



# Etudes structurales et fonctionnelles de lectines et d'adhésines chez *Pseudomonas aeruginosa*

Géraldine Ganne

► **To cite this version:**

Géraldine Ganne. Etudes structurales et fonctionnelles de lectines et d'adhésines chez *Pseudomonas aeruginosa*. Sciences agricoles. Université de Grenoble, 2013. Français. <NNT : 2013GRENV005>. <tel-00949168>

**HAL Id: tel-00949168**

**<https://tel.archives-ouvertes.fr/tel-00949168>**

Submitted on 19 Feb 2014

**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 DE GRENOBLE

Spécialité : Biologie Structurale et Nanobiologie

Arrêté ministériel : 7 août 2006

Présentée par

**Géraldine GANNE**

Thèse dirigée par **Annabelle VARROT**

préparée au sein du **Laboratoire CERMAV (Centre de Recherche sur les MAcromolécules Végétales)**

dans l'**École Doctorale de Chimie et Sciences du Vivant**

**Etudes structurales et fonctionnelles  
de lectines et d'adhésines chez  
*Pseudomonas aeruginosa***

Thèse soutenue publiquement le 1 Février 2013  
devant le jury composé de :

**Andrea DESSEN**

Directeur de recherche, Grenoble (Présidente)

**Isabelle CALLEBAUT**

Directeur de recherche, Paris (Rapporteur)

**Julie BOUCKAERT**

Chargée de recherche, Lille (Rapporteur)

**Sophie de BENTZMANN**

Directeur de recherche, Marseille (Examineur)

**Annabelle VARROT**

Chargée de recherche, Grenoble (Directeur de thèse)







*A ma mère*



*Je remercie très sincèrement le Dr Annabelle Varrot qui a été ma directrice de thèse ainsi que le Dr Anne Imberty, directrice de l'équipe Glycobiologie Moléculaire, pour m'avoir offert l'opportunité de réaliser ma thèse sur ce sujet passionnant et dans des conditions de travail exceptionnelles.*

*Je remercie le Dr Julie Bouckaert et Dr Isabelle Callebaut d'avoir accepté de juger mes travaux de thèse.*

*Je remercie également les membres de mon comité de thèse, le Dr Andrea Dessen et le Dr Sophie de Bentzmann, pour les précieux conseils qui m'ont beaucoup aidés durant ces trois années.*

*Un grand merci à toutes les personnes de l'équipe Glycobiologie Moléculaire. En particulier à Valérie Chazalet, à Emilie Gillon, Julie Arnaud, au Dr Aline Thomas, au Dr Joana Rocha et au Dr Aymeric Audfray, qui ont toujours été présents lorsque j'avais des questions ou pour m'aider à me former à de nouvelles techniques.*

*Je tiens également à remercier le Dr Julien Pérard, le Dr Jens Radzimanowski et le Dr Gianluca Cioci, pour m'avoir aidé à traiter mes données de SAXS et d'avoir pris de leur temps pour répondre à mes questions.*

*Je tiens à remercier toutes les personnes du CERMAV que j'ai eu l'occasion de croiser et notamment le personnel administratif qui m'a simplifié la vie pour les ordres de missions.*

*C'est avec un peu de nostalgie que je tiens à remercier Agathe Belime, Quentin Berrod, Matthieu Fumagalli, Jeanne Bernard, Pan Chen, Emilie Hachet, Antoine Bailly, Antoine Duval, Bruno Frka-Petesic, Axel Ettore, Clélia Martin, Deborah Blanchard et Jimmy Mergy pour tous les bons moments qu'on a partagé durant ces trois années et pour tous ce que vous m'avez apporté.*

*Un très grand merci à Marion Burre-Espagnou, Emilie Gillon et Firas Azzam pour cette belle amitié qui est née au CERMAV.*

*Je remercie toute ma famille (ma mère en particulier, ma grand-mère, ma grande sœur et Cédric, mon grand frère et Julia sans oublier Mael, puis ma tante Guenaelle et ses deux enfants Erwan et Maëlwenn) pour leur soutien.*

*Un très grand merci à Aurélien pour tout ce que tu m'offres.*

*Ma dernière pensée sera pour toutes les personnes atteintes de la mucoviscidose et qui se battent au quotidien pour lutter contre cette maladie.*



<b>LISTE DES ABRÉVIATIONS.....</b>	<b>0</b>
------------------------------------	----------

<b>INTRODUCTION.....</b>	<b>1</b>
--------------------------	----------

1. <i>PSEUDOMONAS AERUGINOSA</i> , UN PATHOGENE DIFFICILE A ERADIQUER .....	1
1.1. <i>Pseudomonas aeruginosa</i> .....	1
1.1.1. Description.....	1
1.1.2. L'apport génomique chez <i>P. aeruginosa</i> .....	1
1.1.3. Habitat et pouvoir pathogène .....	2
1.2. <i>La mucoviscidose</i> .....	3
1.2.1. Une maladie génétique.....	3
1.2.2. Les infections bactériennes .....	4
1.3. <i>Une multi-résistance aux antibiotiques</i> .....	6
1.4. <i>Le biofilm</i> .....	8
1.4.1. La matrice extracellulaire.....	8
1.4.2. La formation du biofilm.....	8
1.4.3. Le biofilm, une forme de résistance .....	11
1.4.4. Perception des signaux environnementaux et régulation.....	11
1.4.4.1. Le c-di-GMP .....	12
1.4.4.2. Les systèmes à deux composants .....	13
2. LA GLYCO-STRATEGIE DE <i>PSEUDOMONAS AERUGINOSA</i> .....	15
2.1. <i>La glycosylation chez l'homme</i> .....	15
2.1.1. Les glycoprotéines .....	15
2.1.1.1. La <i>N</i> -glycosylation.....	16
2.1.1.2. La <i>O</i> -glycosylation.....	17
2.1.1.2.1. La <i>O</i> -glycosylation de type mucine.....	17
2.1.2. Les protéoglycanes.....	18
2.1.3. Les glycolipides .....	19
2.1.4. Les antigènes des groupes sanguins .....	20
2.1.5. Modification de la glycosylation chez les patients CF.....	21
2.2. <i>Les lectines</i> .....	22
2.2.1. Les lectines flagellaires.....	23
2.2.1.1. Un des rôles du flagelle : initiation de l'adhésion .....	23
2.2.1.2. Structure du flagelle .....	24
2.2.1.3. Les lectines FliD et FliC.....	25
2.2.2. Les lectines solubles .....	25
2.2.2.1. La lectine PA-IL.....	25
2.2.2.1.1. Le rôle de PA-IL.....	25
2.2.2.1.2. La spécificité de PA-IL .....	26
2.2.2.1.3. PA-IL : $\beta$ -sandwich à un calcium .....	26
2.2.2.2. La lectine PA-III .....	27
2.2.2.2.1. Le rôle de PA-III .....	27
2.2.2.2.2. Spécificité de PA-III.....	28
2.2.2.2.3. PA-III : $\beta$ -sandwich à deux calciums .....	28
2.2.3. Les lectines fimbriales .....	29
2.2.3.1. Les pili de type IV.....	29
2.2.3.1.1. Le rôle des pili de type IV .....	29
2.2.3.1.2. L'assemblage des pili de type IV.....	29
2.2.3.1.3. La lectine PilA.....	30
2.2.3.2. Les pili de type « CU » .....	31
3. LA VOIE « CHAPERONNE-USHER » (CU).....	32
3.1. <i>La classification des voies CU</i> .....	33
3.2. <i>Les pili de type P et type I chez Escherichia coli</i> .....	34
3.2.1. Description des pili de type I et P .....	35
3.2.2. Les composants de la voie « chaperonne usher ».....	36
3.2.2.1. Les sous-unités structurales.....	36
3.2.2.1.1. Les pilines majeures : le corps rigide du pilus.....	36

3.2.2.1.2. le fimbrillum du pilus: les pilines mineures.....	37
3.2.2.1.3. le fimbrillum du pilus: l'adhésine.....	38
3.2.2.2. La chaperonne périplasmique.....	40
3.2.2.3. La complémentation par brin donneur avec les chaperonnes FGS.....	41
3.2.2.4. L'échange de brin donneur.....	42
3.2.2.5. La protéine de membrane externe « Usher ».....	44
3.2.2.6. Modèle d'assemblage.....	47
3.3. Les voies CUP chez <i>Pseudomonas aeruginosa</i> .....	48
3.3.1. CupA.....	50
3.3.1.1. Le système CupA.....	50
3.3.1.2. La régulation du cluster <i>cupA</i> .....	50
3.3.2. CupB et CupC.....	52
3.3.2.1. Les systèmes CupB et CupC.....	52
3.3.2.2. La régulation des clusters <i>cupB</i> et <i>cupC</i> .....	54
3.3.3. CupE.....	55
3.3.3.1. Le système CupE.....	55
3.3.3.2. La régulation du cluster <i>cupE</i> .....	56
4. VERS LA THERAPIE ANTI-ADHESIVE.....	58
4.1. Relation entre lectines et pathogénicité de <i>P. aeruginosa</i> .....	58
4.2. Les glycomimétiques.....	59
<b>MATÉRIELS ET MÉTHODES .....</b>	<b>63</b>
5. TECHNIQUES DE BIOLOGIE MOLÉCULAIRE.....	63
5.1. Obtention des différentes constructions plasmidiques.....	63
5.1.1. Clonage des gènes <i>cupE6</i> , <i>cupE4</i> , <i>cupB6</i> et <i>birA</i> .....	63
5.1.1.1. Clonage du domaine adhésine de <i>cupE6</i> et <i>cupB6</i> .....	63
5.1.1.2. Clonage du gène <i>cupE6</i> et <i>cupE4</i> dans pETDuet.....	63
5.1.1.3. Clonage du gène <i>birA</i> .....	63
5.1.2. Le plasmide pETDuet-TEV-Biot.....	64
5.1.3. Tests d'expression.....	65
6. CARACTERISATION BIOCHIMIQUE ET BIOPHYSIQUE.....	68
6.1. Purification des protéines recombinantes.....	68
6.1.1. Préparation des extraits bruts.....	68
6.1.2. Purification du domaine adhésine CupB6.....	68
6.1.2.1. Chromatographie sur résine de Nickel ou de Zinc.....	68
6.1.2.2. Chromatographie sur des billes de glutathion.....	68
6.1.2.3. Chromatographie sur Amylose.....	69
6.1.2.4. Clivage de l'étiquette par la TEV.....	69
6.1.3. Purification du complexe CupB6/CupB4.....	69
6.1.4. Purification de la lectine PA-IL.....	69
6.1.5. Purification de la lectine PELa.....	70
6.2. Détermination des domaines protéiques stables de CupB6/B4.....	70
6.3. Diffusion dynamique de la lumière (DLS).....	71
6.4. Thermal Shift Assay (TSA).....	71
6.4.1. Principe de la méthode Thermal Shift Assay.....	71
6.4.2. Protocole permettant l'optimisation des conditions de cristallisation par TSA.....	73
6.5. La diffusion des rayons X aux petits angles (SAXS pour Small-angle X-ray scattering).....	73
6.5.1. Introduction au SAXS.....	73
6.5.2. Paramètres structuraux globaux.....	75
6.5.2.1. La relation de Guinier et le rayon de giration.....	75
6.5.2.2. Détermination de la masse moléculaire.....	75
6.5.2.3. Fonction de distribution des distances $P(r)$ .....	76
6.5.3. Modélisation ab initio.....	77
6.5.4. Etude de la flexibilité par la méthode EOM (Ensemble Optimization Method).....	78
6.5.5. Collecte des données pour le complexe CupB6/B4.....	79
7. OUTILS POUR L'ETUDE DES INTERACTIONS PROTEINES-SUCRE.....	80
7.1. Enzyme-Linked Lectin Assay : ELLA.....	80

7.1.1. Principe du test ELLA .....	80
7.1.2. Test ELLA avec la lectine PA-IL.....	81
7.2. <i>Microcalorimétrie de titration isotherme (ITC)</i> .....	81
7.2.1. Notion de thermodynamique.....	81
7.2.2. Principe de l'ITC.....	82
7.3. <i>La résonance plasmonique de surface (SPR)</i> .....	84
7.3.1. Le principe de la résonance plasmonique de surface .....	84
7.3.2. Etude de la spécificité de CupB6/B4 avec les puces CM5.....	85
7.3.3. Etude de la spécificité de CupB6/B4 avec les puces NTA.....	86
7.4. <i>La cristallographie aux rayons X</i> .....	86
7.4.1. La cristallogénèse.....	86
7.4.1.1. Les paramètres physico-chimiques et les principaux agents précipitants .....	87
7.4.1.2. La cristallisation par diffusion de vapeur .....	88
7.4.1.3. Optimisation.....	88
7.4.1.4. Collecte des données expérimentales .....	90
7.4.2. La cristallographie .....	91
7.4.2.1. Définition d'un cristal .....	91
7.4.2.2. Principe de la diffraction aux rayons X.....	92
7.4.2.3. Indexation, intégration et mise à l'échelle du jeu de données .....	93
7.4.2.4. Facteurs de structure et densité électronique.....	94
7.4.2.5. Le remplacement moléculaire .....	94
7.4.2.6. Affinement des structures.....	95
<b>RÉSULTATS ET DISCUSSION .....</b>	<b>97</b>
8. MISE EN APPLICATION DES METHODES QUI PERMETTENT D'ETUDIER LES INTERACTIONS LECTINE-SUCRE.....	97
8.1. <i>Introduction</i> .....	97
8.2. <i>Production de PELa sous forme recombinante</i> .....	98
8.3. <i>Etude de la spécificité de PELa</i> .....	99
8.3.1. Résultat de l'analyse par puces à sucres.....	99
8.3.2. Etude de la spécificité de PELa par ITC .....	100
8.4. <i>Etudes par cristallographie des interactions mises en place entre PELa et son ligand</i> .....	102
8.4.1. Structure du complexe Pela/composé H.....	102
8.4.2. Essais de cristallogénèse du complexe PELa/composé E.....	106
8.5. <i>Conclusion</i> .....	107
9. A LA RECHERCHE D'INHIBITEURS DE PA-IL DERIVES DU GALACTOSE .....	108
9.1. <i>Introduction</i> .....	108
9.2. <i>Criblage par test ELLA d'inhibiteurs dérivés du galactose</i> .....	110
9.3. <i>Etude par ITC</i> .....	113
9.4. <i>Conclusion</i> .....	114
10. LES ADHESINES CUPE6 ET CUPB6 .....	115
10.1. <i>L'adhésine CupE6</i> .....	116
10.1.1. Expression du domaine adhésine de CupE6.....	116
10.1.1.1. Délimitation du domaine adhésine par analyse de séquence .....	116
10.1.1.2. Tests d'expression du domaine adhésine CupE6.....	118
10.2. <i>Expression du complexe CupE6/E4</i> .....	121
10.3. <i>Conclusion</i> .....	122
10.4. <i>L'adhésine CupB6</i> .....	123
10.4.1. Expression du domaine CupB6.....	123
10.4.1.1. Analyse de séquence pour définir le domaine adhésine .....	123
10.4.1.2. Test d'expression et de purification .....	125
10.5. <i>Le complexe CupB6/B4</i> .....	129
10.5.1. Optimisation de la purification de CupB6/B4 .....	129
10.5.2. Etude de la spécificité de CupB6 pour les oligosaccharides .....	131
10.5.2.1. Identification de ligands saccharidiques par puces à sucres .....	131
10.5.2.2. Etude de la spécificité de CupB6 par SPR .....	132
10.5.2.3. Création d'un tétramère artificiel de CupB6/B4.....	136

10.5.3. Cristallogenèse du complexe CupB6/B4.....	138
10.5.3.1. Optimisation des conditions de cristallisation par TSA.....	140
10.5.3.2. Optimisation des conditions de cristallisations.....	142
10.5.3.3. La protéolyse ménagée de CupB6/B4 .....	145
10.5.4. Caractérisation du complexe CupB4/CupB6 par SAXS .....	148
10.5.4.1. Evaluation et traitement des données collectées.....	148
10.5.4.2. La modélisation ab initio.....	150
10.5.4.3. Etude de la flexibilité par EOM .....	152
10.6. Conclusion.....	153
<b>CONCLUSIONS ET PERSPECTIVES .....</b>	<b>155</b>
11. CONCLUSIONS ET PERSPECTIVES.....	155
<b>RÉFÉRENCES BIBLIOGRAPHIQUES .....</b>	<b>161</b>
<b>ANNEXE.....</b>	<b>174</b>

## Liste des abréviations

**$\alpha$ MeGal:**  $\alpha$ -D-méthyl-galactoside

**ABC transporteurs:** ATP-Binding Cassette

**ASL:** Airway Surface Liquid

**BgA:** Blood group A antigen

**BIMPs:** Bitopic Inner Membrane Proteins

**BSA:** Bovine Serum Albumin

**Cftr:** Cystic Fibrosis Transmembrane conductance Regulator)

**CCD:** Couple Charge Device

**c-di-GMP:** di-guanilate monophosphate cyclique

**CF:** patients atteints de la mucoviscidose

**Cgr:** *cup* gene regulator

**CRDs:** Carbohydrate-Recognition Domains

**CSS:** Clear Strategy Screen

**CS:** Chondroïtines sulfates

**CTD:** domaine C-terminal

**CU:** Chaperonne-Usher

**DLS:** Dynamic light Scattering

**DS:** dermatanes sulfates

**DSL:** disulfide-bonded loop

**EAL:** domaine à activité phosphodiesterase

**ELLA:** Enzyme-Linked-Lectin Assay

**EOM:** Ensemble Optimization Method

**EPS:** exopolysaccharides

**ESRF:** European Synchrotron Radiation Facility

**FHAs:** Filamentous hemagglutinins

**FGL:** chaperonnes à boucle F1-G1 long

**FGS:** chaperonnes à boucle F1-G1 court

**Fuc:** Fucose

**GAGs:** glycosaminoglycanes

**Gal:** galactose

**GalNAc:** N-acétyl-galactosamine

**Gal $\alpha$ 1-6Glc:** mélibiose

**GGDEF:** domaine à activité diguanylate cyclase

**GlcNAc:** N-acétylglucosamine

**GlcN:** glucosamine

**GM3:** Ganglioside sugar 3'Sialyllactose

**GST:** Glutathione S-Transferase

**GyrA:** ADN gyrase

**HCA:** Hydrophobic Cluster Analysis

**H-NS:** Heat-stable nucleoid structuring

**HPt:** Histidine Phospho-Transfert

**HS:** Héparane-sulfates

**IMAC:** Immobilized Metal Affinity Chromatography

**IPTG:** Iso-Propyl- $\beta$ -d-Thiogalactopyranoside

**ITC:** Isotherm Titration Calorimetry

**LacNAc:** N-acetyl-lactosamine

**Man:** Mannose

**MRSA:** Multi Resistant *S. aureus*

**MSD:** Membrane Spanning Domain

**N<sub>A</sub>:** nombre d'Avogadro

**NBD:** Nucleotide Binding Domain

**NSD:** Normalized Spatial Discrepancy

**NTD:** domaine N-terminal

**OM:** membrane externe

**ONM:** l'Observatoire National de la Mucoviscidose

**OPD:** *ortho*-phenylènediamine

**ORF:** Open Reading Frame

**P:** périplasme

**Pap:** pilus-associated with Pyelonephritis

**PBS:** Phosphate Buffer Saline

**PCL:** Periciliary Liquid Layer

**PCR:** Polymerase Chain Reaction

**PNP-  $\beta$ -Gal:** *p*-nitrophenyl- $\beta$ -galactoside

**POTRA:** Polypeptide-Transport-Associated

**Psl:** Polysaccharides Synthesis Locus

**P-usher:** POTRA-like domain-containing Usher

**RD:** Regulation Domain

**RFM:** Registre Français de la Mucoviscidose

**Rg:** le rayon de giration

**RND:** Resistance-Nodulation-Division

**Roc:** Regulation of cup

**RU:** Response Units

**SAXS:** Small Angle X ray Scattering

**SCV:** Small Colony Variant

**SIDA:** Syndrome de l'ImmunoDéficiency Acquis

**SIGNs:** Specific Intercellular adhesion molecule 3 (ICAM-3)-Grabbing Nonintegrin

**SPCU:** Spore Coat Protein U

**SNC:** Symétrie Non Cristallographique

**SPR:** Surface Plasmon Resonance

**TCS:** systèmes à deux composants

**TEV:** Tabacco Etch Virus

**Thr:** thréonine

**TSA:** Thermal Shift Assay

**VNTR:** Variable Number of Tandem Region



# **Introduction**



---

# 1. *Pseudomonas aeruginosa*, un pathogène difficile à éradiquer

---

## 1.1. *Pseudomonas aeruginosa*

### 1.1.1. Description

*Pseudomonas aeruginosa* (« *pseudo* » signifie imitation, « *monas* » unité et « *Aeruginosa* » signifie vert- de gris qui est la couleur d'un des pigments sécrétés par la bactérie) a été étudié pour la première fois en 1882 par Carle Gessard. Ce bacille à Gram négatif appartient à la famille des *Pseudomonadaceae* et à la classe des  $\gamma$ -protéobactéries. C'est une bactérie aérobie stricte mais capable de se développer en anaérobiose. *P. aeruginosa* est une bactérie saprophyte qui vit dans les milieux aquatiques, les milieux terrestres, à la surface des végétaux et à l'état commensal dans l'appareil digestif de l'homme et l'animal. Sa température optimale de croissance est comprise entre 30 et 37°C mais elle peut résister à des températures pouvant aller jusqu'à 42°C.

*P. aeruginosa* peut présenter trois morphotypes différents : (1) des colonies rugueuses, larges et légèrement bombées (souche parentale), (2) des petites colonies rugueuses appelées « Small Colony Variant » (SCV) présentant un comportement autoagrégatif (Häussler et al., 2003), puis (3) des colonies « mucoïdes » de forme bombées, opaques, et visqueuses (Mathee et al., 1999). Les colonies « SCV » et « mucoïdes » sont associées à un état de biofilm, mais *P. aeruginosa* peut également exister sous forme planctonique présentant une mobilité par le biais d'un flagelle polaire unique.

### 1.1.2. L'apport génomique chez *P. aeruginosa*

Les génomes complets de quatre souches (PAO1, PA14, PA7 et LESB58) de *Pseudomonas aeruginosa* ont été séquencés puis annotés. Ils sont accessibles sur le site web : <http://www.Pseudomonas.com/>.

