81. Brautigam, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. in *Methods in Enzymology* vol. 562 109–133 (Elsevier, 2015).
82. Salvay, A. G. & Ebel, C. Analytical Ultracentrifuge for the Characterization of Detergent in Solution. in *Analytical Ultracentrifugation VIII* (eds. Wandrey, C. & Cölfen, H.) vol. 131 74–82 (Springer-Verlag, 2006).
83. Salvay, A. G., Santamaria, M., le Maire, M. & Ebel, C. Analytical Ultracentrifugation Sedimentation Velocity for the Characterization of Detergent-Solubilized Membrane Proteins Ca<sup>++</sup>-ATPase and ExbB. *J. Biol. Phys.* **33**, 399–419 (2007).
84. Pedersen, M. E., Østergaard, J. & Jensen, H. In-Solution IgG Titer Determination in Fermentation Broth Using Affibodies and Flow-Induced Dispersion Analysis. *ACS Omega* **5**, 10519–10524 (2020).
85. Skinner, S. P. *et al.* CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR* **66**, 111–124 (2016).
86. Volk, J., Herrmann, T. & Wüthrich, K. Automated sequence-specific protein NMR assignment using the memetic algorithm MATCH. *J. Biomol. NMR* **41**, 127–138 (2008).
87. Fiorito, F., Herrmann, T., Damberger, F. & Wüthrich, K. Automated amino acid side-chain NMR assignment of proteins using <sup>13</sup>C- and <sup>15</sup>N-resolved 3D [1H, 1H]-NOESY. *J. Biomol. NMR* **42**, 23–33 (2008).
88. Guerry, P., Duong, V. D. & Herrmann, T. CASD-NMR 2: robust and accurate unsupervised analysis of raw NOESY spectra and protein structure determination with UNIO. *J. Biomol. NMR* **62**, 473–480 (2015).
89. Guerry, P. & Herrmann, T. Comprehensive automation for NMR structure determination of proteins. *Methods Mol. Biol. Clifton NJ* **831**, 429–451 (2012).
90. Herrmann, T., Güntert, P. & Wüthrich, K. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J. Biomol. NMR* **24**, 171–189 (2002).



91. Herrmann, T., Güntert, P. & Wüthrich, K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* **319**, 209–227 (2002).
92. Güntert, P., Mumenthaler, C. & Wüthrich, K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283–298 (1997).
93. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**, 2728–2733 (2007).
94. Mareuil, F., Malliavin, T. E., Nilges, M. & Bardiaux, B. Improved reliability, accuracy and quality in automated NMR structure calculation with ARIA. *J. Biomol. Nmr* **62**, 425–438 (2015).
95. Gore, S. *et al.* Validation of Structures in the Protein Data Bank. *Structure* **25**, (2017).
96. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
97. Lakomek, N.-A., Ying, J. & Bax, A. Measurement of <sup>15</sup>N relaxation rates in perdeuterated proteins by TROSY-based methods. *J. Biomol. NMR* **53**, 209–221 (2012).
98. Woźniak-Braszak, A., Jurga, K. & Baranowski, M. The Lipari-Szabo Model-Free Analysis as a Method for Study of Molecular Motion in Solid State Heteronuclear Systems Using NMR Off-Resonance. *Appl. Magn. Reson.* **47**, 567–574 (2016).
99. Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
100. Pawelek, P. D. *et al.* Structure of TonB in Complex with FhuA, *E. coli* Outer Membrane Receptor. *Science* **312**, 1399–1402 (2006).
101. Engilberge, S. *et al.* Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chem. Sci.* **8**, 5909–5917 (2017).
102. Veesler, D. & Cambillau, C. A Common Evolutionary Origin for Tailed-Bacteriophage Functional Modules and Bacterial Machineries. *Microbiol. Mol. Biol. Rev.* **75**, 423–433 (2011).

103. Mirdita, M. *et al.* ColabFold: Making Protein folding accessible to all. *Nat. Methods* (2022)  
doi:10.1038/s41592-022-01488-1.
104. Dali server: structural unification of protein families | Nucleic Acids Research | Oxford  
Academic. <https://academic.oup.com/nar/article/50/W1/W210/6591528?login=false>.
105. Breyton, C. *et al.* Assemblies of lauryl maltose neopentyl glycol (LMNG) and LMNG-solubilized  
membrane proteins. *Biochim. Biophys. Acta BBA - Biomembr.* **1861**, 939–957 (2019).
106. Arechaga, I. *et al.* Characterisation of new intracellular membranes in *Escherichia coli*  
accompanying large scale over-production of the b subunit of F1Fo ATP synthase. *FEBS Lett.* **482**,  
215–219 (2000).
107. van den Berg, B., Silale, A., Baslé, A., Mader, S. L. & Khalid, S. *Structural basis for host  
recognition and superinfection exclusion by bacteriophage T5.*  
<http://biorxiv.org/lookup/doi/10.1101/2022.06.28.497910> (2022) doi:10.1101/2022.06.28.497910.

# ANNEXES



**pET20b-AcLlp D18K21K23-NNQ -His**

Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Nebi : AY543070		Plasmide pEt20b (fuite promoteur) bacterial expression		Resistance Ampiciline
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 92,4 ng/ $\mu$ l V = 20 $\mu$ l	Az60280 1,83
Caractéristiques de la construction				
WT seq C Llp G HIS				
Séquence d'adressage Llp Wild type / Llp : WT partie mature /C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6*His.				
Objectifs : Mutant de Llp basé sur les données RMN pour investiguer l'interaction entre Llp et FhuA.				
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Llp-2 (X049772)	p175-179	Référence séquençage 50-604152729	
Clonage : Mutagenèse dirigée à partir de la construction pET20b-wt-Llp-G-His6 (ou pET20b-AcLlpWTHis) Séquence nucléique annotée :				
<p><i>ttlaagaaaggagatacaatATGAAAAAATATTTAGCTATGGCAGGGTACTCTTTGGGCTTGC</i>  <i>TCTACTTTGGACCTAAAGATAATAAATGTAGGCATACTATATGCAGAACCATGTGAAC</i>  <i>TATCAGGCTAATGTTTTGATAGGAAAGGGGATATGTTTTGGTGTCACCAATATGGCCT</i>  <i>ATGGATCTTCTGGCCCCCGTAAGCTACTTACCGAAGGAAATA CATGCGAGGGAGTTT</i>  <i>TCGGCCATCACCATCACCATCAC</i> TGAgcgcac-cgctgctgctgg</p>				
Forward Primer : 5' -> 3' D18K21K23-NNQ F GTGAGGCATACTATATGCAGAACCATGTGAACT ATCAGGCTAATGTTTTGATAGGAAAG	Reverse Primer : 5' -> 3' D18K21K23-NNQ R CTTTCCTATCAAAAACATTAGCCCTGA TAGTTCACATGGTCTGCATATAGTA TGCCTCAC	Taille construction : 3920		
Séquence protéique annotée :				
MKKLLFLAMAVVLLSACSTFGPKDKICEAYMQNHVNYQANVFDRKGDMLVSPIMAYGSF WAPVSYFTEGNTCEGVFGHHHHHH*				
Nbre AA avec seq signal = 84	MW=9574 Da	$\epsilon$ = 13075 M <sup>-1</sup> .cm <sup>-1</sup> (SS) $\epsilon$ = 12950 M <sup>-1</sup> .cm <sup>-1</sup> (SH)	pl=7,04	
Nbre AA sans seq signal = 69	MW=7956,9Da	$\epsilon$ = 13075 M <sup>-1</sup> .cm <sup>-1</sup> (SS) $\epsilon$ = 12950 M <sup>-1</sup> .cm <sup>-1</sup> (SH)	pl=6,37	

**pET20b-AcLlp FTE52-54YKK-His**

Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Nebi : AY543070		Plasmide pEt20b (fuite promoteur)		Resistance Ampiciline
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 87,5 ng/ $\mu$ l V = 20 $\mu$ l	Az60280 1,83
Caractéristiques de la construction				
WT seq C Llp G HIS				
Séquence d'adressage Llp Wild type / Llp : WT partie mature /C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6*His.				
Objectifs : Mutant de Llp basé sur les données RMN pour investiguer l'interaction entre Llp et FhuA.				
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Llp-2 (X049772)	p175-179	Référence séquençage 50-604152729	
Clonage : Mutagenèse dirigée à partir de la construction pET20b-wt-Llp-G-His6 (ou pET20b-AcLlpWTHis) Séquence nucléique annotée :				
<p><i>ttlaagaaaggagatacaatATGAAAAAATATTTAGCTATGGCAGGGTACTCTTTGGGCTTGC</i>  <i>TCTACTTTGGACCTAAAGATAATAAATGTAGGCATACTATATGCAGGCATATGTGAAA</i>  <i>TATAAGGCTAATGTTTTGATAGGAAAGGGGATATGTTTTGGTGTCACCAATATGGCCT</i>  <i>ATGGATCTTCTGGCCCCCGTAAGCTACTATAAAAA GGAATACATGCGAGGGAGTTT</i>  <i>TCGGCCATCACCATCACCATCAC</i> TGAgcgcac-cgctgctgctgg</p>				
Forward Primer : 5' -> 3' FTE52-54YKK F GGGCCCCGTAAGCTACTATAAAAAAGG AAATACATGCGAGGGAG	Reverse Primer : 5' -> 3' FTE52-54YKK R CTCCCCTGCATGTATTCCTTTTATAGTA GCTTACGGGGCCC	Taille construction : 3920		
Séquence protéique annotée :				
MKKLLFLAMAVVLLSACSTFGPKDKICEAYMQNDHVKYEANVFDRKGDMLVSPIMAYGSF WAPVSYKKGNTEGVFGHHHHHH*				
Nbre AA avec seq signal = 84	MW=9631 Da	$\epsilon$ = 14565 M <sup>-1</sup> .cm <sup>-1</sup> (SS) $\epsilon$ = 14440 M <sup>-1</sup> .cm <sup>-1</sup> (SH)	pl=9,03	
Nbre AA sans seq signal = 69	MW=8014 Da	$\epsilon$ = 14565 M <sup>-1</sup> .cm <sup>-1</sup> (SS) $\epsilon$ = 14440 M <sup>-1</sup> .cm <sup>-1</sup> (SH)	pl=8,5	