Le génome de *P. aeruginosa* est l'un des plus grands génomes bactériens identifiés. Selon les souches, sa taille varie de 5,2 à 7,1 Mb et contient environ 5500 ORF (Open Reading Frame). Ceci confère à cette bactérie de considérables facultés d'adaptation et la possibilité de coloniser un grand nombre de niches environnementales (Croft et al., 2000; Lee et al., 2006). Cette variation s'explique par la perte ou le gain de matériel génétique par transferts

horizontaux (Römling et al., 1997). Une étude comparative entre les génomes des différentes souches a mis en évidence un squelette commun très conservé et une partie variable qui représentent 90 % et 10 % du génome respectivement (Spencer et al., 2003). Cette variation résulte de l'acquisition de plasmides, de transposons, de bactériophages et d'autres régions instables appelées îlot génomiques. Les îlots génomiques, par l'acquisition de nouveaux mécanismes, sont un moyen pour la bactérie de s'adapter rapidement face à la pression de sélection. Parmi les îlots génomiques, les régions spécialisées qui codent pour des facteurs de virulence sont nommés îlots de pathogénicité. Par exemple, dans la souche PA14, il a été identifié deux îlots de pathogénicité: PAPI-1 et PAPI-2 (He et al., 2004). L'îlot PAPI-2 code pour l'exotoxine ExoU sécrétée par le système de sécrétion de type III. Cette phospholipase dégrade les phospholipides de la membrane épithéliale ce qui conduit à la destruction des cellules et à des infections aigües (Finck-Barbançon et al., 1997). Ils expliquent les différences de virulence observées en fonction des souches.

### **1.1.3. Habitat et pouvoir pathogène**

*P. aeruginosa* est un pathogène opportuniste capable d'infecter un large spectre d'hôtes comme les amibes, les nématodes, les insectes, les plantes, les animaux et l'homme (Silby et al., 2011). Chez les individus sains, il est peu virulent mais il est la cause de graves infections chez les individus immunodéprimés (cancéreux, malades du SIDA (syndrome de l'immunodéficience acquise)), les grands brûlés, les patients intubés-ventilés en réanimation et les patients qui souffrent de maladies chroniques respiratoires comme la mucoviscidose. Après *Staphylococcus aureus* et *Escherichia coli*, *P. aeruginosa* figure au troisième rang des pathogènes causant des infections nosocomiales en France. Chez les patients atteints de la mucoviscidose (cystic fibrosis - CF), il représente la principale cause de mortalité entraînant un sévère déclin des fonctions respiratoires.

En fonction de l'environnement, le génotype et les phénotypes de *P. aeruginosa* varient (Hogardt and Heesemann, 2010). A l'étape initiale de l'infection, le pathogène sous la forme planctonique exprime des facteurs de virulence qui induisent des infections aigües. Des molécules toxiques sont sécrétées par différents systèmes de sécrétions ce qui entraîne une destruction rapide des cellules (systèmes de sécrétion de type II et III). Les infections aigües sont transitoires bien qu'elles peuvent conduire à une septicémie. A un stade plus avancé, *P. aeruginosa* va former un biofilm lui permettant de persister au sein des poumons de son hôte. La formation d'un biofilm entraîne une destruction lente des tissus pulmonaires et est souvent associée à des infections chroniques.

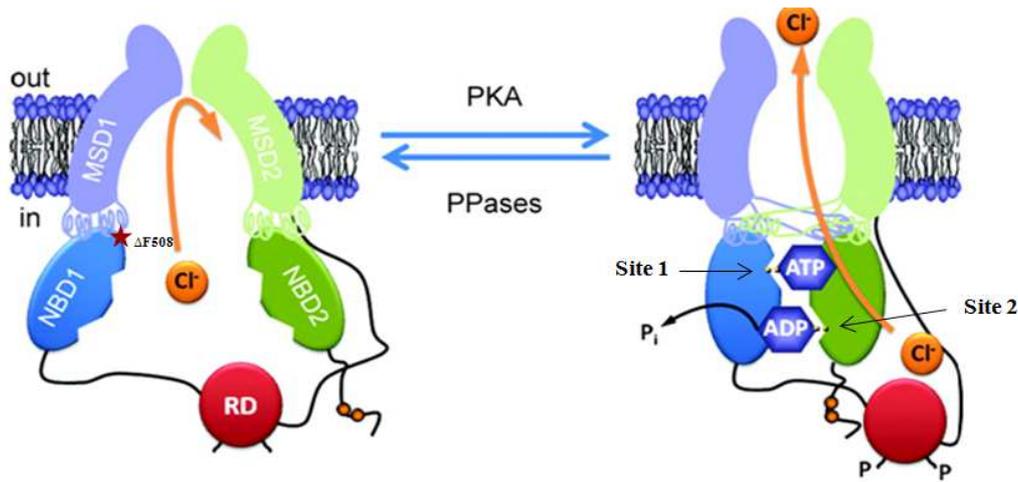
## 1.2. La mucoviscidose

1. La mucoviscidose, qui provient du terme « mucus » et « viscosité », est une maladie génétique qui touche les tissus épithéliaux exocrines (atteinte pulmonaire, digestive, génitale, glandes sudoripares). Cette maladie génétique est létale et de transmission autosomique récessive. Elle est retrouvée le plus fréquemment dans les populations de type caucasien. Elle conduit à une obstruction pulmonaire chronique ainsi qu'à une insuffisance pancréatique exocrine et une sueur contenant une concentration élevée en sodium et en chlore. Aucun traitement curatif n'existe et les traitements actuels sont lourds et contraignants. Elle affecte 2500 nouveaux nés et 70 000 personnes dans le monde par an (Plan national maladies rares 2005-2006 Ministère de la Santé). En France, l'âge moyen de décès est de 27 ans et l'espérance de vie est de 47 ans (Registre Français de la Mucoviscidose (RFM) : Bilan des données 2010 de l'Observatoire National de la Mucoviscidose (ONM)).

### 1.2.1. Une maladie génétique

La mucoviscidose est causée par des mutations sur le gène *cftr* (Cystic Fibrosis Transmembrane conductance Regulator). Le gène *cftr*, localisé sur le chromosome 7, code pour une protéine de transport qui forme un canal à ions chlorures au niveau du pôle apical des cellules épithéliales glandulaires. La protéine CFTR appartient à la famille des ABC transporteurs (ATP-binding cassette). Elle est composée de cinq domaines : (1) deux domaines transmembranaires MSD (Membrane Spanning Domain) chacun composé de six segments transmembranaires, (2) deux domaines NBD (Nucleotide Binding Domain) capables de lier l'ATP, et (3) un unique domaine de régulation RD (Regulation Domain) cytoplasmique qui contient des sites de phosphorylation (Riordan et al., 1989) (Figure 1).

La phosphorylation du RD par une PKA (protéine kinase c-AMP dépendante) est prérequis pour l'ouverture du canal (Ostedgaard et al., 2001). La fixation d'un ATP sur le site 1 va provoquer la formation d'un dimère en tête à queue des deux domaines NBDs. Si l'ATP se fixe sur le site 1, son hydrolyse se fait au niveau du site 2 du dimère NBDs (Lewis et al., 2004). La formation du dimère NBDs va induire un changement de conformation des deux domaines MSDs permettant l'ouverture du canal et l'exportation d'ion chlorure (Vergani et al., 2005) (Figure 1).



**Figure 1: Représentation schématique du fonctionnement de la protéine de transport CFTR (Cystic Fibrosis Transmembrane Regulator).** A gauche est représenté la forme quiescente du CFTR et à droite la forme ouverte qui permet le transport des ions chlorures à l'extérieur de la cellule. L'étoile rouge représente la mutation  $\Delta F508$  (D'après (Hwang and Sheppard, 2009)).

D'après le consortium international « Cystic Fibrosis Genetic Analysis Consortium », plus de 1571 mutations réparties tout au long du gène ont été identifiées. Selon les types de mutations, six classes fonctionnelles ont été définies : (Classe I) pas de production de la protéine CFTR, (Classe II) protéine CFTR mal synthétisée, (Classe III) protéine CFTR mal régulée, (Classe IV) conductance ionique de CFTR altérée, (Classe V and VI) faible quantité de CFTR produite.

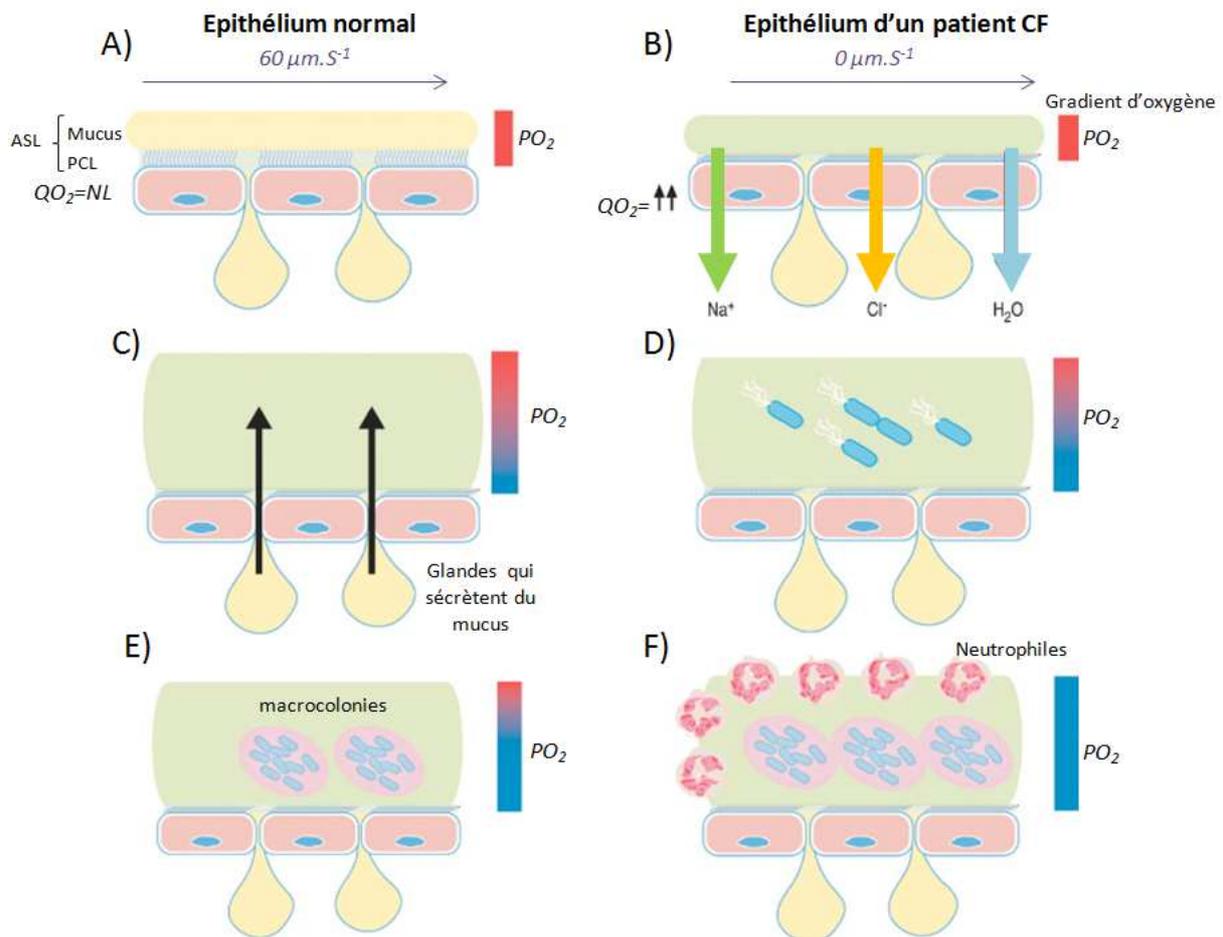
La mutation la plus fréquente dans 70 % des cas est la délétion de 3 paires de bases responsables de l'absence de la phénylalanine en position 508 connue sous le nom de  $\Delta F508$  (Classe IV) (Fekete et al., 1992).

### 1.2.2. Les infections bactériennes

Chez les patients atteints de la mucoviscidose (patients CF), le dysfonctionnement du canal transmembranaire CFTR est responsable de nombreuses infections bactériennes. Les mécanismes antibactériens tels que la clairance mucociliaire et l'activité antimicrobienne qui assurent un environnement stérile au sein des poumons sont inhibés.

Les cellules respiratoires sont recouvertes d'une couche appelée ASL (Airway Surface Liquid). Cette couche est composée d'une couche supérieure visqueuse, le mucus, et d'une couche inférieure plus fluide PCL (Periciliary Liquid Layer) qui permet le battement des cils à la surface de l'épithélium respiratoire (Figure 2). Le mucus forme une barrière de protection sur laquelle se fixent toutes les bactéries inhalées. Les battements ciliaires vont par la mise en mouvement du mucus assurer l'expulsion des bactéries hors des voies respiratoires.

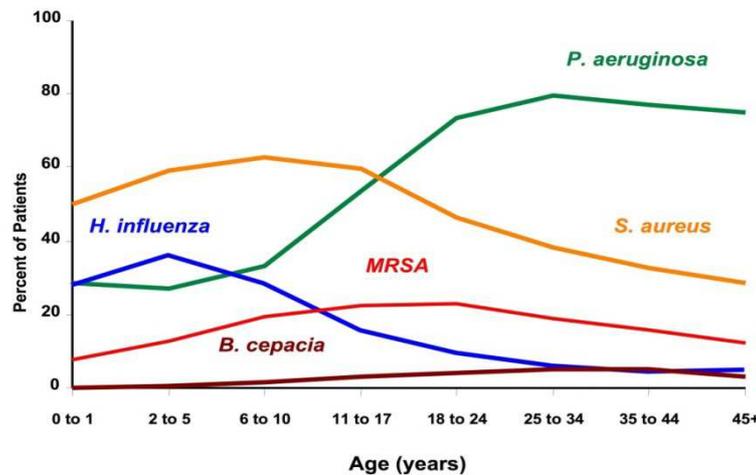
Chez les patients CF, l'augmentation de la viscosité du mucus bloque les battements ciliaires. Les bactéries ne sont plus éliminées et peuvent alors persister au sein des poumons (Boucher, 2003; Chmiel and Davis, 2003). Il a également été observé chez ces patients, une modification ionique ainsi qu'une acidification du mucus responsable de l'inhibition de l'activité des peptides anti-microbiens, des lysozymes, des lactoferrines ainsi que des neutrophiles polymorphonucléaires (Barasch and al-Awqati, 1993; Smith et al., 1996) (Figure 2).



**Figure 2: Schéma représentant le processus infectieux de *Pseudomonas aeruginosa* à la surface de l'épithélium respiratoire des patients atteints de la mucoviscidose.** A) Mécanisme de clairance mucociliaire normale. B) Chez les patients CF, le dysfonctionnement de la protéine CFTR engendre une absence de sécrétion des ions chlore favorisant une hyperabsorption d'eau et d'ions sodium. C) Les battements ciliaires sont bloqués par l'épaississement du mucus qui devient visqueux. D) L'inhibition de la clairance mucociliaire favorise l'infection bactérienne et la formation d'un biofilm. E) La modification du mucus et la formation du biofilm rend inefficace le système immunitaire.  $PO_2$  est la pression partielle d'oxygène. (Réalisée d'après (Boucher, 2004)).

Plusieurs espèces bactériennes sont capables de coloniser les voies respiratoires des patients CF. Au stade précoce de la maladie, *S. aureus* et *Haemophilus influenzae*, sont les deux pathogènes les plus présents. Puis après l'âge de 20 ans, plus de 80 % des patients contractent

une infection à *P. aeruginosa* (Figure 3). Cette dernière est très grave et représente la première cause de mortalité des patients CF suite à déclin de la fonction respiratoire important dans 90 % des cas (Cystic Fibrosis Foundation Patient Registry 2006 Annual Data Report Bethesda, Maryland). Lorsque la structure des poumons s'altère et que des fibroses apparaissent, seule une transplantation d'organe peut sauver la vie du patient. Face à sa multirésistance aux antibiotiques et surtout sa capacité à former un biofilm, *P. aeruginosa* est un pathogène très difficile à éradiquer.



**Figure 3: Prévalence des infections bactériennes respiratoires en fonction de l'âge des patients CF. MRSA : Multi Resistant *S. aureus*** (Observatoire National de la Mucoviscidose- Bilan des données 2004, Vaincre la Mucoviscidose et Ined, Paris, 2006).

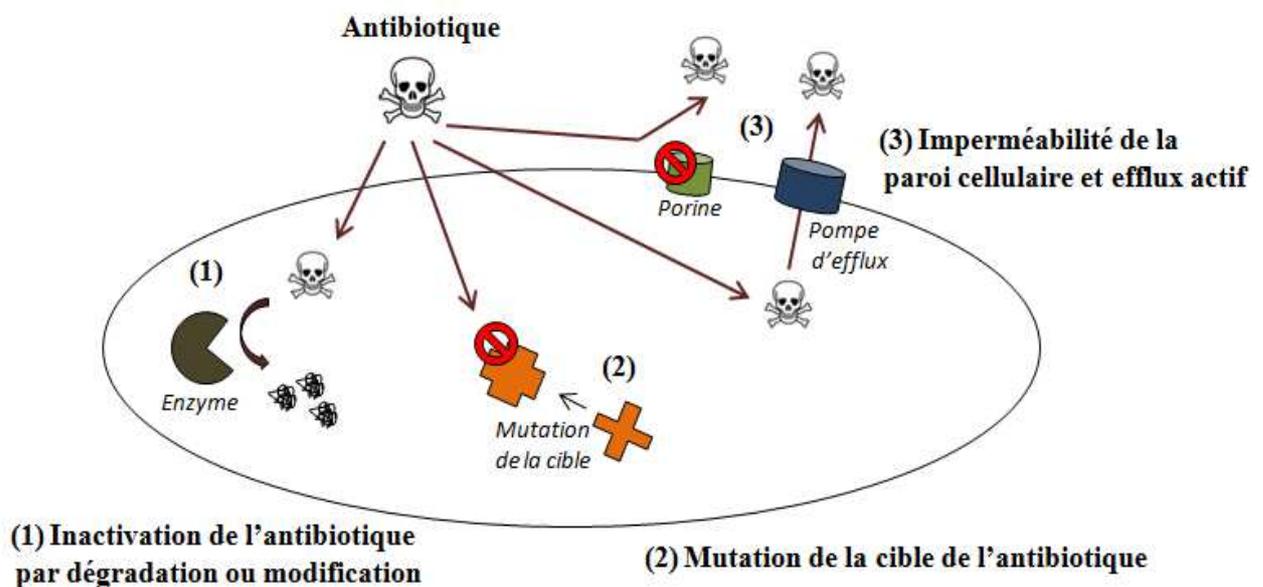
### 1.3. Une multi-résistance aux antibiotiques

La plupart des souches de *P. aeruginosa* présentent une multi-résistance à de nombreuses classes d'antibiotiques. Elle est la résultante d'un ensemble de mécanismes qui sont mis en place dans le but d'inactiver ou d'empêcher l'antibiotique d'atteindre sa cible (Figure 4).

Le premier mécanisme utilise des enzymes qui vont inactiver l'antibiotique telles les  $\beta$ -lactamases qui hydrolysent les  $\beta$ -lactamines (Weldhagen et al., 2003) et les acétyltransférases qui inactivent les aminoglycosides en les modifiant (Poole, 2005). Une autre stratégie consiste à muter la cible de l'antibiotique. Des mutations spécifiques sur l'ADN gyrase, GyrA ont été identifiées sur des isolats cliniques. Elles confèrent à *P. aeruginosa* une résistance pour les fluoroquinolones en provoquant une diminution de l'affinité de GyrA pour l'antibiotique (Mouneimné et al., 1999).

La restriction de la perméabilité membranaire est également une stratégie de résistance efficace. Sur les membranes des bactéries à Gram négatif sont en effet présentes des porines non sélectives et sélectives qui permettent la diffusion de petites molécules hydrophiles

comme les antibiotiques. La porine sélective OprD laisse entrer les carbapénèmes. Une diminution de l'expression ou une altération des porines OprD rend donc *P. aeruginosa* résistant aux carbapénèmes (Köhler et al., 1999; Epp et al., 2001). Parmi les porines non sélectives OprF, seulement 5 % sont fonctionnelles (Benz and Hancock, 1981). Les composés nocifs qui pénètrent par les porines sont également relargués vers le milieu externe par les pompes d'efflux. Chez *P. aeruginosa*, les systèmes d'efflux de la famille RND (pour Resistance-Nodulation-Division) participent ainsi à protéger la bactérie des antibiotiques. La famille RND comporte dix systèmes d'efflux capables de relarguer différentes familles antibiotiques. Parmi les 10 pompes d'efflux caractérisées, quatre ont été observées dans des isolats cliniques (MexAB-OprM, MexCD-OprJ, MexXY-OprM, MexEF-OprN) avec une forte prévalence pour les systèmes MexAB-OprM et MexXY-OprM. Ces systèmes permettent le relargage de nombreux antibiotiques comme les  $\beta$ -lactamines ou les aminoglycosides, deux antibiotiques très utilisés dans les centres hospitaliers (Mesaros et al., 2007). Enfin, il faut noter que cette multi-résistance est augmentée d'un facteur mille en condition de biofilm par rapport aux bactéries planctoniques (Hoyle and Costerton, 1991).



**Figure 4: Représentation schématique des mécanismes de résistance aux antibiotiques mis en place par *P. aeruginosa*.**

## 1.4. Le biofilm

La forme planctonique de *P. aeruginosa* n'est qu'un état transitoire qui lui permet de coloniser et d'initier la formation d'un biofilm. Le biofilm est un mode de vie où le pathogène va pouvoir s'établir et persister. Il constitue une communauté bactérienne organisée qui est englobée dans une matrice extracellulaire auto-produite adhésive et protectrice.

### 1.4.1. La matrice extracellulaire

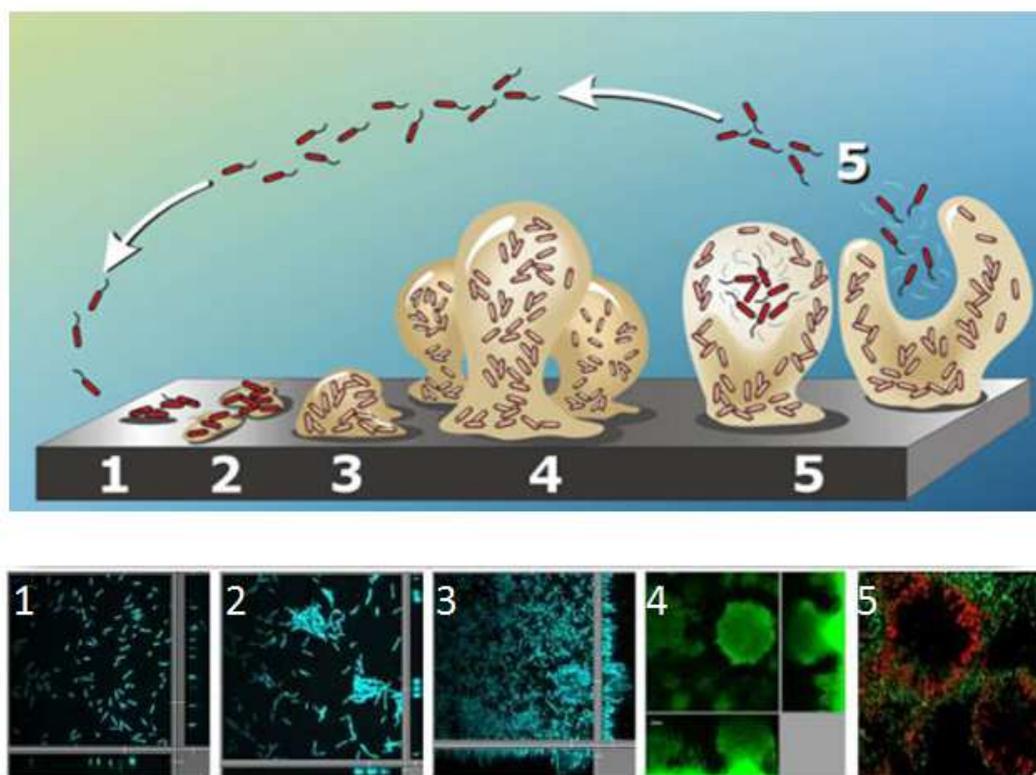
La composition de la matrice extracellulaire peut varier. L'existence d'une matrice extracellulaire composée d'ADN a récemment été observée. Elle est cependant souvent constituée de protéines, d'acides nucléiques, et surtout d'exopolysaccharides. Trois types d'exopolysaccharides (EPS) sont produits par *P. aeruginosa*: l'alginate, le Pel (pour Pellicule) et le Psl (pour Polysaccharides Synthesis Locus).

Durant les premières phases de la colonisation, la bactérie est non-mucoïde et produit les exopolysaccharides Psl et Pel. L'EPS Pel est riche en mannose et en galactose, et l'EPS Psl est une répétition de pentasaccharides qui contiennent du D-mannose, du D-glucose et du L-rhamnose (Byrd et al., 2009).

La production de l'EPS Psl va protéger la bactérie des cellules phagocytaires jusqu'à l'apparition du phénotype mucoïde (Mishra et al., 2012). Le basculement vers le phénotype mucoïde est associé à une phase de colonisation plus tardive. *P. aeruginosa* va produire en grande quantité de l'alginate qui lui confère une meilleure protection contre le système immunitaire. L'alginate est un EPS visqueux constitué d'acides  $\beta$ -D-mannuronique et  $\alpha$ -L-guluronique liés par des liaisons glycosidiques  $\beta$ -1,4.

### 1.4.2. La formation du biofilm

La formation du biofilm se réalise en différentes étapes selon un processus dynamique et hiérarchisé. Sur la base des études de Sauer, un modèle classique de la formation du biofilm est proposé en cinq étapes : (1) l'attachement réversible, (2) l'attachement irréversible, (3) la première étape de maturation, (4) la deuxième étape de maturation et (5) la dispersion (Sauer et al., 2002b) (figure 5).



**Figure 5: Illustration du développement du biofilm en différentes étapes chez *P. aeruginosa*.** La formation du biofilm se réalise en cinq étapes : (1) l'attachement initial et réversible. Les bactéries sont sous forme planctonique et adhèrent à la surface par leur flagelle et leurs pili de type IV. (2) l'attachement irréversible. Perte de la mobilité flagellaire au profit de la « Twitching motility » assurée par les pili de type IV. D'autres structures adhésives sont synthétisées ainsi que la matrice exopolysaccharidique qui permettent l'adhésion de la bactérie à la surface. (3) Première étape de maturation caractérisée par la formation de microcolonies. La production d'exopolysaccharides est plus abondante. Des structures adhésives permettant des interactions bactérie/bactérie sont produites. (4) Seconde étape de maturation avec l'apparition de structures en forme de « champignon ». (5) La dispersion. Une altération de la structure « champignon » est observée avec la lyse de certaines bactéries au niveau du chapeau du « champignon » (représentées en rouge sur les images de microscopie confocale). Les autres bactéries synthétisent *de novo* leur flagelle ce qui leur permet d'aller coloniser de nouvelles niches. (Réalisée d'après (Monroe, 2007; Bordi and de Bentzmann, 2011).

- Le « film conditionnant »

Avant l'initiation du biofilm, l'adsorption des éléments nutritifs contenus dans le milieu sur le support va permettre la formation d'une pellicule appelée « film conditionnant ». Sa présence va devenir un lieu d'attraction chimiotactique ainsi qu'un site d'adhésion spécifique pour les bactéries. Sa formation constitue l'étape 0 de la formation du biofilm.

- L'attachement réversible et irréversible

La première étape de la formation du biofilm est l'attachement initial. Durant cette étape, les bactéries planctoniques se déplacent grâce à leur flagelle afin de sélectionner une surface colonisable. Des contacts vont s'établir entre la surface et les structures adhésives du flagelle et des pili de type IVa. L'attachement qui en résulte est réversible ce qui permet aux bactéries d'explorer la surface et de se décrocher si l'environnement est néfaste pour leur croissance.

Les bactéries perdent ensuite leur flagelle et d'autres structures adhésives comme les pili de type IVb et les pili de type Cup sont exprimés à la surface des bactéries permettant une adhésion irréversible. Un autre type de mobilité (« Twitching motility, glissement par saccade ») est mis en place par les pili de type IVa, ce qui permet d'initier la formation de microcolonies. A ce stade, les bactéries regroupées de manière monostratifiée, synthétisent des exopolysaccharides tels que Pel et Psl qui constituent la matrice extracellulaire.

- La maturation du biofilm

L'étape de maturation implique un regroupement de manière pluristratifiée des bactéries qui forment des microcolonies d'environ 10 µm d'épaisseur. La production de la matrice extracellulaire est de plus en plus abondante et les bactéries génèrent des interactions cellule/cellule grâce à la synthèse de nouvelles structures adhésives.

Durant la seconde étape de maturation, la matrice extracellulaire est composée d'exopolysaccharides, d'ADN et de protéines. Les microcolonies prennent la forme d'un « champignon ». Au sein du « champignon » l'accessibilité des nutriments et de l'oxygène est différente ce qui entraîne une certaine hétérogénéité des bactéries. Les bactéries qui se situent au niveau du « chapeau du champignon » ont accès aux nutriments et à l'oxygène, elles présentent une « Twitching motility ». Alors que les bactéries situées au pied des « champignons » sont en condition anaérobie et ne présentent pas de mobilité. Elles sont les seules à pouvoir sécréter des rhamnolipides (glycolipides constitués d'un acide gras et d'un à deux rhamnosés). Ces derniers sont impliqués dans la structuration du biofilm et forment des canaux aidant à la circulation des nutriments et de l'oxygène nécessaires aux bactéries mais aussi à l'évacuation des déchets.

- La dispersion

La dernière étape est la dispersion du biofilm. Au sein du chapeau du « champignon » est observée la mort de nombreuses bactéries. Les bactéries qui ont survécu synthétisent *de novo*

leur flagelle afin de coloniser un autre endroit ou de rejoindre une autre structure existante (Parsek and Fuqua, 2004).

### **1.4.3. Le biofilm, une forme de résistance**

Le biofilm est un mode de vie qui offre à la bactérie une plus grande résistance aux attaques du système immunitaire et aux agents antimicrobiens. Cette résistance est due à la matrice extracellulaire riche en exopolysaccharides qui agit comme une barrière naturelle. Face à cette structure qui peut atteindre 100 µm d'épaisseur, les cellules phagocytaires sont inefficaces. Le processus de phagocytose reste donc bloqué (Alhede et al., 2009; Mishra et al., 2012). De plus, les anticorps sont incapables de diffuser au travers de cette matrice extracellulaire pour reconnaître les différents antigènes ainsi que les composés hydrophiles et chargés positivement.

Des expériences ont montré qu'une solution contenant 2 % d'alginate bloque totalement la diffusion des antibiotiques tels que la gentamycine et la tobramycine. La lyse de cet EPS (majoritairement produit par des souches mucoïdes de *P. aeruginosa* retrouvés chez les patients CF) augmente l'efficacité des antibiotiques sur le biofilm (Hatch and Schiller, 1998; Alkawash et al., 2006). L'alginate permettrait de retarder principalement la pénétration des aminoglycosides qui sont chargés positivement. Les β-lactamates et des fluoroquinolones seraient retardés avec une moins grande efficacité par l'alginate du fait de la différence de charge.

Le biofilm constitue un environnement propice à une adaptation génétique et une évolution de *P. aeruginosa*. Le stress oxydatif et nutritif conduit à la formation de nouvelles sous populations microbiennes. Il a été identifié au sein du biofilm des colonies « Small Colony Variant » qui présentent une très forte résistance aux antibiotiques (Von Götz et al., 2004). Une autre population de cellules sous forme de dormance appelées « persisters » possède également une multirésistance aux antibiotiques. Cette population, qui émerge sous des conditions de stress, ne représente que 1 % des cellules du biofilm. Ce phénotype est réversible lorsque les conditions de stress cessent (Anderson et al., 2008).

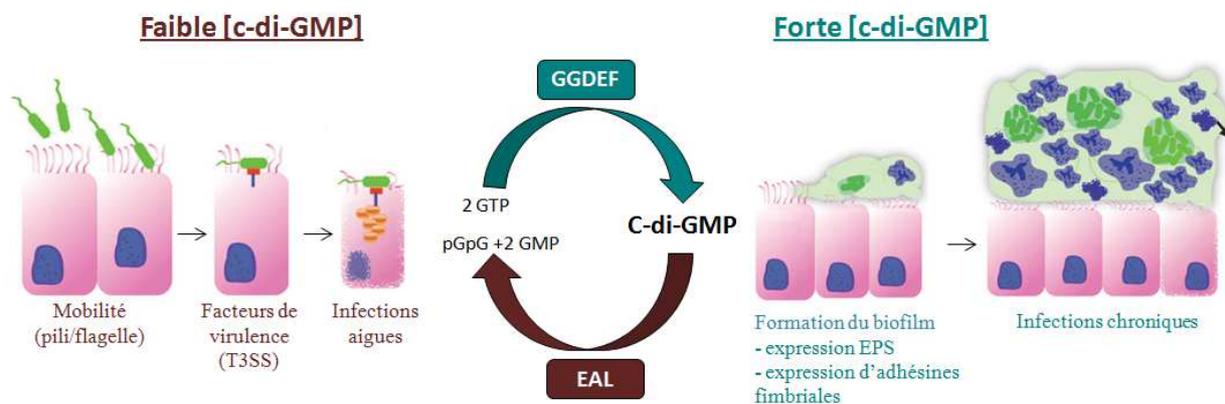
### **1.4.4. Perception des signaux environnementaux et régulation**

En fonction des conditions environnementales, *P. aeruginosa* va déclencher le processus de formation du biofilm. La transition de l'état planctonique à l'état sessile nécessite des mécanismes de régulation. Les systèmes à deux composants (TCS) sont l'un des principaux mécanismes qui confèrent à *P. aeruginosa* la capacité de percevoir des signaux et d'induire

une réponse adaptée. Une molécule de signalisation intracellulaire, le di-guanilate monophosphate cyclique (c-di-GMP), est également impliquée dans l'installation du biofilm.

#### 1.4.4.1. Le c-di-GMP

Le c-di-GMP est une molécule de signalisation intracellulaire spécifique aux bactéries. Chez *P. aeruginosa*, il a été identifié que le contrôle de son taux intracellulaire va induire des changements phénotypiques reliés à la transition entre l'état planctonique et sessile de la bactérie. Une forte concentration du c-di-GMP va induire la biosynthèse des structures adhésives et la production d'exopolysaccharides ce qui va mener à la formation du biofilm. Alors qu'une faible concentration du c-di-GMP induit une forme motile, planctonique qui favorise les infections aiguës par l'expression de facteurs de virulence. La synthèse ou la dégradation du c-di-GMP se fait par la présence de protéines qui contiennent un domaine à activité diguanylate cyclase (GGDEF) ou un domaine à activité phosphodiesterase (EAL). Les protéines qui comportent un domaine GGDEF contrôlent de façon positive la formation du biofilm alors que les protéines qui ont un domaine EAL inhibent la formation du biofilm (Figure 6).



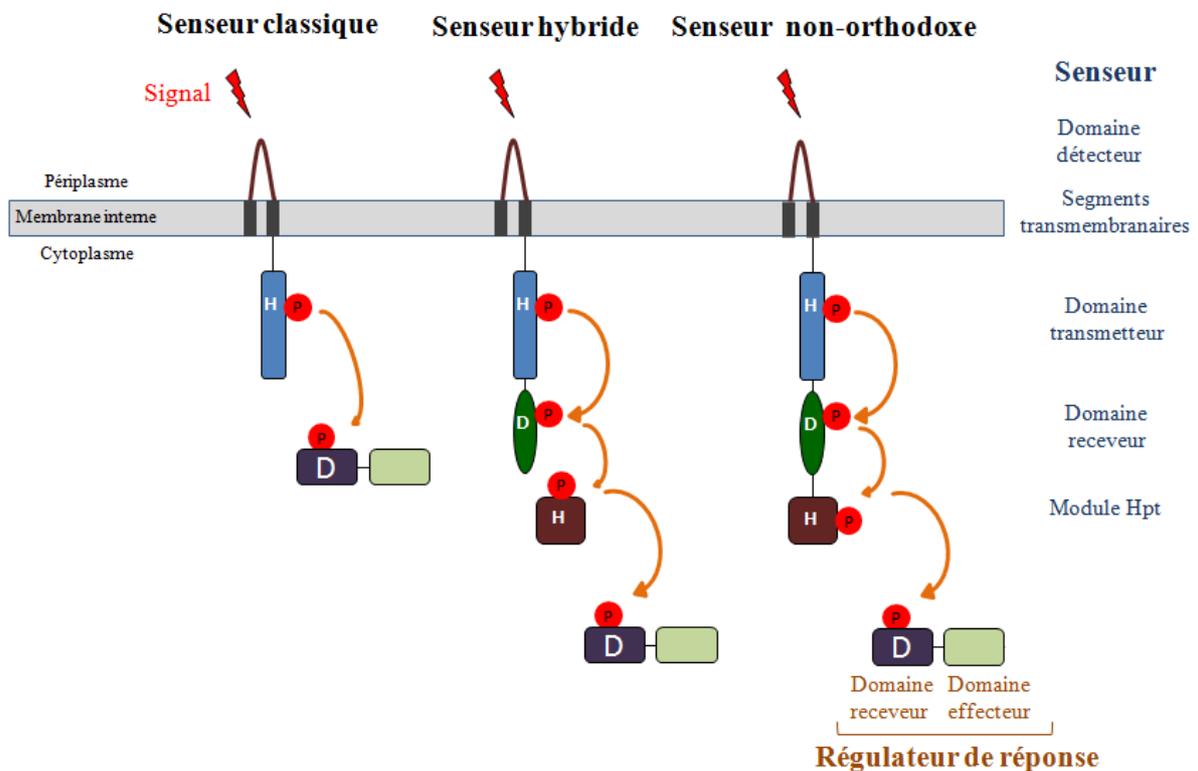
**Figure 6: Représentation schématique de la signalisation du c-di-GMP chez *P.aeruginosa*.** Les protéines à domaine GGDEF synthétisent le c-di-GMP. Il est dégradé par les protéines à domaine EAL. Un taux élevé du c-di-GMP entraîne la formation d'un biofilm ainsi que des infections chroniques. Un faible taux du c-di-GMP est associé à des infections aiguës dues à l'expression de facteurs de virulence à l'état planctonique de *P.aeruginosa*. (Réalisée d'après (Römling et al., 1997; Williams et al., 2010)).

Durant la formation du biofilm, la mobilité est un facteur important qui est finement régulé. Le second messager c-di-GMP intracellulaire participe à cette régulation ainsi qu'à la structuration du « champignon ». Les cellules qui constituent la base du biofilm présentent un fort taux de c-di-GMP. Elles produisent des exopolysaccharides et ne présentent pas de

mobilité alors que les cellules qui présentent une « twitching motility » au niveau du « chapeau » de la microcolonie ont un faible taux de c-di-GMP. Récemment, il a été démontré que l'augmentation du c-di-GMP induit par des pilines mineures impliquées dans la synthèse du pilus de type IV, inhibait la mobilité flagellaire (Kuchma et al., 2012).

#### 1.4.4.2. Les systèmes à deux composants

Les systèmes à deux composants sont formés d'une protéine senseur histidine-kinase généralement intégrée à la membrane et d'un régulateur de réponse cytoplasmique. Il existe trois classes de senseurs histidine-kinase : les senseurs classiques, les senseurs non orthodoxes et les senseurs hybrides (Figure 7).



**Figure 7: Représentation des systèmes à deux composants chez *P. aeruginosa*.** (Reproduite de (Rodrigue et al., 2000)).

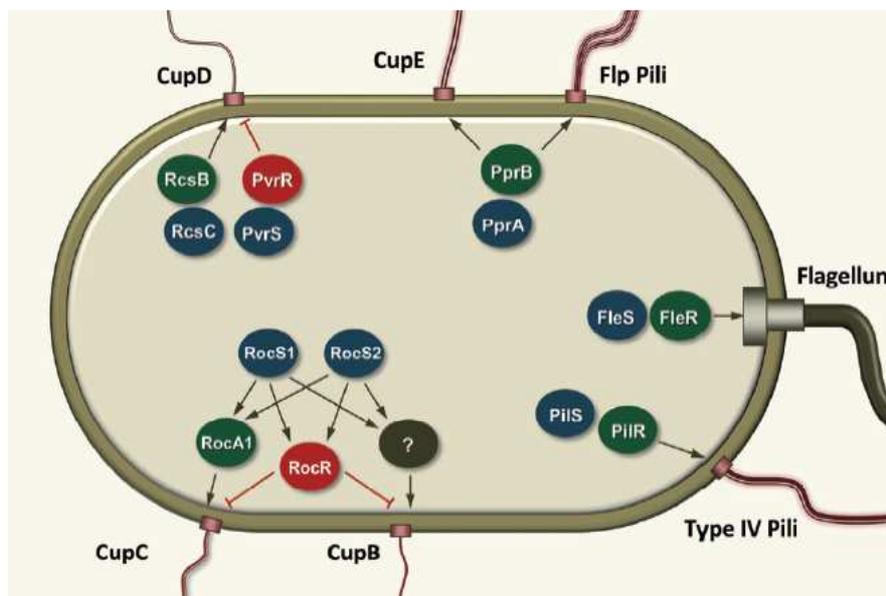
Le senseur va percevoir un signal par son domaine récepteur. La détection d'un stimulus va induire l'autophosphorylation du domaine transmetteur qui va transmettre le groupement phosphate au domaine receveur du régulateur de réponse. L'activation du régulateur de réponse permet une réponse cellulaire adaptée au signal reçu par le senseur. La plupart des récepteurs de régulation sont des facteurs de transcription. Certains régulateurs de réponse

sont capables de synthétiser (domaine GGDEF) ou de dégrader (domaine EAL) le c-di-GMP, et d'autres ont une activité enzymatique spécifique.

Pour les senseurs hybrides et non-orthodoxes la cascade de phosphorylation peut se faire en trois étapes faisant intervenir des domaines additionnels du senseur : un domaine receveur et un module HPT (Histidine Phospho-Transfert). Pour le senseur hybride, le module HPT est une protéine à part entière.

*P. aeruginosa* possède plus de 20 systèmes à deux composants qui lui permettent de s'adapter en fonction des conditions environnementales.

Durant la formation du biofilm, plusieurs TCS sont mis en place afin de déclencher la synthèse d'appendices de surface impliqués dans l'adhésion de la bactérie ainsi que dans la production d'exopolysaccharides (Figure 8). Certains de ces TCS peuvent en même temps inhiber la production de facteurs de virulence. C'est le cas du senseur LadS, qui en permettant la synthèse d'EPS Pel/Psl, réprime l'expression des gènes des systèmes de sécrétion de type II et type III.



**Figure 8: Représentation des différents systèmes à deux composants (TCS) qui contrôlent la synthèse d'appendices extracellulaires chez *P. aeruginosa*.** Les senseurs figurent en bleu, les régulateurs de réponse possédant un domaine effecteur capable de se lier à l'ADN sont en vert, et les régulateurs de réponse possédant un domaine effecteur EAL sont en rouge. (Mikkelsen et al., 2011).