**pET20b-AcLip YK22-23AS-His**

Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Ncbi : AY543070		Plasmide pEt20b (fuite promoteur)		Résistance Ampicilline	
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 121.6ng/µl V = 20 µl	A <sub>260/280</sub> 1,84	
Caractéristiques de la construction					
WT seq C Lp G HIS					
Séquence d'adressage Lp Wild type / Lp : WT partie mature /C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6*His.					
Objectifs : Mutant de Lp basé sur les données RMN pour investiguer l'interaction entre Lp et FhuA.					
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Lp-2 (X049772)p175-179		Référence séquençage 50-604152729		
Clonage :					
Mutagenèse dirigée à partir de la construction pET20b-wt-Lp-G-His6 (ou pET20b-AcLpWTHis)					
Séquence nucléique annotée :					
<p><i>tttaagcagggatatacaATGAAAAAATATTTTAGCTATGGCAGTGGTACTCTTTGGCTTGC</i>  <i>TCTACTTTGGACCTAAAGATATAAATGTGAGGCATATAATGCAGGATCATGTGAAA</i>  <i>CGAGCGCTAAATGTTTTGATAGGAAAGGGATATGTTTTGGTGTCCACCAATTAATGGCC</i>  <i>TATGGATCTTTCTGGCCCCCGTAAGCTACTTTACCGAAGGAAATACATCGGAGGGAGTT</i>  <i>TTCCGGC CATCACCATCACATCACTGAgccgcgctgctgctg</i></p>					
Forward Primer : 5' -> 3' YK22-23AS F		Reverse Primer : 5' -> 3' YK22-23AS R			
CTATATGCAGGATCATGTGAAAGCGAGCGCTA		CTTTCCTATCAAAAACATTAAGCGCTC			
ATGTTTTGATAGGAAAAG		GCCTTTCACATGATCCTGCATATAG			
Taille gène : 255		Taille construction : 3920			
Séquence protéique annotée :					
MKKFLAMAVVLLSACSFTGPKDKICEAYYMQDHWKASANVFDKGMFLVSPIMAYGSF WAPVSYFTEGNTCEGVFGHHHHHHH*					
Nbre AA avec seq signal = 84		MW=9455,9 Da		pI=7,04	
		ε= 11 585 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 11 460 M <sup>-1</sup> .cm <sup>-1</sup> (SH)			
Nbre AA sans seq signal = 69		MW=7838,8 Da		pI=6,37	
		ε= 11 585 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 11 460 M <sup>-1</sup> .cm <sup>-1</sup> (SH)			

**pET20b-AcLip YMQD14-18AFAAK -His**

Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Ncbi : AY543070		Plasmide pEt20b (fuite promoteur)		Résistance Ampicilline	
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 94 ng/µl V = 20 µl	A <sub>260/280</sub> 1,83	
Caractéristiques de la construction					
WT seq C Lp G HIS					
Séquence d'adressage Lp Wild type / Lp : WT partie mature /C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6*His.					
Objectifs : Mutant de Lp basé sur les données RMN pour investiguer l'interaction entre Lp et FhuA.					
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Lp-2 (X049772)p175-179		Référence séquençage 50-604152729		
Clonage :					
Mutagenèse dirigée à partir de la construction pET20b-wt-Lp-G-His6 (ou pET20b-AcLpWTHis)					
Séquence nucléique annotée :					
<p><i>tttaagcagggatatacaATGAAAAAATATTTTAGCTATGGCAGTGGTACTCTTTGGCTTGC</i>  <i>TCTACTTTGGACCTAAAGATATAAATGTGAGGCAGCATTCGCAGCAACATGTGAAA</i>  <i>TATAAGGCTAAATGTTTTGATAGGAAAGGGATATGTTTTGGTGTCCACCAATTAATGGCCT</i>  <i>ATGGATCTTTCTGGCCCCCGTAAGCTACTTTACCGAAGGAAATACATCGGAGGGAGTTT</i>  <i>TCGGCCATCACCATCACATCACTGAgccgcgctgctgctg</i></p>					
Forward Primer : 5' -> 3' YMQD14-18AFAAK F		Reverse Primer : 5' -> 3' YMQD14-18AFAAK R			
CTAAAGATATAAATGTGAGGCAGCATTCGCAG		CATATAGTATGCCTCACATTAATTC			
CAAAAACATGTGAAATATAAGGCTAATG		TTGCAGGTCCAAAAGTAGAGCAAGC			
Taille gène : 255		Taille construction : 3920			
Séquence protéique annotée :					
MKKFLAMAVVLLSACSFTGPKDKICEAAFAAKHVKYKANVFDKGMFLVSPIMAYGSF WAPVSYFTEGNTCEGVFGHHHHHHH*					
Nbre AA avec seq signal = 84		MW=9376,9 Da		pI=8,87	
		ε= 10095 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 9970 M <sup>-1</sup> .cm <sup>-1</sup> (SH)			
Nbre AA sans seq signal = 69		MW=7759,8 Da		pI=7,96	
		ε= 10095 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 9970 M <sup>-1</sup> .cm <sup>-1</sup> (SH)			



**pET20b-AcLlp KD7-8AR -His**

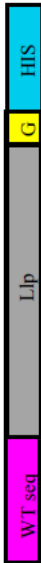
Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Ncbi : AY543070		Plasmide pEt20b (fuite promoteur) bacterial expression		Resistance Ampiciline
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 74 ng/µl V = 20 µl	A <sub>260/280</sub> 1,83
Caractéristiques de la construction				
<p>WT seq <span style="background-color: #FF00FF; padding: 2px;">C</span> <span style="background-color: #FF00FF; padding: 2px;">Lip</span> <span style="background-color: #FF00FF; padding: 2px;">G</span> <span style="background-color: #00FF00; padding: 2px;">His</span></p> <p>Séquence d'adressage Lip Wild type / Lip : WT partie mature / C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6* His.</p> <p>Objectifs : Mutant de Lip basé sur les données RMN pour investiguer l'interaction entre Lip et FhuA.</p>				
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Lip-2 (X049772) p175-179		Référence séquençage 50-604152729	
Clonage : Mutagenèse dirigée à partir de la construction pET20b-wt-Lip-G-His6 (ou pET20b-AcLlpWTHis) Séquence nucléique annotée :				
<p>TTTAAGAAGGAGATACATATGATAAATAATTTTACGCTGGCAGTGTACTTTCGGGCTGC TCTACTTTGGACCTGCAGCATTAAATGTGAGGCATATAATGAGGCATATAATGAAATATAAGGC TAATGTTTTGATAGGAAAGGGATATGTTTGTGTCACCAATATGGCTATGGATCTTCTGGG CCCCGTAAGTACTTTACCGGAAAGAAATACATGCGGAGGATTTTCGGCCATCACCATCAT GAGCCGACCGCTGCTGCTGG</p>				
Primer direct : 5' -> 3' KD7-8AR F GGCTTGCTACTTTGGACCTGCAGCATTAATGCTG AGGCATACATATG	Primer reverse : 5' -> 3' KD7-8AR R CATATAGTATGCTCAGCATTTAAATGGGT GCAGGTCAAAAGTAGAGCAAGCC		Taille construction : 3920	
Séquence protéique annotée : MKKLF <del>LA</del> MAV <del>VLL</del> SACSTFGPARIKCEAYYMQDHVKYKANVDFRKGDMFLVSPIMAYGSFWAIPVSYFT EGNTCEGVFGHHHHHH*				
Nbre AA avec seq signal = 84		MW=9573,12 Da		pI=8,5
Nbre AA sans seq signal = 69		MW=7956 Da		pI=7,21

**pET20b-AcLlp RK29-30SS-His**


Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Ncbi : AY543070		Plasmide pEt20b (fuite promoteur)		Resistance Ampiciline
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 121.6ng/µl V = 20 µl	A <sub>260/280</sub> 1,83
Caractéristiques de la construction				
<p>WT seq <span style="background-color: #FF00FF; padding: 2px;">C</span> <span style="background-color: #FF00FF; padding: 2px;">Lip</span> <span style="background-color: #FF00FF; padding: 2px;">G</span> <span style="background-color: #00FF00; padding: 2px;">His</span></p> <p>Séquence d'adressage Lip Wild type / Lip : WT partie mature / C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6* His.</p> <p>Objectifs : Mutant de Lip basé sur les données RMN pour investiguer l'interaction entre Lip et FhuA.</p>				
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Lip-2 (X049772) p175-179		Référence séquençage 50-604152729	
Clonage : Mutagenèse dirigée à partir de la construction pET20b-wt-Lip-G-His6 (ou pET20b-AcLlpWTHis) Séquence nucléique annotée :				
<p>TTTAAGAAGGAGATACATATGATAAATAATTTTACGCTGGCAGTGTACTTTCGGGCTGC TCTACTTTGGACCTAAAGATAATAATGTGAGGCATATAATGAGGCATATAATGAAATATAAGGC TATAAGGCTAATGTTTGTATAGCAGCGGGGATATGTTTGGTGTACCAATATAGGCCT ATGGATCTTTCTGGCCCCCGTAAGCTACTTTACCGAAGGAAATAATGCGGAGGAGGATTT TCGGCCATCACCATCACATCACATGAGCGACGCTGCTGCTGG</p>				
Forward Primer : 5' -> 3' RK29-30SS F GAAATATAAGGCTAATGTTTGTATAGCAGCAGCGG GGATATGTTTGTGTC	Reverse Primer : 5' -> 3' RK29-30SS R GACACCAAAAACATATCCCGCTGC TATCAAAAACATTAGCCTATATTTTC		Taille construction : 3920	
Séquence protéique annotée : MKKLF <del>LA</del> MAV <del>VLL</del> SACSTFGPKDIKCEAYYMQDHVKYKANVDFSSGDMFLVSPMAYGSF WAPVSYFTTEGNTCEGVFGHHHHHH*				
Nbre AA avec seq signal = 84		MW=9478,9 Da		pI=6,63
Nbre AA sans seq signal = 69		MW=7861,8 Da		pI=6,12



**pET20b-AcLlp YGSFW42-46STRAS -His**

Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Ncbi : AY543070		Plasmide pET20b (fuite promoteur)		Resistance Ampiciline
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 80,5 ng/µl V = 20 µl	A <sub>260/280</sub> 1,84
				
Caractéristiques de la construction				
Séquence d'adressage Llp Wild type / Llp : WT partie mature /C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6*His.				
Objectifs : Mutant de Llp basé sur les données RMN pour investiguer l'interaction entre Llp et FhuA.				
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Llp-2 (X049772) p175-179	Référence séquençage 50-604152729		
Clonage : Mutagenèse dirigée à partir de la construction pET20b-wt-Llp-G-His6 (ou pET20b-AcLlpWTHis)				
Séquence nucléique annotée : <i>tttaagaggagataatacatATGAAAAAATTTTAGCTATGGCAGTGGTACTCTTTTCGGGCTTGC</i> TCTACTTTTGGACCTAAAGATATAAATGTAGGCATCTATATGAGGATCATGTGAAA TATAAGGCTAATGTTTTGATAGGAAAGGGATATGTTTTGGTGTACCAATATGGCC <b>AGCACCGCGCCAGCGCC</b> CCCGTAAGCTACTTACCGAAGGAATACATGCGAGGGAGT TTTCGGCCATCACCATCACCATCACATGAGCCAGCCGCTGCTGCTGG				
Forward Primer : 5' -> 3' YGSFW42-46STRAS F <b>CGCCAGCGCC</b> CCCGTAAGCTACTTAC	Reverse Primer : 5' -> 3' YGSFW42-46STRAS R GTAAAGTAGCTTACGGGGCGCTGGCGCGG GTGCTGGCCATAATGGTGACACC	Taille construction : 3920		
Séquence protéique annotée : <b>MKKLFLAMAVLLS</b> ACSTFGPKDKICEAYYMQDHVKYKANVFDKRDGMFLVSPIMAS <b>STRAS</b> APVSYFTEGNTCEGVFGHHHHHHH*				
Nbre AA avec seq signal = 84	MW=9450,95Da	ε= 6085 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 5960 M <sup>-1</sup> .cm <sup>-1</sup> (SH)	MW=7833,8Da	pl=8,5 pl=7,21
Nbre AA sans seq signal = 69		ε= 6085 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 5960 M <sup>-1</sup> .cm <sup>-1</sup> (SH)		