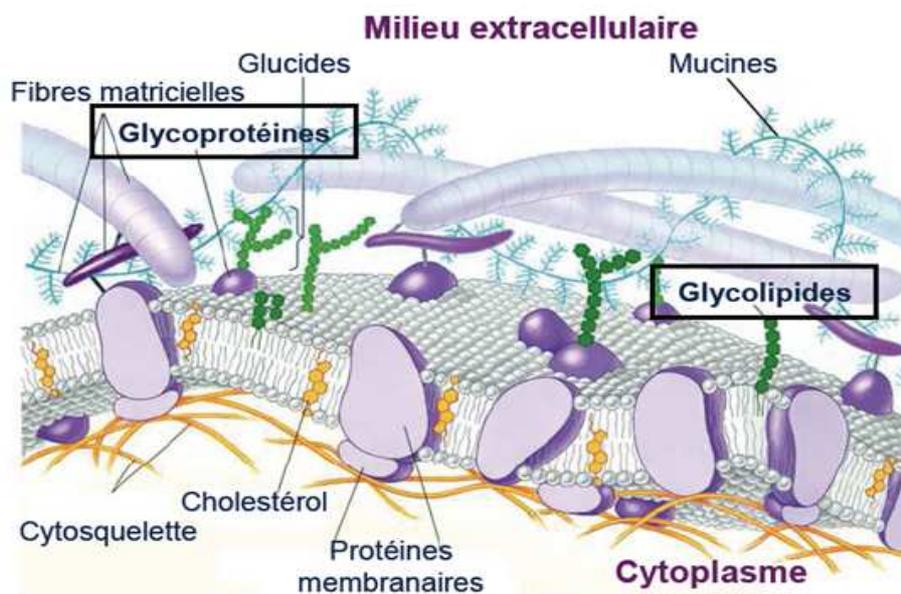
---

## 2. La Glyco-stratégie de *Pseudomonas aeruginosa*

---

### 2.1. La glycosylation chez l'homme

La membrane externe des cellules est tapissée par une couche appelée « glycocalix » dont la structure filamenteuse légère, maillée est composée par des glycoconjugués. Les glycoconjugués sont des chaînes de sucres, appelées glycanes, liés de manière covalente à des protéines (glycoprotéines ou protéoglycanes) ou à des lipides (glycolipides) (Figure 9). Les glycanes sont à la fois des récepteurs mais également des marqueurs qui vont donner une spécificité et une identité à la cellule. Leur présence à la surface des cellules permet de nombreux processus biologiques indispensables à la vie de la cellule comme l'adhésion, le trafic cellulaire, la transduction du signal ou encore l'endocytose (Ohtsubo and Marth, 2006).



**Figure 9: Illustration des différents glycoconjugués présents à la surface d'une membrane cellulaire.** (D'après Copyright 2002 Person Education, Inc., publishing as Benjamin Cummings)

La biosynthèse des glycoconjugués se fait par des glycosyltransférases qui transfèrent un monosaccharide associé à un nucléotide sur un accepteur protéique, lipidique ou saccharidique.

#### 2.1.1. Les glycoprotéines

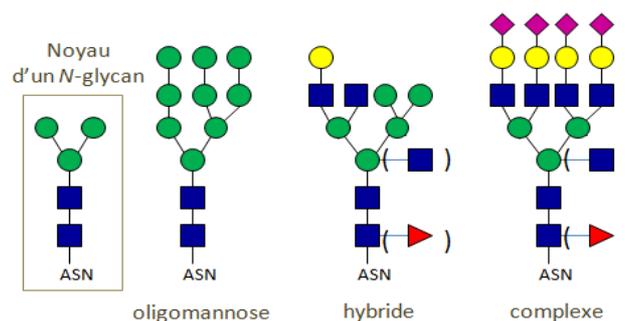
La glycosylation des protéines est l'une des modifications post-traductionnelles les plus courantes puisque qu'il est estimé qu'environ 70 % des protéines humaines sont glycosylées.

En fonction du type de liaison qui relie le glycané à la partie peptidique, il existe deux types de glycoprotéines: les *N*-et *O*-glycoprotéines.

### 2.1.1.1. La *N*-glycosylation

La *N*-glycosylation est initiée dans le réticulum endoplasmique (modification co-traductionnelle) par l'établissement d'un lien *N*-glycosidique sur une asparagine d'une protéine comportant un séquen N-X-S/T (X peut représenter n'importe quel acide aminé excepté la proline). Elle se poursuit dans l'appareil de Golgi (modification post-traductionnelle) où les *N*-glycanes de la protéine sont maturés. La maturation des *N*-glycoprotéines conduit à la synthèse de trois types d'oligosaccharides : (1) les *N*-glycanes de type oligomannosidique, (2) les *N*-glycanes de type hybride et (3) les *N*-glycanes de type complexe.

Tous les *N*-glycanes sont composés d'un noyau pentasaccharidique commun qui comporte deux résidus *N*-acétylglucosamine (GlcNAc) et trois mannoses (Man) :  $\text{Man}_3\text{GlcNAc}_2$ . En plus du noyau  $\text{Man}_3\text{GlcNAc}_2$ , les *N*-glycanes de type oligomannosidique sont composés de résidus mannoses. Les *N*-glycanes de type complexe comportent des résidus galactose et *N*-acétylglucosamine et peuvent posséder à leurs extrémités des résidus galactose, *N*-acétylgalactosamine, fucose et /ou acide sialique. Ces *N*-glycanes peuvent exister sous la forme de structures bi-, tri-, tétra- ou penta-antennées. Les *N*-glycanes de type hybride possèdent les caractéristiques des *N*-glycanes de type oligomannosidique sur la branche Man ( $\alpha$ 1-6) et des *N*-glycanes de type complexe sur la branche Man ( $\alpha$ 1-3) (Figure 10).



**Figure 10: Illustration des trois types de *N*-glycanes chez l'homme.** Le noyau commun au *N*-glycane est encadré. Les résidus Gal et Man sont représentés par un rond jaune et vert respectivement, le GlcNAc est représenté par un carré bleu, le NeuAc est schématisé par un losange violet, et le Fuc par un triangle rouge. Les résidus entre parenthèse montrent une éventuelle présence.

Une protéine peut contenir plusieurs sites de *N*-glycosylation et plusieurs structures glycaniques différentes. La séquence mais surtout la conformation de la protéine peut

influencer les différentes étapes réticulaires et golgiennes de la *N*-glycosylation en favorisant ou non l'accessibilité du substrat aux différentes glycosyltransférases. Le type de cellule et l'état physiologique de la cellule ont également une influence sur la glycosylation (Hubbard and Ivatt, 1981; Goochee and Monica, 1990).

#### 2.1.1.2. La *O*-glycosylation

La *O*-glycosylation a lieu dans l'appareil de Golgi où le glycane est lié au niveau d'une sérine ou d'une thréonine. Selon la nature du sucre qui est lié à la protéine, plusieurs sous-groupes ont été définis (Chevance and Hughes, 2008) :

- La *O*-glycosylation de type mucine (Ser/Thr-*O*- $\alpha$ -N-acétylgalactosamine (GalNAc))
- La glycosylation de type *O*-GlcNAc (Ser/Thr-*O*- $\beta$ GlcNAc)
- La *O*-fucosylation (Ser/Thr-*O*- $\alpha$ Fuc)
- La *O*-glucosylation (Ser/Thr-*O*- $\alpha$ Glc)
- La *O*-mannosylation (Ser/Thr-*O*- $\alpha$ Man)

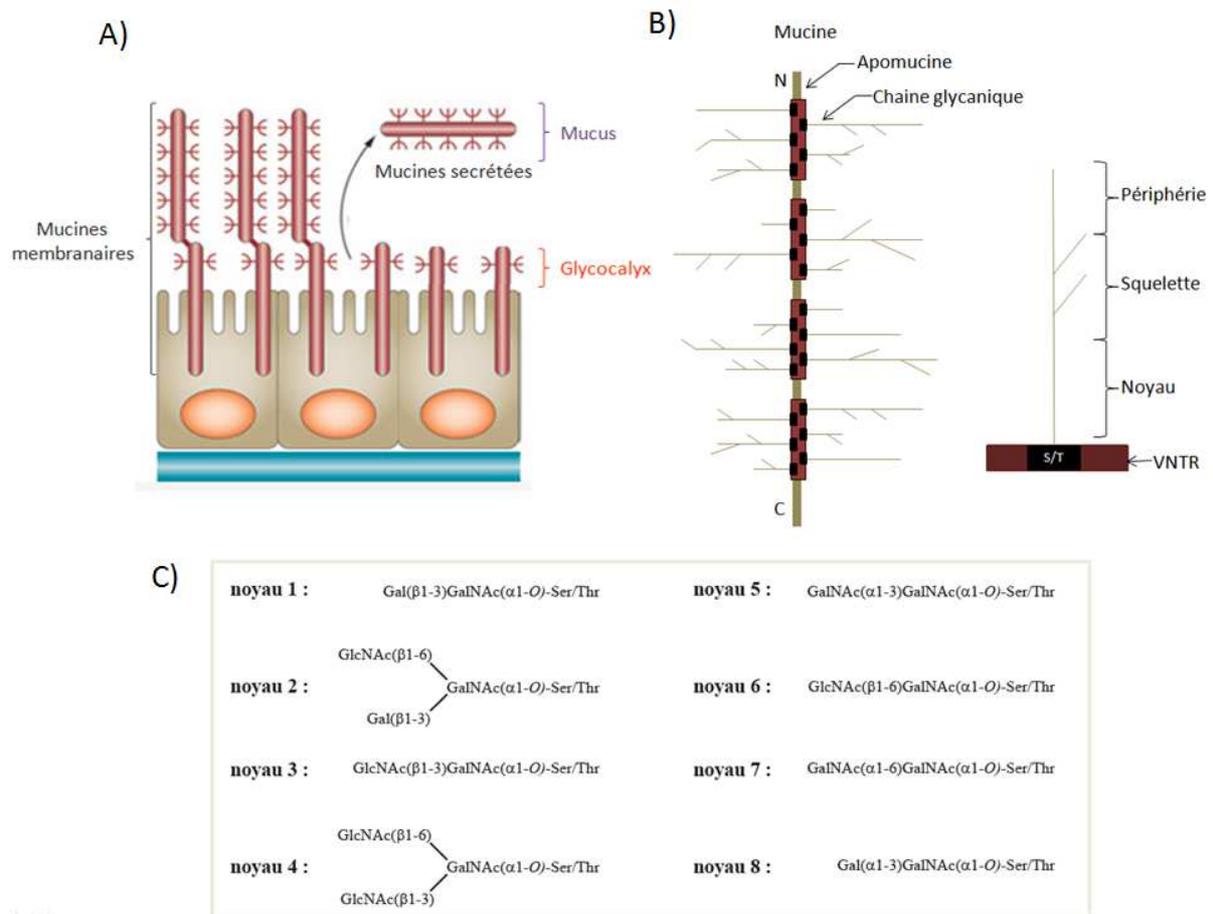
##### 2.1.1.2.1. La *O*-glycosylation de type mucine

Les mucines sont des glycoprotéines très fortement *O*-glycosylées situées dans le mucus qui tapissent la lumière des organes creux comme les voies respiratoires. Il existe deux classes de mucines, les mucines sécrétées et les mucines membranaires (Figure 11 A). Chaque mucine possède un cœur protéique appelé apomucine sur lequel se greffent des chaînes glucidiques. L'apomucine codé par les gènes MUC (Baldus et al., 2004), présente des régions tandem extrêmement riches en sérine et thréonine et/ou proline appelé VNTR (variable number of tandem region) (Figure 11 B).

Ces régions VNTR offrent un grand nombre de sites pour la *O*-glycosylation qui consiste en un transfert d'un résidu GalNAc en configuration  $\alpha$  sur l'hydroxyle de la chaîne latérale d'une sérine ou d'une thréonine par une glycosyltransférase. Ensuite, une cascade de réactions enzymatiques va permettre l'élongation des chaînes glycaniques qui seront composées de 2 à 20 monosaccharides. Les structures *O*-glycaniques qui représentent environ 80 % du poids moléculaire de la glycoprotéine sont composées de trois parties : le noyau (8 différents types de noyaux), le squelette qui peut être linéaire ou branché et la périphérie (Figure 11 B et C). Les extrémités non réductrices de ces chaînes peuvent contenir différentes structures de carbohydrates comme les épitopes glucidiques des systèmes sanguin ABO (antigène A, B, H et H-sulfaté) et lewis (Lewis a, Lewis b, Lewis x, Lewis y, sialyl et/ou sulfo Lewis a et Lewis

x) (cf 2.1.4). La synthèse de la périphérie des chaînes glycaniques implique des glycosyltransférases mais également et des sulfotransférases.

Les mucines par leurs chaînes glycaniques présentent une très grande diversité. Elles peuvent être impliquées dans les phénomènes d'adhésion cellulaire, dans le contrôle de l'antigénicité et du système immunitaire mais les mucines constituent surtout une barrière protectrice (Lamblin et al., 2001).



**Figure 11: Description des mucines présentes à la surface des cellules.** A) Représentation du rôle des mucines dans la protection cellulaire (Réalisée d'après (Niv and Fass, 2011)). Le mucus est constitué des mucines sécrétées qui forment la première ligne de défense. Les mucines membranaires constituent une deuxième ligne de défense. B) Schéma illustrant la structure d'une mucine. L'apomucine comporte des unités de répétition (VNTR) riche en sérine et thréonine où sont attachés des *O*-glycans. C) Structures des noyaux identifiés dans les chaînes glycaniques de type mucine.

### 2.1.2. Les protéoglycans

Les protéoglycans sont des glycoprotéines caractérisées par l'attachement covalent d'une ou plusieurs chaînes de polysaccharides appelés glycosaminoglycans (GAGs) sur une protéine via une zone de liaison tétrasaccharidique commune. Cette chaîne tétrasaccharidique est

composée de trois sucres différents, une unité de D-xylose, deux unités de D-galactose et une unité de D-glucuronique. Cette zone initie la polymérisation des chaînes GAGs qui sont liés de façon covalente au résidu L-sérine par une liaison *O*-glycosidique. Les glycosaminoglycanes sont de longs polysaccharides linéaires fortement polaires constitués d'une unité disaccharidique répétée. Le nombre de ces unités varie de 10 à 5000 ou plus. Ils sont formés d'un sucre amine, le glucosamine (GlcN) ou le N-acétyl-galactosamine (GalNAc), et d'un acide uronique.

Les GAGs sont classés en plusieurs catégories selon la nature du motif glycanique: les chondroïtines sulfates (CS) et les dermatanes sulfates (DS) qui contiennent un GalNAc, et les héparane-sulfates (HS) et les kératane-sulfates (KS) qui contiennent du GlcN. Certaines modifications comme la *N*-déacétylation-*N*-sulphatation, la 2-*O*- et 6-*O*-sulphatation et l'épimérisation de l'acide uronique offrent une grande variété au sein des chaînes GAGs.

### 2.1.3. Les glycolipides

Selon la nature des lipides, les glycolipides sont classés en deux groupes : les glycoylcérolipides et les glycosphingolipides.

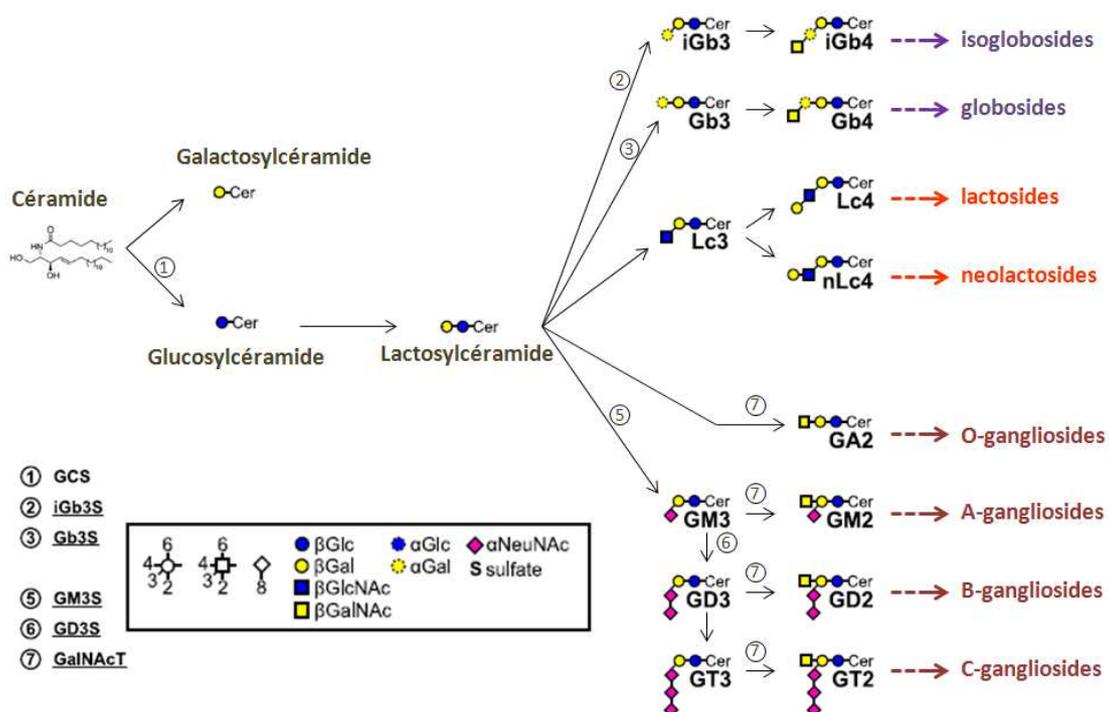


Figure 12: Voie de synthèse des différents glycosphingolipides chez les mammifères. (Réalisée d'après (Porubsky et al., 2012)).

Les glycolipides sont très rares dans le règne animal. Les glycosphingolipides sont composés d'une partie hydrophobe appelé céramide, et d'une partie hydrophile composée d'un monosaccharide ou oligosaccharide lié par un groupe hydroxyle à la céramide. La partie glucidique des glycolipides est synthétisée dans l'appareil de Golgi où des glycosyltransférases spécifiques catalysent l'ajout des monosaccharides. Il existe trois groupes majeurs : (1) les ganglio- et isoganglio (2) les lacto- et néolacto- (3) les globo et isoglobo- séries (Figure 12). Les glycolipides, localisés dans la membrane externe, ont des rôles variés dans les interactions cellule/cellule et cellule/matrice (Maccioni et al., 2011; Spessott et al., 2012).

#### 2.1.4. Les antigènes des groupes sanguins

Le système ABO a été découvert en 1900 par Karl Landsteiner et le système Lewis en 1940 par Mourant. Ces deux systèmes sanguins comportent des déterminants antigéniques de nature glycanique. Ils sont présents à la surface des cellules sanguines mais également à la surface de nombreux tissus.

La biosynthèse des antigènes des groupes sanguins se fait à partir de quatre précurseurs. Elle nécessite la présence de glycosyltransférases qui catalysent l'ajout de monosaccharides à différentes étapes de la synthèse. Il existe 4 types de précurseurs :

-type 1 : Gal $\beta$ 1-3GlcNAc $\beta$ 1-R

-type 2 : Gal $\beta$ 1-4GlcNAc $\beta$ 1-R

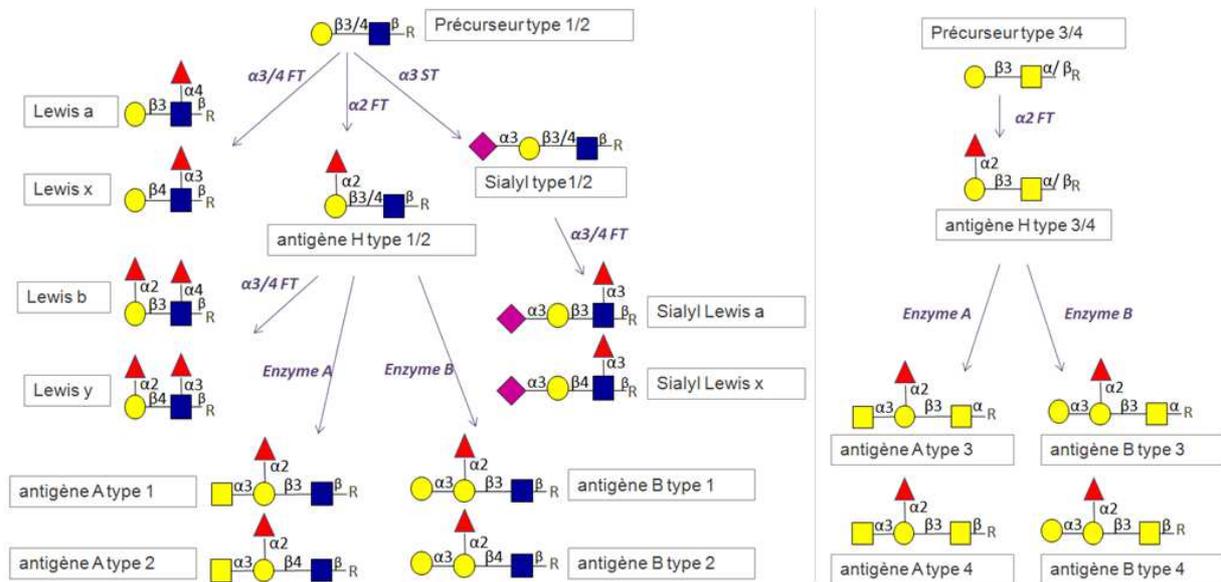
-type 3 : Gal $\beta$ 1-3GalNAc $\alpha$ 1-R

-type 4 : Gal $\beta$ 1-3GalNAc $\beta$ 1-R

Les précurseurs de type 1 et 2 sont présents à la surface de protéines *N*- et *O*-glycosylées et à la surface des glycolipides de la série lacto (représenté par la lettre R). Le précurseur de type 3 est rencontré exclusivement sur les *O*-glycanes et le précurseur de type 4 sur les surfaces des glycolipides des séries globo et ganglio (Marionneau et al., 2001).

L'addition d'un fucose en  $\alpha$ 1-2 par une fucosyltransférase sur le Gal des quatre précurseurs permet la synthèse de l'antigène H dont les porteurs auront le groupe sanguin O. La synthèse de l'antigène A va se faire par l'ajout d'un résidu GalNAc par une N-acétylgalactosaminyltransférase sur le Gal de l'antigène H. Tandis que la synthèse de l'antigène B résulte de l'addition d'un résidu Gal par une galactosyltransférase sur le Gal de l'antigène H. L'ajout d'un fucose en  $\alpha$ 1-4 sur le GlcNAc du précurseur de type 1/2 ou de l'antigène H de type 1/2 catalyse la synthèse du Lewis a et du Lewis b, respectivement. Lorsque la liaison se fait en  $\alpha$ 1-3 sur le GlcNAc du précurseur de type 1/2 ou de l'antigène H de type 1/2, c'est le Lewis x et Lewis y qui sont synthétisés. La synthèse du Sialyl Lewis a et du Sialyl Lewis x se

fait par l'ajout d'un fucose en  $\alpha$ 1-3 et  $\alpha$ 1-4, respectivement, sur le GlcNAc du précurseur de type 1/2, qui possède un acide sialique lié en  $\alpha$ 1-3 sur le résidu Gal (Figure 13) (Marionneau et al., 2001).



**Figure 13: Voies de biosynthèse majeures des groupes sanguins ABH et Lewis.** Les résidus Gal sont représentés par un rond jaune les GlcNAc par un carré bleu, les NeuAc par un losange violet, et le Fuc par un triangle rouge. (Réalisée d'après (Marionneau et al., 2001)).