**pET20b-AcLlp D28K-His**

Gène : Lipoprotéine du phage T5 Opéron seq ref EMBL/DDBJ : X75922 Ncbi : AY543070		Plasmide pET20b (fuite promoteur) bacterial expression		Resistance Ampiciline
Date 20/10/21	Boite stock Clone SDX plasmide	Clone x	Concentration C= 70,5 ng/µl V = 20 µl	A <sub>260/280</sub> 1,83
				
Caractéristiques de la construction				
Séquence d'adressage Llp Wild type / Llp : WT partie mature /C- présence de la cystéine +1 / G : glycine pour donner accès au tag His / His : tag 6*His.				
Objectifs : Mutant de Llp basé sur les données RMN pour investiguer l'interaction entre Llp et FhuA.				
Nom Degroux Séraphine	Référence cahier Labo T5/FhuA/Llp-2 (X049772) p175-179	Référence séquençage 50-604152729		
Clonage : Mutagenèse dirigée à partir de la construction pET20b-wt-Llp-G-His6 (ou pET20b-AcLlpWTHis)				
Séquence nucléique annotée : <i>tttaagaggagataatacatATGAAAAAATTTTAGCTATGGCAGTGGTACTCTTTTCGGGCTTGCCTCTAC</i> TTTTGGACCTAAAGATATAAATGTAGGCATCTATATGAGGATCATGTGAAAATAAAGGCTAATG TTTTTAAAGGAAAGGGATATGTTTTGGTGTACCAATATGGCCTATGGATCTTTCTGGGGCCCC GTAAGCTACTTTACCGAAGGAATACATGCGAGGGAGTTTT <b>CGCCATCACCATCACCATCAGCC</b> AGCC <i>gaccgctgctgctgg</i>				
Primer direct : 5' -> 3' D28K F GTGAAATATAAGGCTAATGTTTT <b>AAA</b> AGGAAAGGGGA TATGTTTTG	Primer reverse : 5' -> 3' D28K R CAAAAACATATCCCTTTCCTTTAAAA ACATTAGCCTTATATTCAC	Taille construction : 3920		
Séquence protéique annotée : <b>MKKLFLAMAVLLS</b> ACSTFGPKDKICEAYYMQDHVKYKANVFKRKGDMFLVSPIMAYGSFWAPVSYFT EGNTCEGVFGHHHHHHH*				
Nbre AA avec seq signal = 84	MW=9602 Da	ε= 13075 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 12950 M <sup>-1</sup> .cm <sup>-1</sup> (SH)	MW=7985 Da	pl=8,83 pl=7,95
Nbre AA sans seq signal = 69		ε= 13075 M <sup>-1</sup> .cm <sup>-1</sup> (SS) ε= 12950 M <sup>-1</sup> .cm <sup>-1</sup> (SH)		

















Annexe 2 : Table des entrées pour le calcul de structure et la caractérisation de la structure RMN minimisée en énergie pour Sol-Llp.

Quantity	Value
NOE-derived upper distance restraints	511
Intraresidual ( $ i - j  = 0$ )	130
Short-range ( $ i - j  = 1$ )	140
Medium-range ( $2 \leq  i - j  \leq 4$ )	62
Long-range ( $ i - j  \geq 5$ )	179
Dihedral angle constraints	106
Residual target function <sup>(b)</sup> , Å <sup>2</sup>	1.53 ± 0.37
Residual NOE violations	
Number > 0.2 Å	0
Residual angle violations	
Number > 5 °	0
Amber energies, kcal/mol	
Total	-1909.7 ± 69.3
Van der Waals	-215.5 ± 12.5
Electrostatic	-2268.3 ± 69.5
RMSD from ideal geometry	
Bond lengths, Å	0.0158 ± 0.0007
Bond angles, °	1.80 ± 0.09
RMSD to the mean coordinates <sup>(c)</sup> , Å	
bb	1.02 ± 0.19
ha	1.58 ± 0.2
Ramachandran plot statistics <sup>(d)</sup>	
Most favored regions (%)	66.7
Additionally allowed regions (%)	28.5
Generously allowed regions (%)	3.1
Disallowed regions (%)	1.7