### 2.1.5. Modification de la glycosylation chez les patients CF

Chez les patients atteints de la mucoviscidose, il a été observé une modification de la glycosylation. Leurs glycoconjugués possèdent une altération au niveau des monosaccharides en position terminale des branches oligosaccharidiques. Une augmentation des fucoses liés en  $\alpha$ 1, 3/4 et  $\alpha$ 1,6 à un GlcNAc est constatée. L'augmentation de la fucosylation a été observée sur des glycoconjugués qui proviennent de différents types cellulaires (poumons, glandes salivaires, intestins) (Rhim et al., 2001). Les glycoprotéines des cellules des voies aériennes (excepté les mucines) comportent moins d'acide sialique par rapport à des cellules de patients non CF (Rhim et al., 2001). Une plus forte concentration de glycolipides asialo GM1 est observée pour les cellules de l'épithélium respiratoire des patients CF par rapport aux non CF (de Bentzmann et al., 1996).

Les mucines des voies respiratoires, présentent une surexpression de composants sialylés et sulfatés. Une étude a mis en évidence que les épitopes 6-sulfo-sialyl-Lewis x et sialyl-Lewis x présents en abondance permettaient la fixation de plusieurs souches de *P. aeruginosa*. Ces dernières ont une affinité plus forte pour l'épitope 6-sulfo-sialyl Lewis x (Scharfman et al., 2000).

La modification de la glycosylation favorise la colonisation des voies respiratoires par les pathogènes et ainsi l'infection, mais le processus d'infection va amplifier l'altération de la glycosylation. Des mucines qui proviennent de patients CF infectés par *P. aeruginosa* contiennent plus d'épitopes sialyl-Lewis x par rapport aux mucines qui appartiennent aux patients CF non infectés (Davril et al., 1999). Cette modification est due aux cytokines pro inflammatoires TNF qui induisent l'expression de la sialyltransférase ST3Gal IV impliquée dans la synthèse du sialyl-Lewis x (Colomb et al., 2012). Ces modifications favorisent la fixation de *P. aeruginosa* car les glycoconjugués ainsi modifiés sont reconnus par des récepteurs de la bactérie : les lectines.

## 2.2. Les lectines

A la fin du XIX<sup>ème</sup> siècle, des protéines capables d'agglutiner naturellement des hématies ou érythrocytes sont découvertes et appelées hémagglutinines. La capacité de ces hémagglutinines à reconnaître spécifiquement différents groupes sanguins humains mena Boyd et Shapleigh en 1954 à nommer ces hémagglutinines des lectines, du latin *legere*, qui signifie « sélectionner » (Sharon and Lis, 2004).

Aujourd'hui sont appelées lectines toutes protéines d'origine non immunitaire capables de reconnaître et de se fixer spécifiquement sur des structures oligosaccharidiques de manière non covalente, réversible et sans les modifier (Sharon and Lis, 2004).

A quelques exceptions près, les lectines ne présentent pas une très grande affinité pour leurs ligands. Cette faible affinité, avec une constante d'affinité de l'ordre du millimolaire (mM), est souvent compensée par la multivalence de ces protéines soit par la présence de plusieurs sites de fixation aux sucres. Les lectines sont capables de reconnaître des monosaccharides mais aussi des sucres plus complexes de nature oligosaccharidique.

Cette reconnaissance protéines-glucides menée par les lectines permet de nombreux phénomènes de reconnaissance dans l'adhésion cellulaire. Les lectines sont présentes dans tout le règne vivant avec une grande diversité structurale. Plus de 1096 structures cristallographiques sont répertoriées dans une base de données développée au laboratoire (<http://lectin3d.cermav.cnrs.fr/>). Le rôle des lectines peut être très varié. Chez les bactéries, elles sont utilisées pour la reconnaissance et l'adhésion du pathogène à la cellule hôte. Elles sont classées en trois familles : les adhésines (lectines associées aux organelles telles les pili ou les flagelles), les lectines associées à des toxines et les lectines solubles.

*P. aeruginosa* code pour un arsenal de lectines aux spécificités variées qui définit sa stratégie d'adhésion ou glycostratégie. Elle se compose de deux lectines solubles, PA-IL et PA-IIL,

deux adhésines flagellaires, FliC et FliD, et de plusieurs adhésines fimbriales associées à des pili (Figure 14).

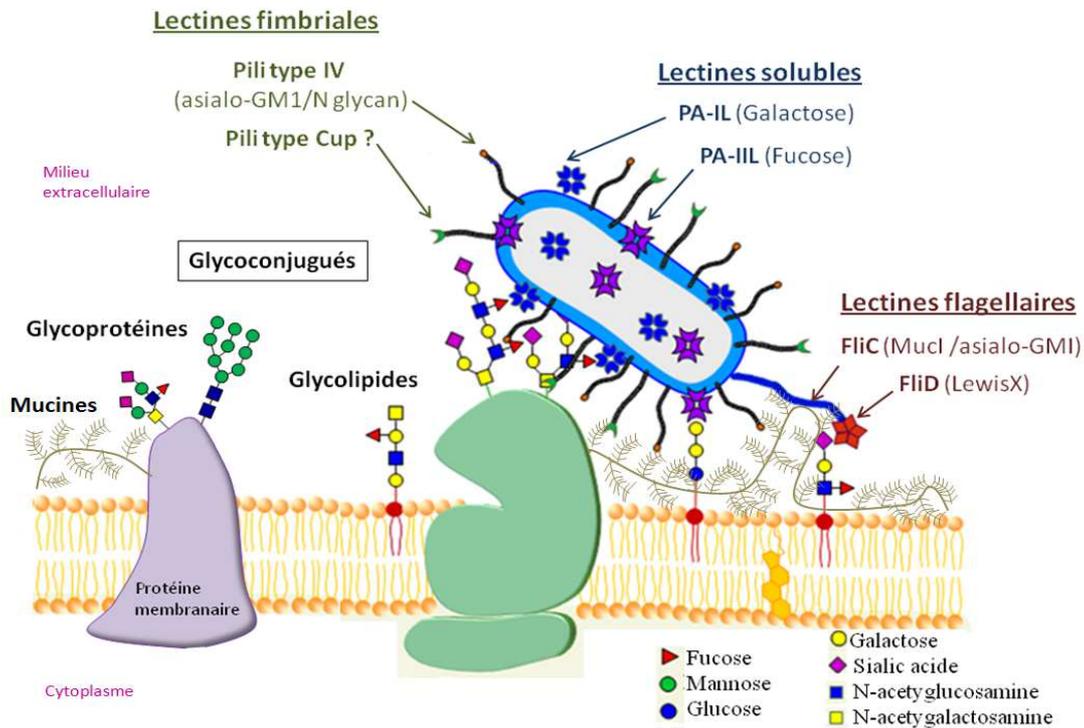


Figure 14: Processus d'adhésion de *P. aeruginosa* par des interactions lectines/sucres.

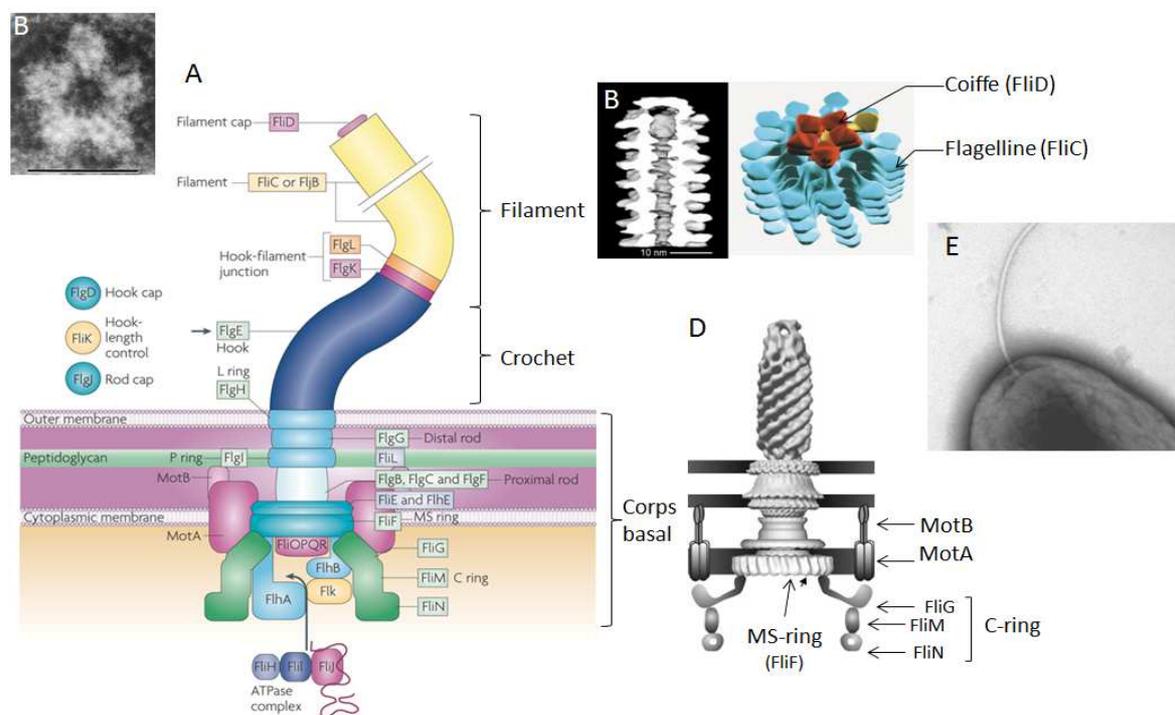
## 2.2.1. Les lectines flagellaires

### 2.2.1.1. Un des rôles du flagelle : initiation de l'adhésion

Le flagelle est un appendice extracellulaire d'environ 15 à 20  $\mu\text{m}$  qui sert à la nage de la bactérie. Cet appendice lui permet également de s'orienter vers des supports biotiques ou abiotiques pour initier son adhésion. *P. aeruginosa* possède un flagelle unique à une de ces extrémités (Figure 15). Selon les différents stimuli chimiotactiques présents dans l'environnement, il adapte sa nage. Le flagelle permet un déplacement en milieu liquide ou sur agar mou (0,3 %) par la nage « Swimming » et sur agar (0,5-1 %) par l'essaimage « Swarming ». Par sa mobilité et un attachement réversible, le flagelle va explorer la surface afin de sélectionner l'environnement le plus propice à la croissance de la bactérie. L'attache du flagelle rapproche la bactérie de la surface ce qui favorise l'interaction avec d'autres structures et l'installation d'une adhésion irréversible. Durant cette étape, la mobilité flagellaire est réprimée par plusieurs mécanismes de régulation puis par la suite la bactérie va perdre son flagelle. En initiant l'adhésion de la bactérie à la surface, le flagelle joue un rôle important dans l'installation du biofilm.

## 2.2.1.2. Structure du flagelle

Le flagelle est une structure complexe qui comporte plus de 41 protéines. Le flagelle est organisé en trois parties distinctes : le corps basal, le crochet et le filament (Chevance and Hughes, 2008) (Figure 15).



**Figure 15: Structure du flagelle de *P. aeruginosa*.** A) Représentation schématique de la structure du flagelle (Chevance and Hughes, 2008). B) Image en microscopie électronique de la lectine FliD. La barre représente 10 nm. C) Représentation graphique de l'extrémité du filament et D) du corps basal du flagelle (Brown et al., 2011). E) Image en microscopie électronique de *P. aeruginosa* monoflagellé.

Ancrés dans l'enveloppe de la bactérie, le corps basal est la première partie à être assemblée. L'anneau C est le rotor du flagelle, associé à l'anneau MS, est capable de tourner dans les deux sens et permet à la bactérie de changer de sens. L'énergie nécessaire est fournie par deux pompes à protons, MotAB et MotCD. Une pompe à sodium MotY est nécessaire au fonctionnement de MotAB (Brown et al., 2011) (Figure 15 D). Une fois le corps basal assemblé, les protéines qui constituent le crochet et le filament sont secrétées puis polymérisées à l'extérieur. Le crochet (FlgD et FlgE) et les protéines de jonction (FlgK et FlgL) permettent de transmettre le mouvement du corps basal au filament. Le filament, de structure hélicoïdale, est un cylindre creux et rigide composé de nombreux monomères de la flagelline FliC au bout duquel est présente une coiffe composée de la protéine FliD. La protéine FliD est nécessaire à la polymérisation de la flagelline (Ikeda et al., 1996).

### 2.2.1.3. Les lectines FliD et FliC

La protéine FliD permet l'adhésion précoce de *P. aeruginosa* dans les voies respiratoires des patients atteints de la mucoviscidose. Des mutations dans le gène de FliD ne permettent plus la fixation de *P. aeruginosa* sur des mucines qui proviennent de patients CF. En apportant le gène FliD en trans, le phénotype d'adhésion est restauré (Arora et al., 1998). Chez *P. aeruginosa*, il existe deux types de coiffes et de flagellines, le type A pour la souche PAK et apparenté et le type B dans la souche PAO1 et apparenté (Arora et al., 2000). Une étude a mis en évidence que FliD de type B permettait l'adhésion de *P. aeruginosa* aux mucines par la reconnaissance des glycoconjugués qui portent les épitopes Lewis x ou sialyl-Lewis x. FliD de type A ne présente pas de spécificité pour le LewisX, le sialyl-Lewis x ou le sulfosialyl-Lewis x. La flagelline FliC de type b reconnaît le Lewis x présent sur les mucines MUC1 (Lillehoj et al., 2002) ainsi que les glycolipides GM1, GD1a et asialoGM1 (Feldman et al., 1998). Récemment, il a été identifié que le flagelle (FliC-FliD) était capable de se fixer sur le domaine basolatéral des cellules épithéliales de poumons polarisées par la reconnaissance de l'HSPGs (Heparan sulfate chains of proteoglycans) (Bucior et al., 2012).

### 2.2.2. Les lectines solubles

Par des expériences d'hémagglutination sur des érythrocytes traités à la papaïne en présence d'extraits de *P. aeruginosa*, Gilboa-Gardner a découvert en 1972 la lectine PA-IL puis en 1977 la lectine PA-III. Les protéines PA-IL et PA-III sont respectivement codées par les gènes *lecA* et *lecB* qui sont sous le contrôle du Quorum-sensing, un système de communication entre les bactéries qui repose production de petites molécules chimiques. Les deux protéines sont principalement présentes dans le cytoplasme mais également à la surface de la membrane externe de la bactérie. Elles ne présentent pas d'identité de séquence mais forment toutes les deux des tétramères en solution et nécessitent la présence de calcium pour être actives.

#### 2.2.2.1. La lectine PA-IL

##### 2.2.2.1.1. Le rôle de PA-IL

La protéine PA-IL joue un rôle dans la stabilisation du biofilm mais elle n'est pas essentielle à sa formation. Des études statiques d'épaisseur de biofilm sur du polystyrène avec une souche dont le gène *lecA* est délété permet la formation d'un biofilm plus fragile, moins couvrant et moins épais (Diggle et al., 2006). Cette expérience montre que la forme tétramérique de PA-IL pourrait permettre de relier les bactéries entre elles. PA-IL pourrait avoir un rôle dans

l'adhésion car elle interagit avec la fibronectine présente dans la matrice extracellulaire et qui est glycosylée. L'ajout de monosaccharides permet d'inhiber l'adhésion de *P. aeruginosa* à la fibronectine (Rebiere-Huët et al., 2004). De plus, PA-IL se fixe au glycocalix (Mewe et al., 2005).

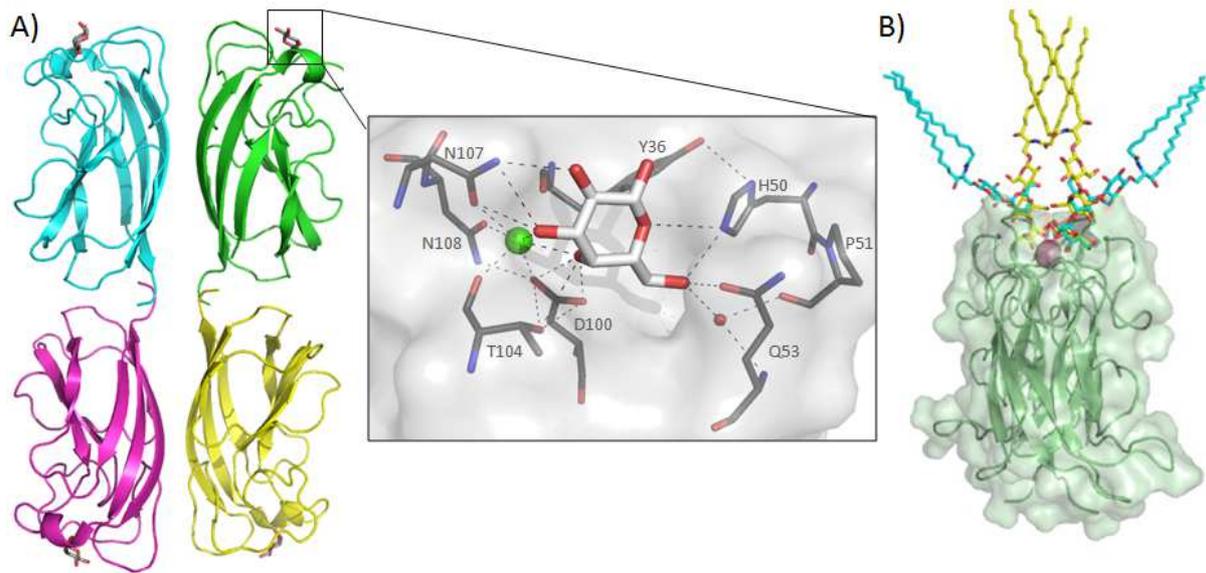
Bien que les fonctions de cette lectine ne soient pas entièrement connues, PA-IL constitue un facteur de virulence notamment par sa cytotoxicité. Au niveau des voies aériennes, PA-IL en synergie avec PA-III, bloquent les battements ciliaires impliqués dans la clairance mucociliaire (Mewe et al., 2005). PA-IL inhibe la croissance des cellules respiratoires et à une forte concentration les cellules sont détruites (Bajolet-Laudinat et al., 1994). En combinaison avec l'exotoxine ou l'élastase, PA-IL détruit les jonctions serrées entraînant une désorganisation et la destruction des cellules. Des travaux utilisant des modèles murins ont permis de montrer la diminution de la destruction de la barrière épithéliale au niveau des capillaires alvéolaires lorsque la souche n'exprime plus PA-IL (Chemani et al., 2009).

#### 2.2.2.1.2. La spécificité de PA-IL

La lectine PA-IL possède une grande spécificité pour les galactosides. La constance d'association ( $K_a$ ) de PA-IL pour le D-galactose est de  $3,6 \cdot 10^{-4} \text{ M}^{-1}$  (Garber et al., 1992). Une étude basée sur des tests d'hémagglutination a montré que PA-IL interagit avec les antigènes du groupe sanguin ABO, avec une plus grande affinité pour le groupe B ( $\text{Gal}\alpha 1-3(\text{Fuc}\alpha 1-2)\text{Gal}\beta 1-4\text{GlcNAc}\beta 1-3\text{Gal}\beta 1-\text{Glc}\beta 1-1\text{Cer}$ ). PA-IL a également une bonne affinité pour les antigènes du groupe P<sup>K</sup> ( $\text{Gal}\alpha 1-4\text{Gal}\beta 1-4\text{Glc}\beta 1-1\text{Cer}$ ) et P1 ( $\text{Gal}\alpha 1-4\text{Gal}\beta 1-4\text{GlcNAc}\beta 1-3\text{Gal}\beta 1-4\text{Glc}\beta 1-1\text{Cer}$ ) (Gilboa-Garber et al., 1994). Cette lectine est donc capable de se lier de manière efficace aux oligosaccharides qui présentent les épitopes du type :  $\text{Gal}\alpha 1-2\text{Gal}$ ,  $\text{Gal}\alpha 1-3\text{Gal}$ ,  $\text{Gal}\alpha 1-4\text{Gal}$  et  $\text{Gal}\alpha 1-6\text{Gal}$ . Une récente étude a montré que l'antigène globotriaosylcéramides Gb3 (glycolipide) était un ligand naturel de PA-IL. Le Gb3 est présent dans de nombreux organes (rein, cœur) ainsi que sur les érythrocytes (antigène P<sup>K</sup>). La constance d'association ( $K_a$ ) de PA-IL pour le trisaccharide du Gb3 ( $\text{Gal}\alpha 1-4\text{Gal}\beta 1-4\text{Glc}$ ) est de  $68-77\mu\text{M}$  (Blanchard et al., 2008).

#### 2.2.2.1.3. PA-IL : $\beta$ -sandwich à un calcium

PA-IL est un homotétramère et sa structure révèle un repliement de type  $\beta$ -sandwich. Chaque monomère présente à son apex un site de fixation spécifique au galactose qui contient un ion calcium (Figure16).



**Figure 16: Structure de PA-IL.** A) Structure cristallographique du tétramère PA-IL complexé avec le galactose. L'ion calcium nécessaire à la fixation du galactose est représenté par une sphère verte et la molécule d'H<sub>2</sub>O par une sphère rouge. B) Modèle représentant la superposition des docking du  $\alpha$ Gal1-3 $\beta$ Gal1-4Glc-cer (en bleu) et du  $\alpha$ Gal1-4 $\beta$ Gal1-4 $\beta$ Glc-cer (en jaune) avec PA-IL (Blanchard et al., 2008).

La structure de PA-IL avec le D-galactose montre que tous les hydroxyles du sucre, à l'exception de l'hydroxyle O1, forment des liaisons avec la protéine. Les hydroxyles O3 et O4 participent à la coordination du calcium. Des liaisons hydrogènes entre les hydroxyles O4, O5 et O6 du sucre et les acides aminés Tyr36, His50, Gln53 situés sur une boucle de surface permettent le maintien du sucre. Une molécule d'eau renforce la position du sucre en créant un pont par un réseau de liaisons hydrogènes entre l'hydroxyle O6 et les acides aminés Pro51 et Gln53 (Figure 16) (Cioci et al., 2003).

La structure de PA-IL complexé avec le trisaccharide de l'isoglobotriaosylcéramide l'iGb3 ( $\alpha$ Gal1-3 $\beta$ Gal1-4Glc), a permis d'établir un modèle de reconnaissance entre PA-IL et son ligand naturel Gb3 (Figure 16 B) (Blanchard et al., 2008).