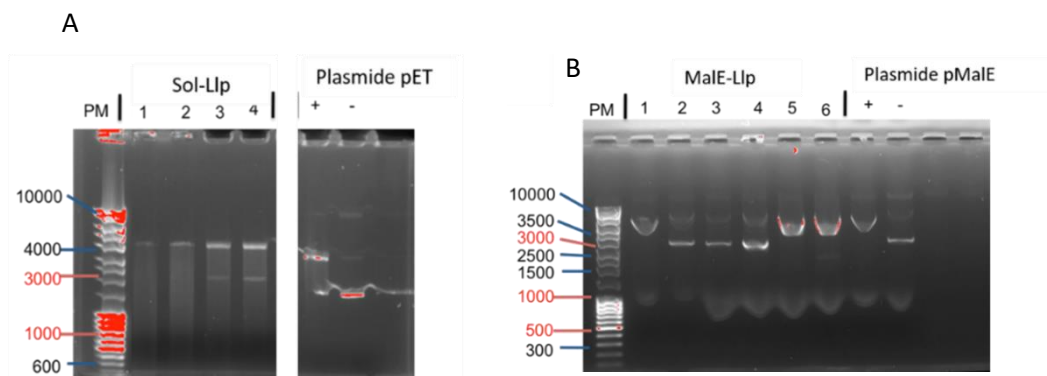


- (a) Number of meaningful, non-redundant upper distance restraints between residues with sequence positions  $i$  and  $j$ .
- (b) Average residual target function value of all 20 structures calculated by CYANA.
- (c) 'bb' indicates backbone atoms N, C, C $\alpha$  and C', 'ha' indicates 'all heavy atoms'. Root-mean square deviations (RMSD) are reported for the well-defined structural regions of the protein, which include residues 6–38 and 45–62.
- (d) Ramachandran plot statistics were determined using the program PROCHECK (Laskowski et al. 1993).

### Annexe 3 : vérification de la présence de l'insert dans les clonages.

Le gène *Llp* a été cloné au niveau du MCS (multiples cloning site) contenant en outre BamHI dans le plasmide pET20b ou pMAL-p2E, car les amorces pour l'amplification sont placées en aval et amont de ce site (voir matériel et méthode). La présence de l'insert est donc vérifiée par digestion enzymatique (BamHI) avant l'envoi en séquençage : la taille des plasmides digérés avec l'enzyme BamHI est comparés au plasmide d'origine digéré (linéaire de taille attendue 3716 pb pour pET20b et 6724 pb pour pMAL-p2E) ou non (sur-enroulé, taille attendue plus faible avec présence possible de différente conformation).

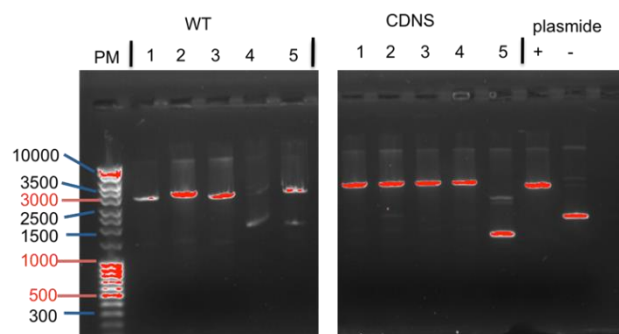
Pour la protéine soluble les clones N° 3 et 4 contiennent l'insert du gène *Llp*, car le profile sur gel est identique témoin du plasmide non digéré. De même, le clone N° 2, 3 et 4 pour la fusion du gène MalE avec *Llp* contient le gène d'intérêt, car ils ne sont pas digérés en présence de BamHI (Figure 7-1 -B). Le séquençage pour les plasmides Sol-Llp et MalE-Llp a confirmé l'absence de mutation dans la séquence codant pour la protéine.



**Figure 7-0-1 : Digestion enzymatique par BamHI des constructions Sol-Llp (A) et MalE-Llp (B).**

PM poids moléculaires, numérotation de 1 à 4/6 : numérotation arbitraire des clones présents sur la boîte de Pétri après transformation pour les constructions Sol Llp et MalE-Llp. La migration des plasmides vides digérés (+) ou non (-) par BamHI a été contrôlée dans les deux dernières pistes des gels. Taille linéaire attendu : 3716 pb pour pET20b et 6724 pb pour pMAL-p2E

Pour Ac-Llp le clones N° 4 contient l'insert du gène *Llp*, car le profile sur gel est identique témoin du plasmide non digéré avec des formes sur-enroulées (figure 7-2 WT). De même, le clone N°5 pour PuIA-Ac-Llp contient le gène d'intérêt, car ils ne sont pas digérés en présence de BamHI (Figure 7-2 CDNS). Le séquençage pour le plasmide Ac-Llp a confirmé l'absence de mutation dans le gène de Llp. Le clonage des constructions PuIA-Ac-Llp et de l'opéron ont échoué et le choix a été fait de ne pas poursuivre les efforts de clonage sur ces deux constructions.



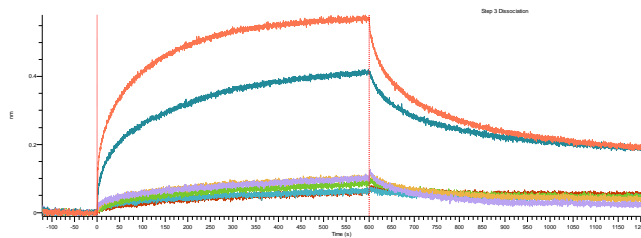
**Figure 7-2: Digestion enzymatique par BamHI des constructions Ac-Llp (WT), et PuIA-Ac-Llp (CDNS).**

PM : poids moléculaires, numérotation de 1 à 4/5 : numérotation arbitraire des clones présents sur la boîte de Petri après transformation. La migration des plasmides vides digérés (+) ou non (-) par BamHI a été contrôlée dans les deux dernières pistes des gels.