## 2.2.2.2. La lectine PA-IIL

### 2.2.2.2.1. Le rôle de PA-IIL

PA-IIL qui est localisée sur la membrane externe de la bactérie joue un rôle dans la formation du biofilm en se fixant sur les glycoconjugués localisés à la surface des bactéries. Une souche déficiente de PA-IIL n'est pas capable de former un biofilm contrairement à la souche sauvage (Tielker et al., 2005). La lectine PA-IIL comme PA-IL est capable d'inhiber le battement ciliaire des cellules pulmonaires (Mewe et al., 2005). Une étude réalisée sur des

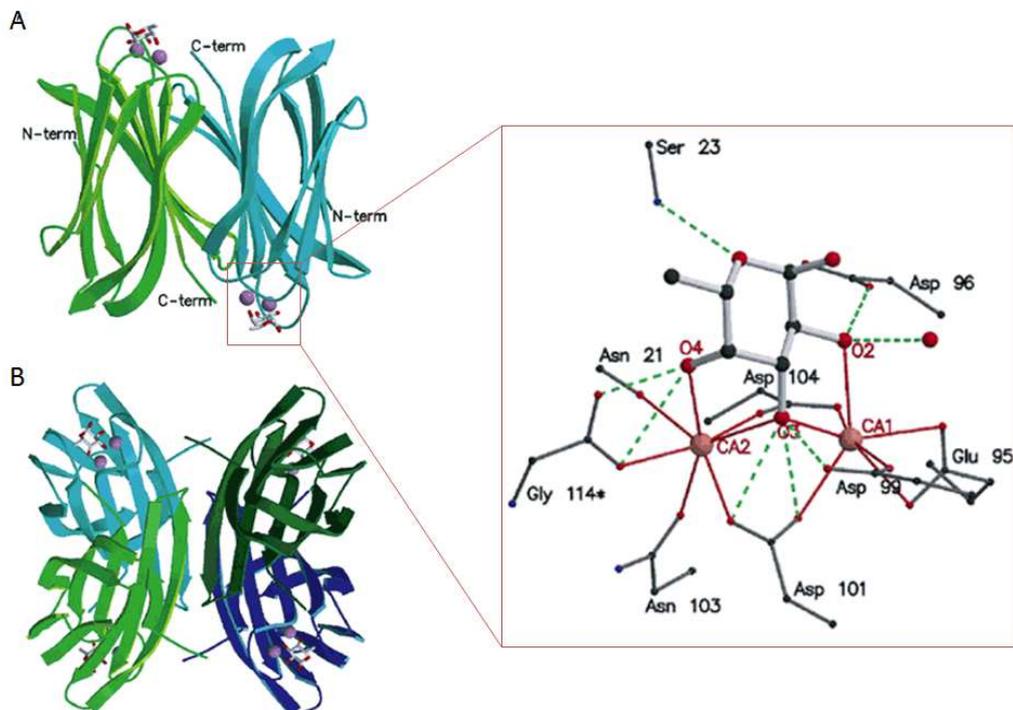
modèles murins a permis de mettre en évidence qu'une souche délétée de PA-IIL, entraîne une diminution de la perméabilité de la barrière alvéolaire (Chemani et al., 2009).

#### 2.2.2.2.2. Spécificité de PA-IIL

La lectine PA-IIL possède une très forte affinité avec le L-fucose (constance d'affinité de l'ordre de  $1.6.10^6 \text{ M}^{-1}$ ) rarement rencontrée dans les interactions lectines/sucres et ceci grâce à la présence de deux ions calcium dans le site fixation (Avichezer and Gilboa-Garber, 1987). PA-IIL reconnaît également mais avec une plus faible affinité le L-galactose > D-arabinose > D-fructose > D-mannose (Imberty et al., 2004). Le trisaccharide Lewis a ( $\text{Gal}\beta 1-3(\text{Fuc}\alpha 1-4)\text{GlcNAc}$ ) est le meilleur ligand naturel de PA-IIL (Wu et al., 2006).

#### 2.2.2.2.3. PA-IIL : $\beta$ -sandwich à deux calciums

PA-IIL est un homotétramère où chaque monomère est composé de neuf brins  $\beta$  antiparallèles en sandwich  $\beta$ . La dimérisation de PA-IIL se fait par l'interaction de deux monomères en tête bêche par l'intermédiaire de feuillet  $\beta$ . Deux dimères forment un tétramère par l'association antiparallèle d'un feuillet  $\beta$  de chaque monomère (Figure 17).



**Figure 17: Structure de la lectine PA-IIL.** A) Représentation d'un dimère de PA-IIL complexé au fucose. Les deux atomes de calcium sont représentés par une sphère rose, et le fucose par des bâtons. Les interactions mises en place entre le sucre et le site de reconnaissance de PA-IIL figurent dans l'encadré. Les traits pleins rouges montrent les liaisons de coordination et les traits en pointillés vert

représentent les liaisons hydrogènes. La sphère rouge schématise la molécule d'eau. B) Tétramère de PA-III. (Mitchell et al., 2002).

La structure de PA-III complexée au fucose met en évidence une poche acide comportant les deux ions calcium. L'extrémité C-terminale du monomère adjacent participe également à la coordination d'un des deux atomes de calcium (résidu Gly114). Trois des hydroxyles du fucose ainsi que les chaînes latérales des acides aminés Glu95, Asp99, Asp101, Asn103 and Asp104 participent à la coordination des atomes de calcium. Les hydroxyles O2 et O4 interagissent avec un atome de calcium chacun tandis que l'hydroxyle O3 avec les deux (Figure 17) (Mitchell et al., 2002).

### **2.2.3. Les lectines fimbriales**

#### 2.2.3.1. Les pili de type IV

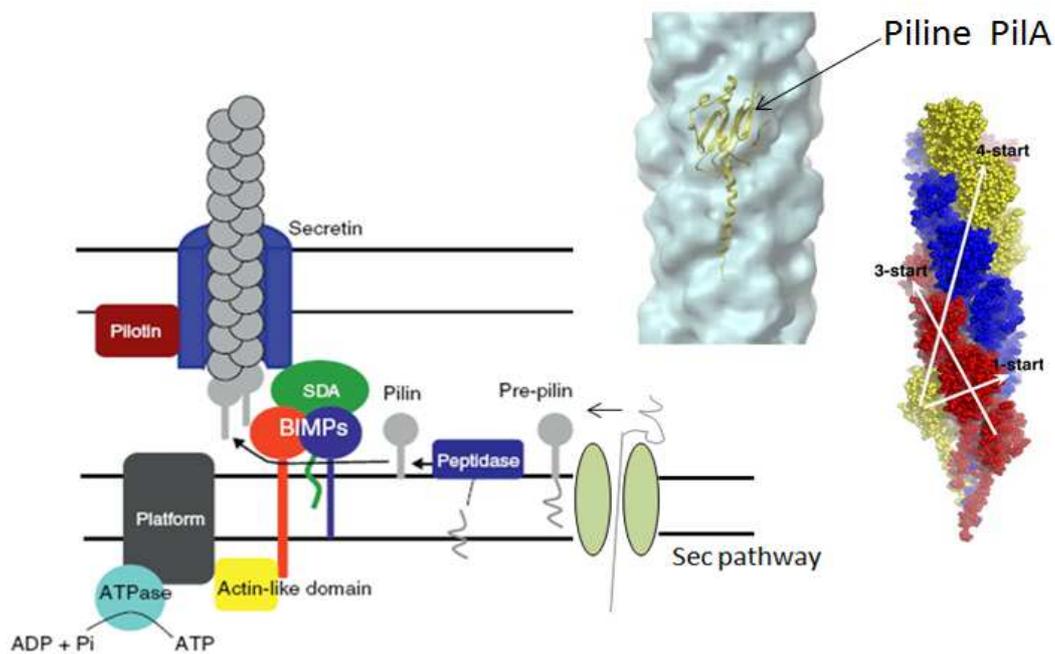
##### *2.2.3.1.1. Le rôle des pili de type IV*

Les pili de type IV, sont localisés au pôle et mesurent plusieurs micromètres de long avec 6-9 nm de diamètre. *P. aeruginosa* est capable d'assembler des pili de type IVa et de type IVb. Ils sont impliqués dans l'adhésion irréversible de la bactérie à la surface par la formation de microcolonies. Cette étape qui constitue la première étape de formation du biofilm est rendue possible par la mobilité et la capacité adhésive des pili de type IV. Leur mobilité par « twitching » ne permet pas à la bactérie de se déplacer mais c'est un moyen rapide pour coloniser un milieu et d'établir l'initiation de la formation du biofilm en formant des microcolonies. La « twitching motility », régulée par des régulateurs transcriptionnels et le chimiotactisme, favorise le processus d'adhésion.

##### *2.2.3.1.2. L'assemblage des pili de type IV*

Les pili de type IVa résultent de l'oligomérisation de la piline PilA de manière hélicoïdale. Lorsque la pré-piline PilA est insérée dans la membrane interne par le système Sec, elle est maturée par une pré-piline peptidase avant sa polymérisation non-covalente en pilus. La membrane interne comprend un complexe formé d'une plateforme de protéines qui interagit avec une protéine cytoplasmique contenant un domaine « actin-like », une ATPase qui permet de fournir de l'ATP nécessaire à la formation du pilus et à sa rétractation, puis avec des protéines BIMPs (bitopic inner membrane proteins). Par un processus d'assemblage encore peu clair mais qui nécessite de l'ATP, les pilines PilA sont assemblées au niveau de la sécrétine qui forme un pore. Les protéines SDA (secretin dynamic-associated protein) forment

un complexe avec les protéines BIMPs pour créer un conduit permettant le passage du pilus au travers des membranes (Figure 18).

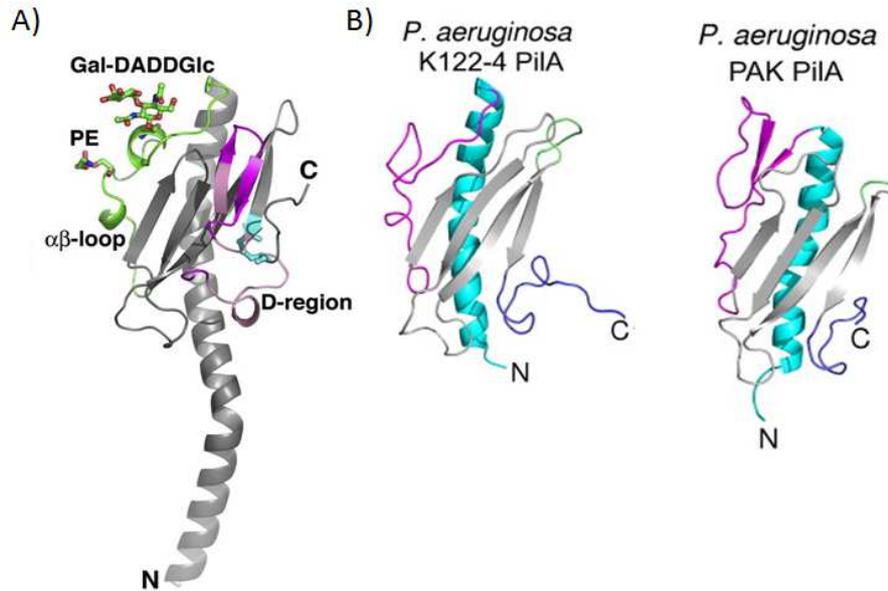


**Figure 18: Modèle d'assemblage des pili de type IVa et représentation de la disposition des pilines PilA polymérisées.** (Réalisée d'après (Craig et al., 2006)).

### 2.2.3.1.3. La lectine PilA

La première structure d'une sous-unité piline de type IV provient de la souche MS11 de *Neisseria gonorrhoeae* (Parge et al., 1995). Toutes les pilines sont constituées d'une hélice  $\alpha$  N-terminale, hydrophobe et très conservée, et d'une tête globulaire. La tête globulaire comporte quatre feuillets  $\beta$  antiparallèles, une boucle qui relie l'hélice  $\alpha$  au feuillet  $\beta$ , puis une boucle DSL (disulfide-bonded loop) en C-terminal (Figure 19) (Nguyen et al., 2010).

Les structures de PilA dans des souches PAK et K122-4 de *P. aeruginosa* ont été résolues mais n'ont pas permis de mettre en évidence un site de reconnaissance à un carbohydrate. Une étude a montré que PilA reconnaît et se fixe au glycolipide asialo-GM1 (Gupta et al., 1994). Puis récemment il a été identifié que les pili de type IVa sont également capables de se fixer aux *N*-glycanes présents au domaine apical des cellules épithéliales de poumons polarisées (Bucior et al., 2012) (Figure 19).



**Figure 19: Structures cristallographiques de différentes sous-unités PilA.** A) Piline MS11 de *N.gonorrhoeae* complexé avec le Gal-DADDGlc. Le carbohydrate interagit au niveau de la  $\alpha\beta$ -loop de PilA (Parge et al., 1995). B) Structures tronquées de K122-4 et PAK PilA (Nguyen et al., 2010).

#### 2.2.3.2. Les pili de type « CU »

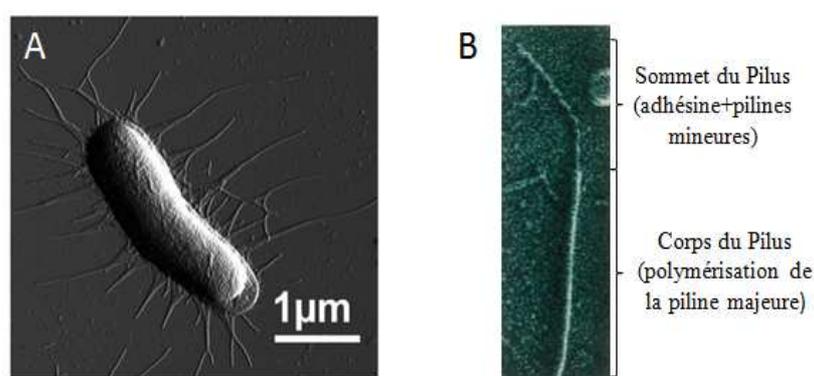
Les pili de type CU sont assemblés par le système « chaperone-Usher » et présentent à leur apex une adhésine. Tous leurs éléments sont codés dans un opéron et plusieurs ont été identifiés dans le génome de *P. aeruginosa*. Ils sont impliqués à des degrés et des stades différents dans la formation du biofilm et seront décrits plus en détails dans la partie suivante.

---

### 3. La voie « chaperonne-usher » (CU)

---

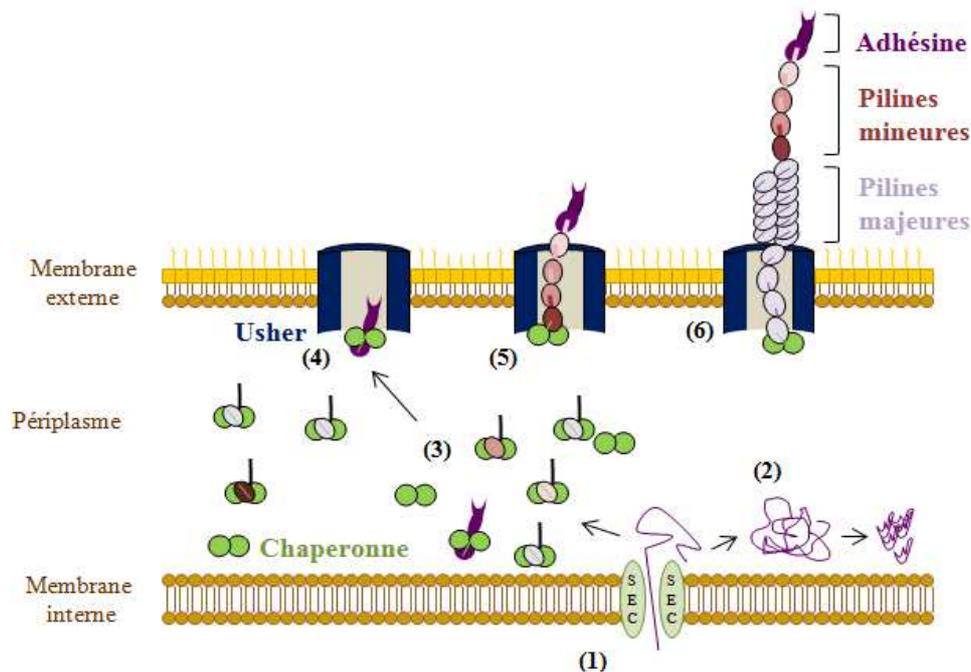
Les pili assemblés par la voie « chaperonne-usher » (CU pour « Chaperonne-Usher») sont uniquement présents chez les bactéries à Gram négatif et forment le groupe d'appendices bactériens le plus abondant. Ces pili sont considérés comme un facteur de virulence car ils jouent un rôle direct dans l'adhésion de la bactérie aux cellules hôtes et dans la formation du biofilm. Les gènes impliqués dans la biosynthèse des pili de type CU sont organisés au sein d'un opéron qui code au minimum pour trois composants différents : (1) une chaperonne périplasmique ; (2) une protéine de membrane externe appelée « usher » qui forme un pore de translocation et (3) au moins une sous-unité structurale. Il existe trois types de sous-unités structurales, les pilines majeures, les pilines mineures et les adhésines. Les pilines majeures, synthétisées en grande quantité, forment le corps du pilus (figure 20). Les pilines mineures, présentes en petite quantité, constituent la partie flexible du pilus. Enfin, une sous-unité adhésine qui est composée de deux domaines chez *P. aeruginosa* - un domaine piline et un domaine lectine - est située à l'extrémité du pilus. Le domaine piline va permettre l'ancrage de l'adhésine au pilus alors que le domaine lectine va permettre la reconnaissance spécifique et la fixation du pathogène aux glycoconjugués de l'hôte (figure 20). Différentes structures de pili CU ont été identifiées allant de la simple exposition d'une adhésine à la surface bactérienne à l'assemblage d'un pilus composé de pilines majeures avec ou sans la présence de pilines mineures et/ou d'adhésine (Thanassi et al., 1998).



**Figure 20: Photo de microscopie électronique A) d'une souche d'*E.coli* qui exprime des P pili à sa surface et B) d'un pilus P (D'après (Hung et al., 2002; Klinth et al., 2012)).**

Les sous-unités sont synthétisées dans le cytoplasme puis exportées dans le périplasm via la machinerie de translocation Sec (Figure 21). Dans le périplasm, elles sont immédiatement

prises en charge par une chaperonne qui a pour rôle de stabiliser les sous-unités structurales afin d'éviter leur polymérisation précoce dans le périplasme. Une fois que les complexes chaperonne/sous-unités sont formés, ils sont recrutés selon un ordre précis par la protéine « usher » également appelée protéine « trieuse ». La protéine « usher » va constituer une plateforme où le complexe chaperonne/sous-unité se dissocie et où les sous-unités pilines s'assemblent. La polymérisation des sous-unités en pilus ne nécessite pas d'ATP (Sauer et al., 2000).



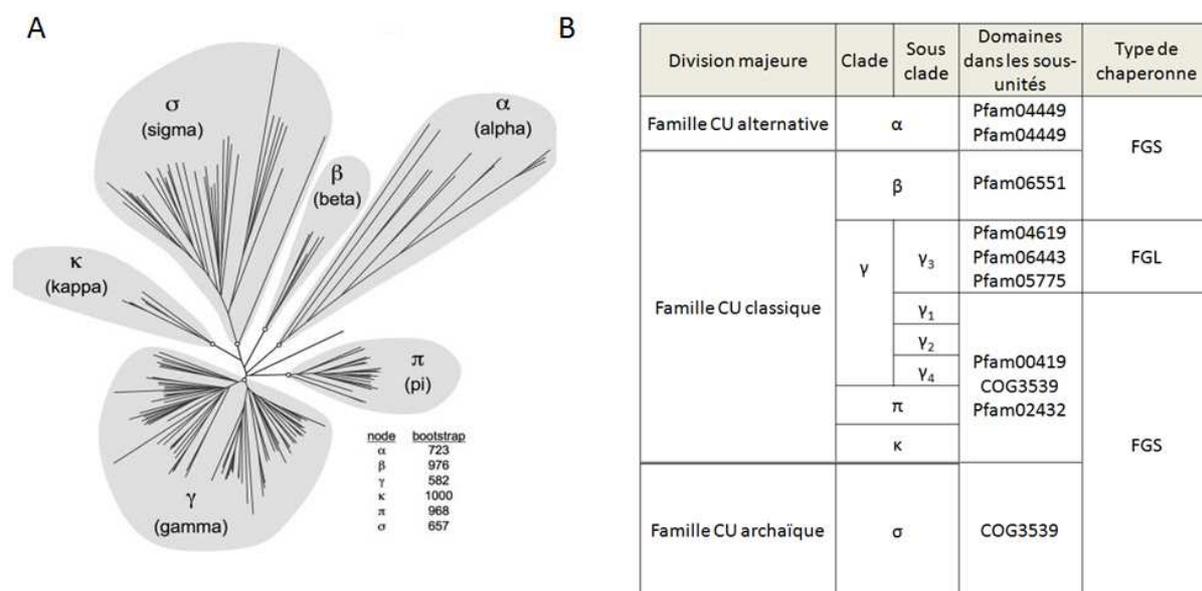
**Figure 21: Modèle de la synthèse de fimbriae par la voie « chaperonne usher ».** (1) Les sous-unités sont transloquées dans le périplasme par le système Sec. (2) En absence de chaperonne les sous-unités s'agrègent et sont dégradées. (3) Les sous-unités sont prises en charge par une chaperonne ce qui permet de les stabiliser. (4-5-6) Les complexes sous-unité/chaperonne sont recrutés par la protéine « usher » qui permet la polymérisation des sous-unités à la surface de la bactérie. L'adhésine est la première sous-unité assemblée suivie des pilines mineures puis des pilines majeures. (Réalisée d'après (Sauer et al., 2000)).

Les pili CU les plus étudiés et les mieux caractérisés d'un point de vue structural et mécanique sont les pili de type P (ou Pap pour pilus-associated with Pylonephritis) et les pili de type I d'*E. coli*.

### 3.1. La classification des voies CU

La première classification des voies CU comportait deux sous-groupes définis selon la structure des chaperonnes. Les pili dont l'opéron codent pour des chaperonnes à boucle F1-G1 courte (FGS) constituaient le sous-groupe des pili « typiques » composés d'une base

rigide et d'un fimbrillum flexible en partie distale. Les pili dont l'opéron codent pour les chaperonnes à boucle F1-G1 longue (FGL) formaient le groupe des pili adhésifs ne contenant qu'une ou deux sous-unités piline différentes. A l'aide d'une étude phylogénétique basée sur l'analyse de séquences de 189 protéines « usher », Nuccio et Bäumlér ont montré qu'il existait différents clades phylogénétiques pour les pili dont l'opéron code pour une chaperonne FGS. Cette étude phylogénétique a fait naître une nouvelle classification composée de 6 clades majeurs :  $\alpha$ ,  $\beta$ ,  $\gamma$  (divisé en sous-clades  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ ,  $\gamma_4$ ),  $\kappa$ ,  $\pi$  et  $\sigma$  (figure 22) (Nuccio and Bäumlér, 2007). Cette classification tient compte des caractéristiques communes au sein d'un même clade ou sous-clade comme la conservation de domaines dans les sous-unités et les chaperonnes, le type de chaperonne. Les pili *cupA*, *cupB*, *cupC* et *cupD* de *P. aeruginosa* décrits plus bas appartiennent au sous-clade  $\gamma_4$  tandis que les pili *cupE* appartiennent au clade  $\sigma$ . Les P pili et le pili de type I d'*E. coli* font partie des clades  $\pi$  et  $\gamma_1$  respectivement.



**Figure 22:** A) Arbre phylogénétique des protéines « usher ». B) Tableau représentant les différentes divisions, clade, sous-clade, les domaines conservés dans les sous-unités et type de chaperonne (Réalisée d'après (Nuccio and Bäumlér, 2007)).

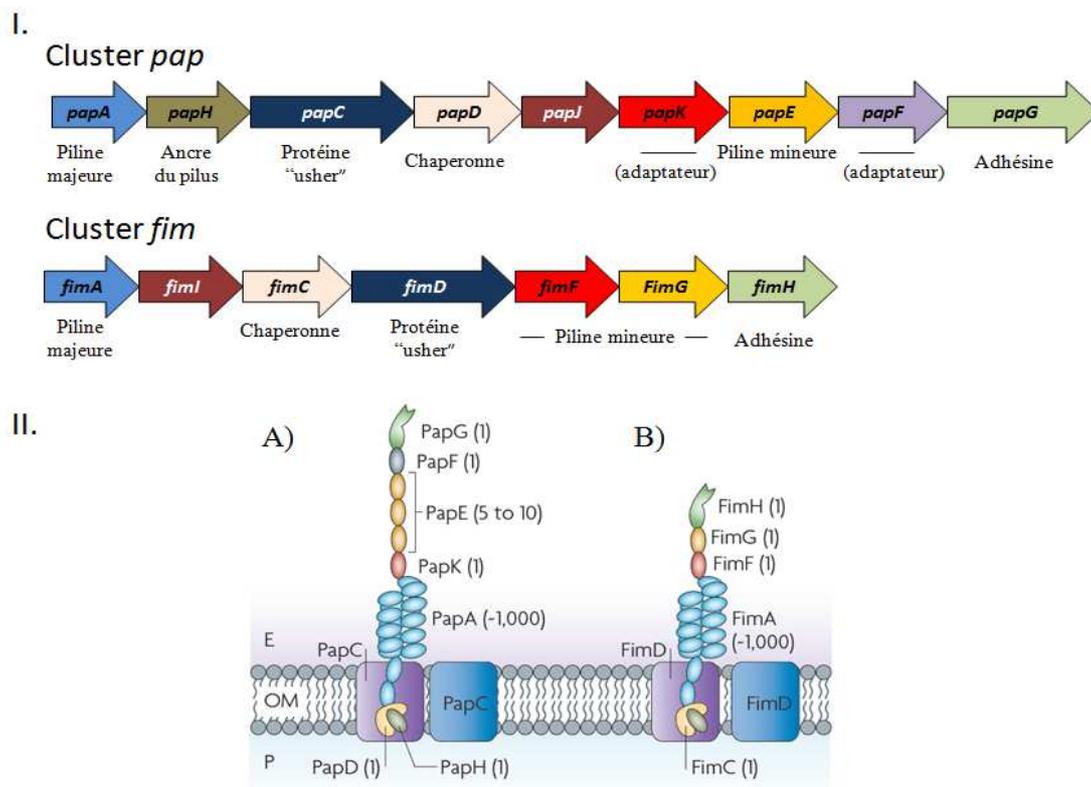
### 3.2. Les pili de type P et type I chez *Escherichia coli*

Les pili de type P et de type I chez *Escherichia coli* sont impliqués dans des infections urinaires. Les pili de type I permettent l'attachement de la bactérie au niveau de la vessie, sa colonisation ainsi que le développement d'un biofilm (Wright and Hultgren, 2006). Les pili

de type P sont retrouvés dans les souches impliquées dans l'établissement de pyélonéphrites (Roberts et al., 1994).

### 3.2.1. Description des pili de type I et P

Les pili de type P et de type I sont codés par les opérons *pap* et *fim* respectivement (figure 23 I.). Les pilines majeures, PapA et FimA, sont oligomérisées en hélice droite ce qui forme le corps rigide du fibrillum (figure 23 II.). L'extrémité flexible du pili de type P est composée de 5 à 10 copies de la piline mineure PapE, de deux pilines mineures adaptatrices PapF et PapK et de l'adhésine PapG. L'extrémité flexible du pili type I est composée des deux pilines mineures FimF et FimG et de l'adhésine FimH. Les sous-unités sont prises en charge par les chaperonnes PapD et FimC. La synthèse du pilus s'effectue par les protéines « usher » PapC et FimD. PapH nommée « sous-unité de terminaison » termine la polymérisation du pilus de type P (Waksman and Hultgren, 2009).



**Figure 23: I. Représentation des clusters *pap* et *fim*. II. A) Structure du pili type P. B) du pili de type I d'*Escherichia coli*.** Les adhésines, en une seule copie au sommet du pilus, sont représentées en vert. Les pilines mineures sont en orange, bleu et rouge. Les pilines majeures sont en bleu. La protéine PapH qui termine la polymérisation du pilus est en gris. Les chaperonnes sont représentées en rose pastel et la protéine de membrane externe « usher » est schématisée par un rectangle violet et bleu. (Waksman and Hultgren, 2009).

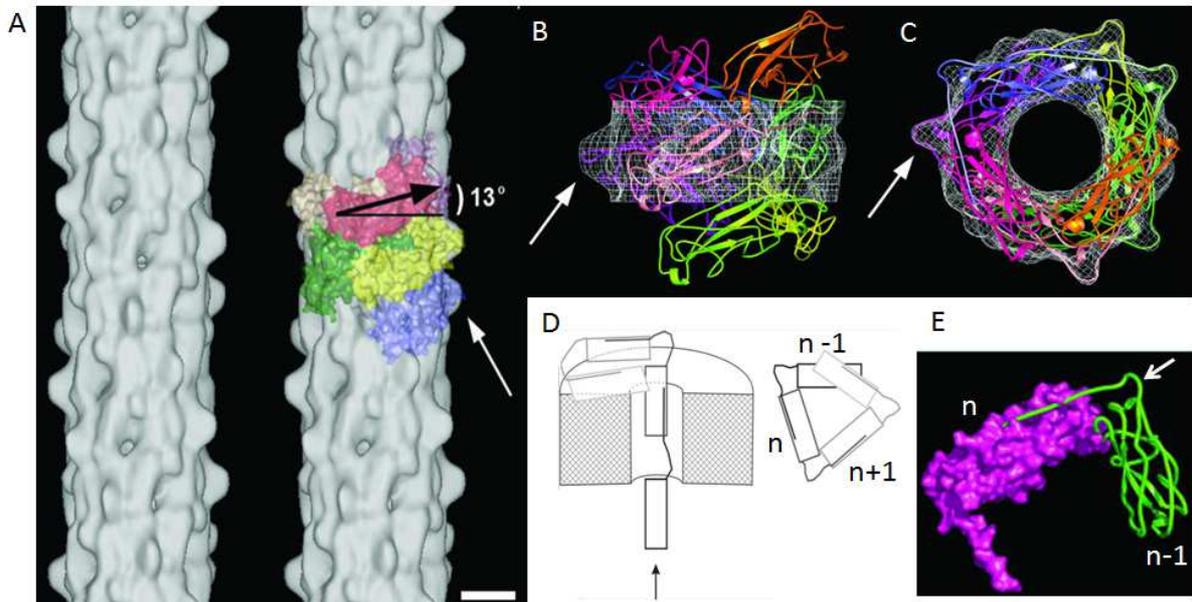
## 3.2.2. Les composants de la voie « chaperonne usher »

### 3.2.2.1. Les sous-unités structurales

Les sous-unités sont composées d'un domaine piline caractérisé par un repliement de type immunoglobuline incomplet. Contrairement à toutes les structures secondaires immunoglobuline, le domaine piline n'est constitué que de 6 brins  $\beta$  antiparallèles A-F. Le brin G C-terminal n'est pas présent et ceci crée un sillon hydrophobe qui déstabilise les pilines non polymérisées. A l'exception des adhésines, toutes les sous-unités pilines possèdent une « extension N-terminale » qui permet la polymérisation du pilus (Figure 24 E).

#### 3.2.2.1.1. Les pilines majeures : le corps rigide du pilus

Les pilines majeures PapA et FimA des pili de type P et I respectivement, forment la partie basale de pilus qui est rigide. Cette partie rigide également appelée « tige » mesure entre 5 et 7  $\mu\text{m}$  de longueur. La tige du P pilus est formée de plus de 1000 pilines majeures PapA polymérisées en hélice droite avec 3, 3 sous-unités PapA par tour d'hélice. Une étude en cryo-microscopie électronique a permis de mesurer le diamètre extérieur qui est de 6,8 nm, le diamètre intérieur qui est de 25 Å et d'identifier la présence de protrusions de 7 Å (Figure 24).

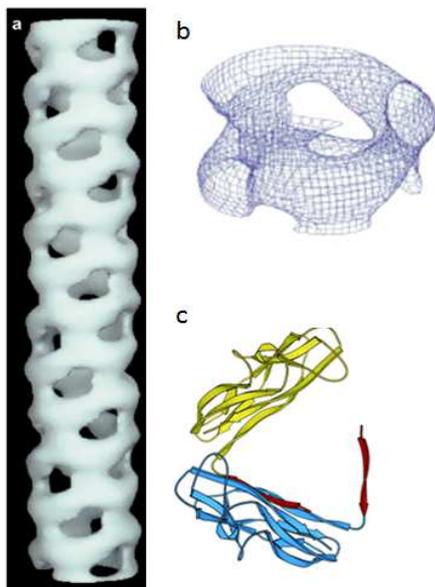


**Figure 24: La tige du pilus de type P d'*Escherichia coli*.** A) Reconstruction 3D de la tige à partir de cryo-microscopie électronique comportant 8 sous-unités PapA. La flèche blanche indique les protrusions à la surface du pilus. La barre blanche représente une échelle de 25 Å. B) et C) Reconstruction du fimbriae, vue latéralement et vue en coupe, respectivement. La densité électronique est représentée en blanc dans laquelle est représenté l'assemblage en hélice de plusieurs pilines majeures PapA. D) Modèle de repliement hélicoïdal de la tige. E) Polymérisation de deux sous-unités PapA. L'extension N-terminale de la sous-unité PapA (n-1) s'insère dans le sillon hydrophobe de la

sous-unité PapA (n). La boucle qui permet cette structure en hélice est montrée par une flèche blanche (Mu and Bullitt, 2006).

Ces protrusions qui sont visibles à la surface de la tige correspondent à une boucle formée par l'extension N-terminal de la piline majeure (n-1) insérée dans le sillon hydrophobe de la piline majeure (n) (Figure 24) (Mu and Bullitt, 2006). Cette boucle permet la rotation des pilines majeures de la position verticale, où elles traversent la protéine « usher », à la position horizontale, où elles forment l'hélice. PapA forme un angle de 13° avec la perpendiculaire de l'axe de l'hélice (Figure 24). Des mutations dans la région de cette boucle empêchent cette rotation et la formation de l'hélice, générant de fins filaments fimbrillaires.

La tige du pili type I correspond à l'assemblage de 500 à 3000 pilines majeures FimA en hélice droite. Un tour d'hélice contient 3,4 sous-unités FimC et a pour diamètre externe 6,9 nm (Hahn et al., 2002) (Figure 25).

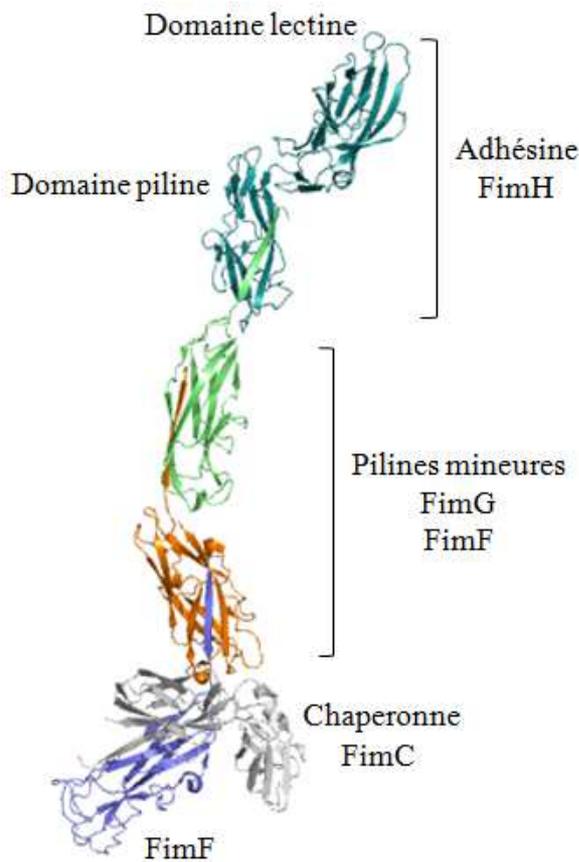


**Figure 25: La tige du pilus de type I d'*Escherichia coli*.** a) Reconstruction 3D de la tige du pilus type I contenant 40 pilines majeures FimA. b) Carte de densité électronique de 5 sous-unités FimA calculée à partir de la reconstruction 3D en a) (D'après (Hahn et al., 2002)). c) Représentation de l'interaction entre deux sous-unités FimA, l'une en jaune et l'autre en bleu, dans la tige. Les extensions N-terminales sont représentées en rouge. (D'après (Choudhury et al., 1999)).

#### 3.2.2.1.2. le fimbrillum du pilus: les pilines mineures

Le fimbrillum qui résulte d'un assemblage linéaire des pilines mineures est flexible (Hahn et al., 2002). Il mesure environ 42 nm et est composé de 5 à 10 copies de la piline mineure PapE, de deux pilines mineures adaptatrices PapF et PapK et de l'adhésine PapG pour les pili de type P. La protéine adaptatrice PapK relie le fimbrillum à la tige et est impliquée dans le contrôle de la longueur du fimbrillum (Jacob-Dubuisson et al., 1993). La surproduction de PapK conduit à un fimbrillum plus court en stoppant l'assemblage des pilines mineures PapE et en déclenchant la polymérisation de la piline majeure PapA (Hahn et al., 2002).

Le fimbrillum du pilus de type I mesure 10 à 19 nm de longueur. Il est composé d'une seule protéine adaptatrice FimF, de la piline mineure FimG et de l'adhésine FimH (Figure 26). Le fimbrillum du pilus de type I est plus flexible que le fimbrillum du P pilus grâce à une plus faible contrainte d'association entre FimF et ses partenaires (Gossert et al., 2008). Plus le fimbrillum est flexible et plus la reconnaissance impliquant l'adhésine est facilitée ce qui se traduit par une meilleure fixation de la bactérie aux cellules hôtes.

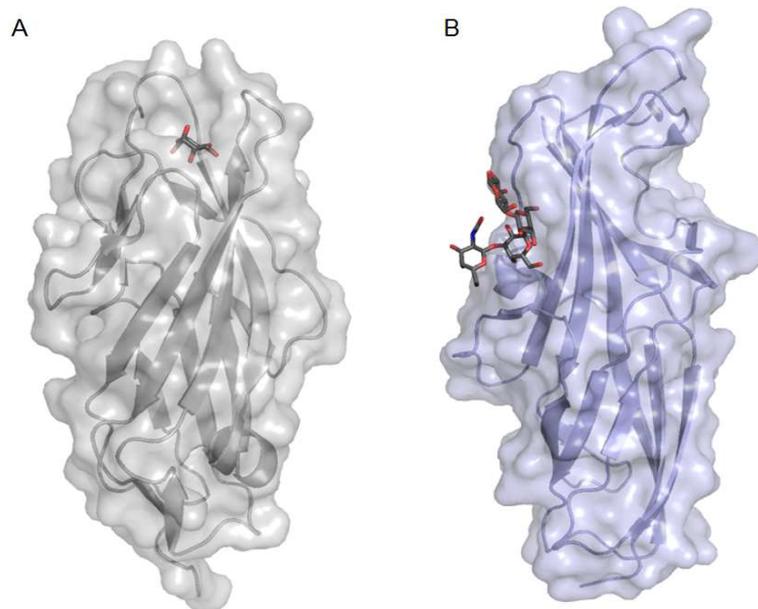


**Figure 26: Structure du fimbrillum du pili type I.** A l'extrémité du fimbrillum est représenté en bleu cyan l'adhésine FimH. Les deux pilines mineures du fimbrillum FimG et FimF sont représentées en vert et orange, respectivement. La chaperonne FimC, en gris complexée à la sous-unité FimF, en bleu, permettent de stabiliser la production du fimbrillum. Toutes les extensions N-terminale complètent le sillon hydrophobe de la sous-unité en aval permettant alors la polymérisation du fimbrillum (D'après (Le Trong et al., 2010)).

### 3.2.2.1.3. le fimbrillum du pilus: l'adhésine

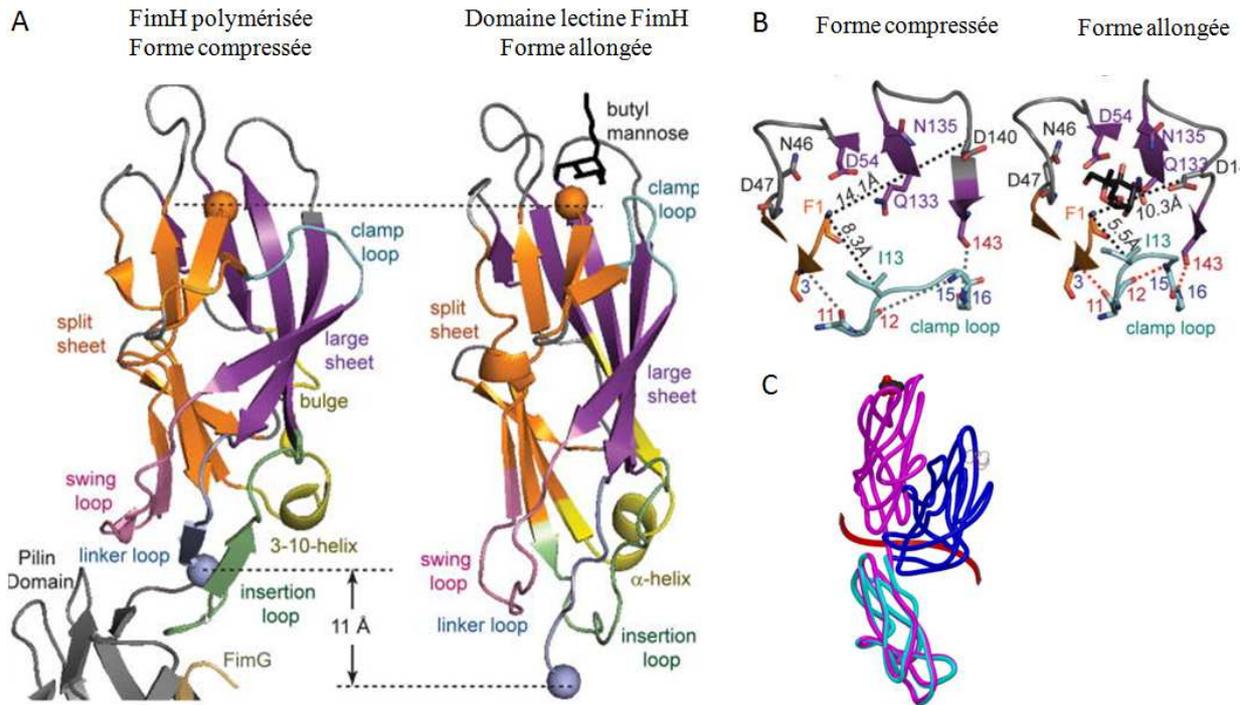
Les adhésines de la voie CU situées au sommet du pilus sont composées de deux domaines séparés par un linker, le domaine piline et le domaine lectine. Le domaine piline, à l'extrémité C-terminale, possède un repliement de type immunoglobuline incomplet comme toutes les sous-unités pilines décrites précédemment qui lui permet d'être associé au pilus (Figure 26). Le domaine lectine forme un tonneau bêta allongé avec un repliement de type jelly roll. Il présente un site de reconnaissance au sucre spécifique à chaque adhésine (Figure 27). La lectine FimH possède une poche à mannose située au sommet du domaine adhésine ce qui lui permet de reconnaître le mannose présent sur les surfaces des uroépithélia (Hung et al., 2002) (Figure 27). Les lectines PapG qui forment trois classes: PapGI, PapGII et PapGIII, possèdent

une affinité pour les globosides GbO3, GbO4 et GbO5, respectivement. La structure cristallographique de PapGII complexé avec GbO4, glycolipide qui domine à la surface des cellules du rein humain, montre que le site de fixation au sucre est positionné latéralement (Figure 27). La fixation du récepteur GbO4 ne se réalise que lorsque PapGII est parallèle à la surface, ce qui est rendu possible grâce à la grande flexibilité du fimbrillum. La fixation de GbO4 n'induit pas de changement de conformation de PapGII (Dodson et al., 2001).



**Figure 27: Comparaison des domaines de reconnaissances des sucres entre les deux adhésines FimH et de PapGII.** Structures atomiques A) du complexe FimH/ mannose et B) du complexe PapGII/ GbO4. Les sucres sont représentés en bâtons noirs.