## Annexe 4 : Test d'interaction BLI et Fidabio

### « Bilayer Interferometry » BLI

La technique de BLI (Interférométrie de bicouches) permet également de détecter des interactions entre protéines. Comme la SPR elle permet d'enregistrer des sensorgrammes sur des gammes de concentration de protéine et de calculer des  $K_d$ . Cependant, dans ce système, le flux de protéine n'est pas assuré par de la microfluidique, mais par le mouvement du liquide dans des puits sous agitations. Nous avons donc testé cette technique, car le système de détection et de flux étant différent nous aurions pu obtenir de meilleur résultat. Comme dans le cas de SPR nous avons fixé FhuA sur les puces puis incubé avec Ac-Llp. Nous notons une interaction entre FhuA et Ac-Llp, mais avec des sauts de signal pour les plus fortes concentrations (Figure 7-3). Le  $K_d$  n'a pas été calculé. Dans le cas de Sol-Llp aucune interaction n'a pu être détectée avec FhuA (non montré). Les résultats préliminaires étant peu satisfaisant nous n'avons pas investigué plus longuement l'interaction à l'aide de cette technique.

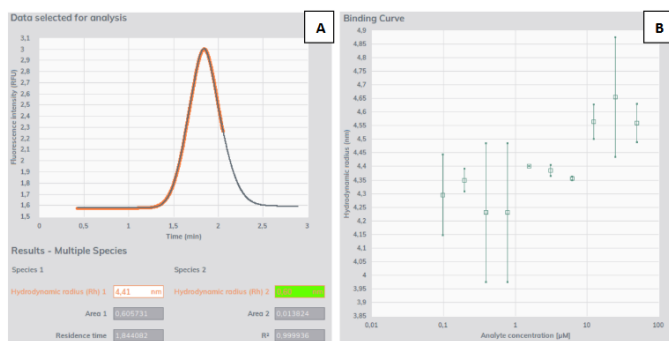


**Figure 7-0-3: Sensorgramme de l'interaction entre Ac-Llp et FhuA immobilisé, en DDM.**

Chaque sensorgramme représente la réponse enregistrée pour une puce normalisée par la réponse obtenue avec la BSA.

### Fidabio

La méthode fidabio est basée sur l'analyse de la dispersion des protéines dans un capillaire soumis à un flux. Le profil parabolique hydrodynamique est analysé pour définir la taille des particules. Selon ce principe, la formation du complexe doit être visible par l'augmentation de taille des particules. Pour une analyse de  $K_d$ , un des deux partenaires doit être marqué par le fluorophore FITC absorbant à 490 nm. Ce type de fluorophore étant très hydrophobe, il pose souvent des problèmes en se dissolvant dans les micelles de détergent. Un tamis moléculaire peut en général séparer les protéines des micelles libres, mais ce n'est pas le cas pour Ac-Llp qui est trop petite et co-migre avec les micelles libres. FhuA a donc été marquée et purifiée sur tamis moléculaire. Le signal de FhuA été très bon avec très peu le fluorophore libre (Figure 7-4 - A). Le complexe FhuA-détergent apparaît comme une particule de Rh (rayon hydrodynamique) de 4,41 nm, ce qui est cohérent avec les données d'AUC et de SEC-MALS. En revanche, aucune différence de Rh significative n'a pu être observée pour une gamme de Ac-Llp de 0,1 à 50  $\mu$ M.

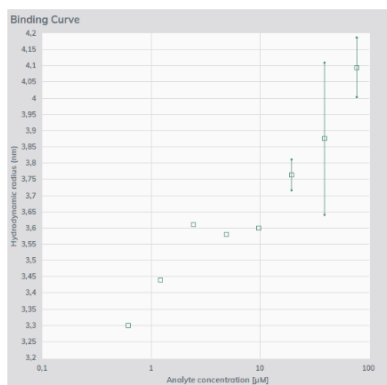


**Figure 7-0-4 :** Courbe d'affinité de Ac-Llp pour FhuA par Fidabio.

Ajustement A. signal brut de FhuA à 490 nm à une concentration de 100 nM. B. changement de Rh de FhuA à 490 nm en fonction d'une gamme de Ac-Llp de 0,1 à 50 μM.

Ac-Llp est une petite protéine, et la formation du complexe peut engendrer une réorganisation de la micelle autour de FhuA. Ces deux éléments peuvent expliquer qu'on ne voit pas de changement de taille au cours de la titration. Pour tester notre hypothèse, nous avons marqué Ac-Llp malheureusement nous n'avons pas réussi à retirer le fluorophore libre ce qui n'a pas permis d'enregistrer une courbe de titration.

Contenant une quantité donnée de détergent. Étonnamment, le kit de nettoyage du fluorophore a été efficace et peu de fluorophore libre était présent permettant l'analyse ! Le complexe Ac-Llp-détergent a une taille apparente de 3,3 nm. Les expériences semblent montrer qu'il y a une augmentation de taille pour une gamme de FhuA de 78 à 0,61 μM (Figure 7-5). Malheureusement, un problème technique n'a pas permis d'enregistrer l'ensemble des points en triplicata ne permettant pas le calcul d'un Kd.



**Figure 7-5: Courbe d'affinité de Ac-Llp pour FhuA.**

Changement de Rh de Ac-Llp à 490 nm en fonction d'une gamme de FhuA de 0,61 à 78 μM.

Une nouvelle campagne de mesure a été réalisée avec un échantillon de Ac-Llp qui avait été concentré. Cette fois le kit de lavage du fluorophore n'a pas été efficace, et le fluorophore libre majoritaire a masqué le signal de Ac-Llp. La taille du fluorophore libre semble augmenter avec la titration de FhuA ce qui suggère que le fluorophore libre s'associe aux micelles de détergent présent qui ne contiennent pas de Llp. Ce résultat montre que dans le cas des protéines membranaires, le fluorophore libre doit être minimisé à tout prix. Cette nouvelle série de mesure n'a pas permis d'obtenir de Kd pour le complexe FhuA-Ac-Llp. Il serait certainement possible d'obtenir un Kd pour ce complexe par cette méthode, mais une optimisation du marquage est nécessaire. Une possibilité pour enlever tout le fluorophore libre de Ac-Llp serait de refaire une colonne NiNTA après le marquage.



# Résumés

Les bactériophages ou phages sont des virus qui infectent les bactéries. Ils représentent l'entité biologiques les plus abondantes de la planète et sont présents dans tous les biotopes. Le phage T5 est un phage strictement lytique qui appartient à la famille des *Siphoviridae* et qui infecte *E. coli*. L'infection par le phage T5 est déclenchée par la reconnaissance de l'hôte grâce à pb5, la protéine de liaison au récepteur (RBP), qui se lie à FhuA, transporteur de la membrane externe de *E. coli*. Cette interaction déclenche la perforation de la paroi bactérienne et l'injection de l'ADN viral dans le cytoplasme de l'hôte. Cette première étape de l'infection est suivie par la réplication virale puis la libération des nouveaux virions. Pendant cette période de vulnérabilité, T5 protège la nouvelle usine virale d'une surinfection par la production d'une lipoprotéine périplasmique, Llp, qui cible le feuillet interne de la membrane externe, se lie à FhuA et l'inactive. La principale fonction biologique de Llp est probablement d'empêcher l'inactivation des virions par les FhuA présente dans les débris des cellules lysées, augmentant ainsi leurs chances d'infecter un nouvel hôte. Pendant ma thèse, nous avons résolu la structure du complexe FhuA-pb5 par microscopie électronique à une résolution de 2,4 Å. La structure de pb5 partage des homologies avec d'autres protéines de phage dans sa partie proximale seulement. La structure du complexe FhuA-pb5 montre que l'interaction entre FhuA et pb5 est principalement due à un réseau de liaisons hydrophobes. Nous avons également résolu la structure de Llp par RMN pour la forme soluble (Sol-Llp), révélant un repliement principalement en feuillet bêta avec une boucle flexible. Cette structure est une nouvelle structure de protéine de phage. De plus, j'ai étudié le complexe FhuA-Llp avec diverses méthodes biochimiques et biophysiques permettant de définir un  $K_D$  de 1,5 mM pour Sol-Llp et une estimation de  $K_D$  pour la protéine acylé de 20  $\mu$ M, ce qui implique de l'acylation joue un rôle important dans la formation du complexe. La titration de Sol-Llp par FhuA en RMN montre des perturbations d'intensité et de déplacement chimique pour les résidus 12 à 24, et 51 à 54. Des mutants de ces résidus ainsi que des résidus 39 à 45 - non accessibles en RMN - confirment l'importance de ces résidus dans l'interaction avec FhuA, les protéines mutantes étant incapables de protéger la cellule d'une infection à T5 contrairement à la protéine sauvage.

# Abstract

Bacteriophages or phages are viruses that infect bacteria. They are the most abundant biological entity on the earth and are present in all biotopes. Phage T5 is a strictly lytic phage that belongs to the family Siphoviridae and infects *E. coli*. Infection by the T5 phage is triggered by host recognition through the receptor binding protein (RBP) pb5, which binds to the *E. coli* outer membrane transporter FhuA. This interaction triggers the perforation of the bacterial wall and the injection of viral DNA into the host cytoplasm. This first stage of infection is followed by viral replication and the release of new virions. During this vulnerable period, T5 protects the new viral factory from superinfection by producing a periplasmic lipoprotein, Llp, which targets the inner leaflet of the outer membrane, binds to and inactivates FhuA. The main biological function of Llp is probably to prevent the inactivation of virions by FhuA present in the debris of lysed cells, thus increasing their chances of infecting a new host. During my thesis, we solved the structure of the FhuA-pb5 complex by electron microscopy at a resolution of 2.4 Å. The structure of pb5 shares homologies with other phage proteins only in its proximal part. The structure of the FhuA-pb5 complex shows that the interaction between FhuA and pb5 is mainly due to a network of hydrophobic bonds. We also solved the structure of Llp by NMR for the soluble form (Sol-Llp), revealing a predominantly beta-sheet folding with a flexible loop. This structure is a novel phage protein structure. In addition, I studied the FhuA-Llp complex with various biochemical and biophysical methods to define a  $K_D$  of 1.5 mM for Sol-Llp and an estimated  $K_D$  for the acylated protein of 20  $\mu$ M, which implies that acylation plays an important role in the formation of the complex. NMR titration of Sol-Llp by FhuA shows intensity and chemical shift perturbations for residues 12-24, and 51-54. Mutants of these residues as well as residues 39-45 - not accessible by NMR - confirm the importance of these residues in the interaction with FhuA, with the mutant proteins being unable to protect the cell from T5 infection in contrast to the wild type.