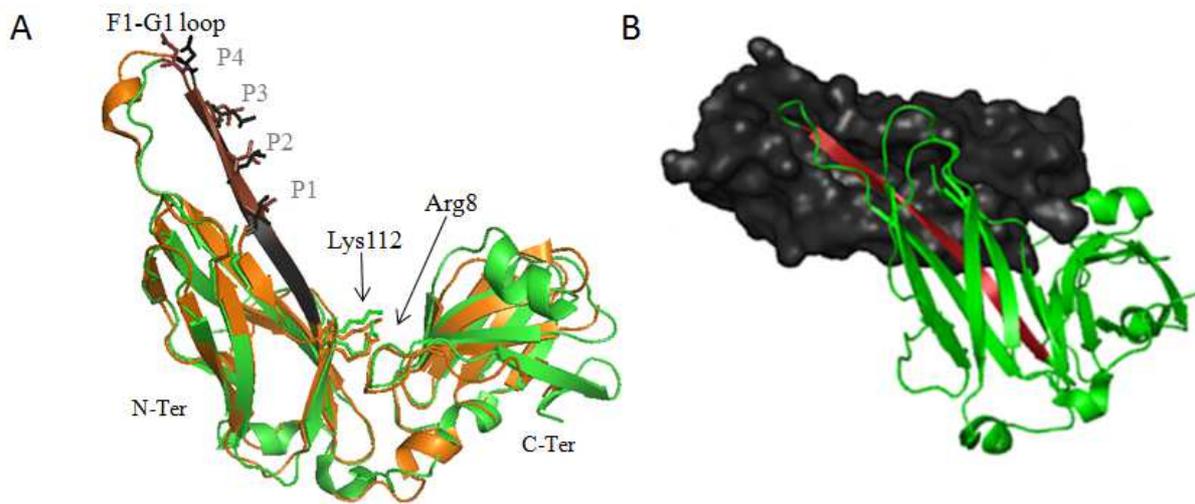
Une étude récente a mis en évidence une régulation allostérique de l'adhésine FimH. Assemblé au pilus, le domaine lectine de FimH est maintenu, par son domaine piline, dans une conformation compressée (Figure 28). La conformation compressée ne permet pas le maintien de la poche à mannose. La présence de mannose peut induire un changement de conformation de l'état compressé à l'état allongé. Cependant le changement reste transitoire car le domaine piline peut induire la conformation compressée à tout moment. Cet effet d'auto-inhibition causé par le domaine piline doit certainement être régulé *in vivo* par un mécanisme encore inconnu qui par l'induction d'une force de torsion permettrait une conformation allongée du domaine lectine à haute affinité avec le mannose (Le Trong et al., 2010)



**Figure 28: Mise en évidence d'une auto-inhibition induite par le domaine piline de l'adhésine FimH.** A) La forme compressée du domaine lectine de FimH lorsqu'elle est assemblée au pilus, et la forme allongée du domaine lectine FimH sans la présence de son domaine piline. B) Vue de dessus de la poche à mannose dans les deux types de conformations. C) Superposition de FimH associée au pilus (domaine lectine en bleu foncé et domaine piline en cyan) avec FimH associée à la chaperonne (en magenta). Le mannose est représenté en mode cpk en noir. (D'après (Le Trong et al., 2010)).

### 3.2.2.2. La chaperonne périplasmique

Les chaperonnes périplasmiques de la voie d'assemblage CU sont composées de deux domaines orientés en L (Figure 29). Chaque domaine est composé de 7 brins  $\beta$  antiparallèles (A1-G1 et A2-G2) dont le repliement est de type immunoglobuline en  $\beta$ -barrel (feuillet  $\beta$  qui forment un tonneau) (Holmgren and Brändén, 1989; Eidam et al., 2008). La boucle qui relie le brin F1 au brin G1 de la chaperonne diffère en longueur et définit la classe de la chaperonne (FGS pour F1-G1 short et FGL pour F1-G1 long). 10 à 20 acides aminés constituent la loop F1-G1 pour les chaperonnes FGS contre 21 à 29 pour les chaperonnes FGL. Les chaperonnes PapD et FimC appartiennent aux chaperonnes de type FGS pour lesquelles une grande similarité de structure est observée avec la conservation d'un motif de résidus hydrophobes sur le brin G1 exposé au solvant ainsi que deux acides aminés basiques, une lysine sur le brin G1 et une arginine sur le brin A1, (Figure 29) (Sauer et al., 2004). Les chaperonnes de la voie CU ont pour fonction de stabiliser les sous-unités structurales afin d'éviter leur polymérisation et leur dégradation dans le périplasma (Figure 29).

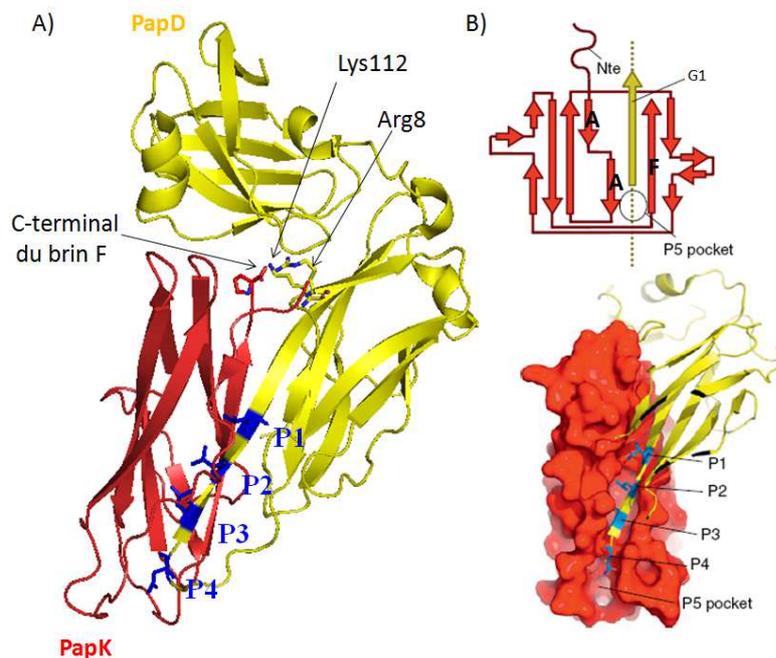


**Figure 29: Structures des chaperonnes PapD et FimC d'*Escherichia coli*.** A) Superposition des chaperonnes PapD et FimC. Les résidus hydrophobes (P1 à P4) conservés sur le brin G1 de PapD et FimC sont représentés en brun et noir, respectivement. Les résidus lysine et arginine de PapD et FimC sont représentés en orange et vert, respectivement. B) Structure du complexe chaperonne PapD/ piline PapK. La surface de la piline PapK est en gris foncé et PapD est représenté en vert. Le brin G de PapD qui stabilise la piline en complétant sa structure immunoglobuline est en rouge (D'après (Sauer et al., 1999)). (Figures réalisées avec Pymol).

### 3.2.2.3. La complémentation par brin donneur avec les chaperonnes FGS

Dans le périplasma, la formation du complexe chaperonne/piline implique deux acides aminés basiques de la chaperonne, l'arginine 8 et la lysine112, qui vont interagir avec le groupement carboxyle C-terminal du brin  $\beta$  F de la piline (Figure 30). Ces deux acides aminés sont nécessaires à la formation du complexe chaperonne/piline. En effet, les interactions qu'elles réalisent favorisent le bon positionnement du brin  $\beta$  F de la piline, ce qui facilite l'insertion d'une partie du brin  $\beta$  G1 de la chaperonne dans le sillon hydrophobe de la piline. Sur le brin  $\beta$  G1 de la chaperonne est présent un motif constitué de quatre résidus hydrophobes (P1, P2, P3 et P4) qui vont se loger dans des poches hydrophobes (P1, P2, P3 et P4) du sillon hydrophobe de la piline (Figure 30) (Sauer et al., 2002a). Contrairement aux poches P1, P2, P3 qui sont profondes et bien définies, la poche P4 est superficielle obligeant le résidu P4 à rester parallèle au sillon hydrophobe de la piline. Cette interaction appelé « bêta zipper » forme de nombreuses liaisons hydrogènes entre les atomes du brin  $\beta$  G1 de la chaperonne et ceux des brins  $\beta$  A et F qui encadrent le sillon hydrophobe de la piline. Le mécanisme durant lequel la chaperonne complète le repliement immunoglobuline de la piline par l'apport d'une partie de son brin  $\beta$  G1 dans le sillon hydrophobe de la piline est appelé « complémentation par brin donneur ». La formation du complexe chaperonne/piline, donne à la sous-unité piline

un repliement immunoglobuline atypique où le brin  $\beta$  G1 de la chaperonne est parallèle au brin  $\beta$  F de la piline (contrairement au repliement immunoglobuline classique où le brin  $\beta$  G est antiparallèle au brin  $\beta$  F) (Figure 30). Ce repliement atypique entraîne une conformation ouverte de la piline, et permet également à la poche hydrophobe P5 d'être accessible pour le remplacement du brin  $\beta$  G1 de la chaperonne par l'extrémité N-terminale des pilines durant la polymérisation du pilus (Figure 30). Le mécanisme de « complémentation par brin donneur » offre aux sous-unités pilines une stabilité provisoire avant leur polymérisation impliquant le mécanisme « d'échange de brin donneur ».

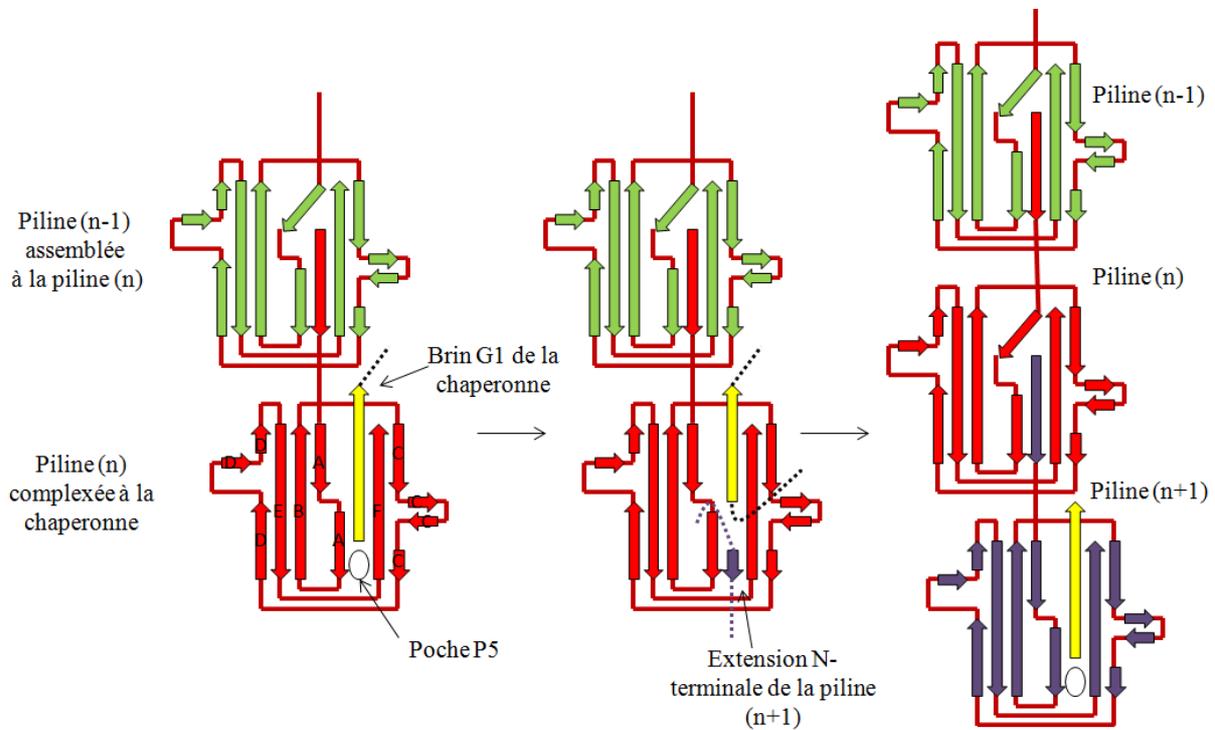


**Figure 30: Complémentation par brin donneur.** A) Structure du complexe chaperonne PapG/ sous-unité PapK. La chaperonne est en jaune. Les résidus hydrophobes, P1, P2, P3 et P4 sur le brin G sont représentés en bâtons bleus. L'arginine 8 et la lysine 112 sont représentés en bâtons jaunes. La piline est en rouge et le groupement carboxyle (Coo-) C-terminal du brin F est représenté en bâton. (D'après (Sauer et al., 2002)). B) Représentation topographique et structurale de l'insertion du brin G de la chaperonne dans le sillon hydrophobe de la piline PapK (Allen et al., 2012).

#### 3.2.2.4. L'échange de brin donneur

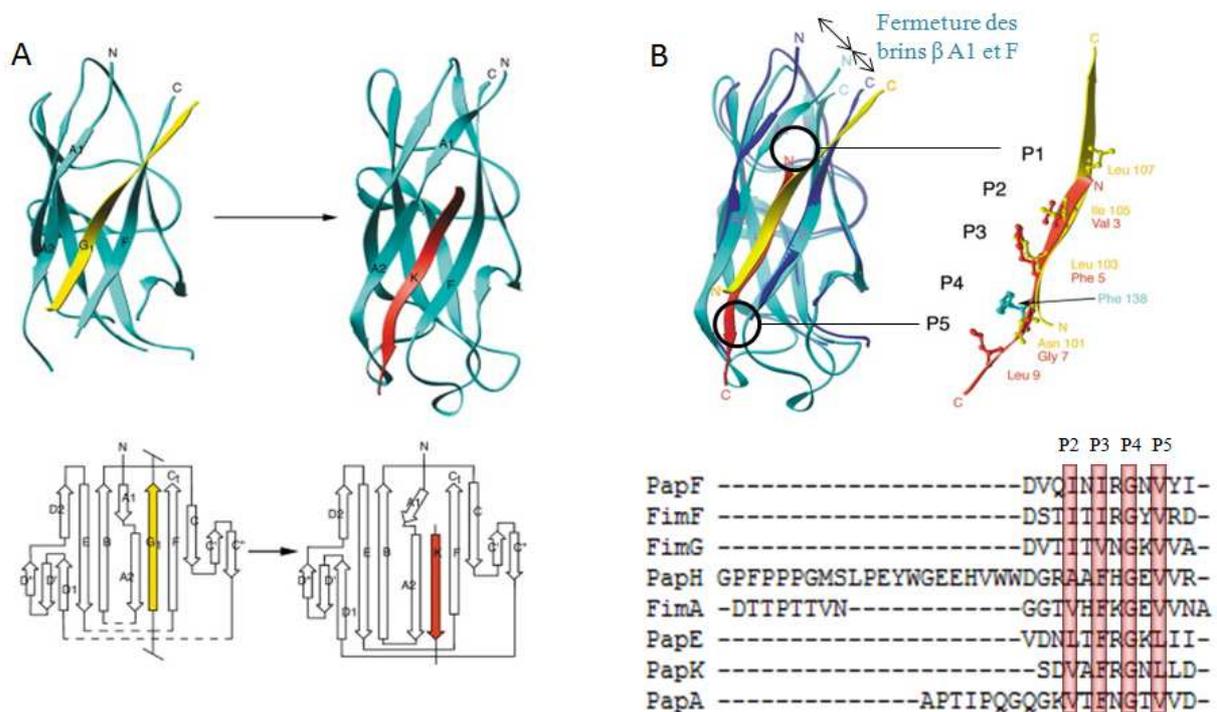
Toutes les sous-unités pilines, à l'exception de l'adhésine, possèdent une extension N-terminale, qui n'intervient pas dans leur repliement. Cette extension N-terminale contient un motif contenant 4 résidus hydrophobes appelés P2, P3, P4, et P5. Lors de la polymérisation des sous-unités pilines, les résidus P2 à P5 de l'extension N-terminale de la piline à assembler ( $n+1$ ) vont venir se loger dans les poches P2 à P5 du sillon hydrophobe de la sous-unité déjà assemblée au fimbriae ( $n$ ). Le brin  $\beta$  G1 de la chaperonne, qui est complexé avec la piline ( $n$ ), est alors délogé du sillon hydrophobe et remplacé par l'extension N-terminale de la piline entrante ( $n+1$ ). Ce mécanisme est appelé « échange de brin donneur ». Il est initié par l'insertion du résidu P5 de l'extension N-terminale de la piline entrante ( $n+1$ ) dans la poche hydrophobe P5 laissée accessible par le complexe chaperonne/ piline ( $n$ ) (Figure 31). L'interaction de l'extrémité N-terminale avec le sillon hydrophobe de la piline permet un

repliement immunoglobuline classique. Le brin qui correspond à l'extension N-terminale de la piline entrante (n+1) se positionne de manière antiparallèle au brin  $\beta$  F de la piline déjà assemblée (n). La poche hydrophobe P1 n'étant plus occupée, les brins  $\beta$  F1 et A1 vont se rapprocher et conduire à la fermeture du sillon hydrophobe (Figure 32).



**Figure 31: Représentation topologique illustrant l'assemblage des sous-unités par le mécanisme « d'échange par brin donneur ».** ( Reproduit d'après (Sauer et al., 2002a; Allen et al., 2012)).

Le mécanisme « d'échange de brin donneur » permet le passage d'un complexe chaperonne/sous-unité stable mais dans un état semi-replié de haute énergie, à un complexe piline/piline correctement replié de faible énergie (Figure 32 A) (Sauer et al., 2002a). La polymérisation des sous-unités par le mécanisme d'échange de brin donneur se fait au niveau de la protéine «usher ».

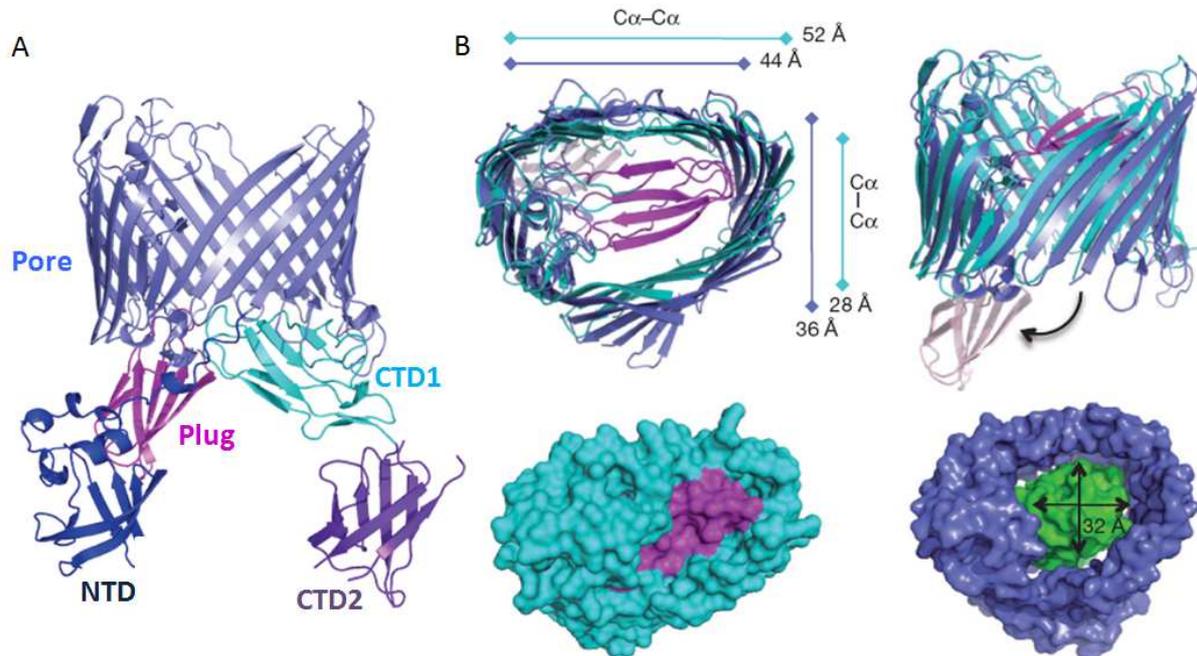


**Figure 32: De la « complémentation par brin donneur » à « l'échange de brin donneur ».** A) Représentation structurale et topologique de la « complémentation par brin donneur » à « l'échange de brin donneur ». A gauche est représentée la structure de PapE en bleu complexée avec la chaperonne dont seul le brin G est représenté (en jaune). La piline PapE a une conformation ouverte. A droite PapE est assemblée avec l'extrémité N-terminale de PapK (représentée en rouge). La piline PapE a une conformation fermée. Les représentations topologiques des structures mettent en évidence le brin G de la chaperonne qui s'insère parallèlement au brin F de la piline, alors que l'extension N-terminale de PapK s'insère de manière antiparallèle au brin F de la piline. L'insertion de l'extension N-terminale induit un changement de conformation du brin A1 de la piline PapE (Sauer et al., 2002a). B) Superposition de PapE/ brin G1 chaperonne avec PapE/ Extension N-terminale de PapK. La piline PapE complexée au brin G1 chaperonne (en jaune) est représentée en bleu foncé et la piline PapE assemblée à l'extension N-terminale de PapK (en rouge) est représentée en bleu clair. La superposition du brin G de la chaperonne et de l'extension N-terminale met en évidence les résidus qui interagissent avec les poches hydrophobes P1 à P5 (d'après (Sauer et al., 2002a)). Un alignement de séquences des extensions N-terminales des sous-unités du pili de type P et de type I, montre la conservation de la glycine qui occupe la poche P4 aux dimensions restreintes.

### 3.2.2.5. La protéine de membrane externe « Usher »

La protéine de membrane externe appelée « usher » ou « trieuse », est composée de cinq domaines: (1) Le domaine N-terminal (NTD) qui constitue le premier site de recrutement des complexes chaperonne/sous-unités. (2) Un domaine qui constitue un pore de translocation membranaire où les sous-unités récemment assemblées vont être transportées à la surface bactérienne. (3) Un domaine bouchon (domaine plug) qui bloque le pore de translocation

quand il n'est pas utilisé. (4 et 5) Les deux domaines C-terminaux (CTD1 et CTD2) qui constituent le second site de fixation des complexes chaperonnes/sous-unités (Figure 33).

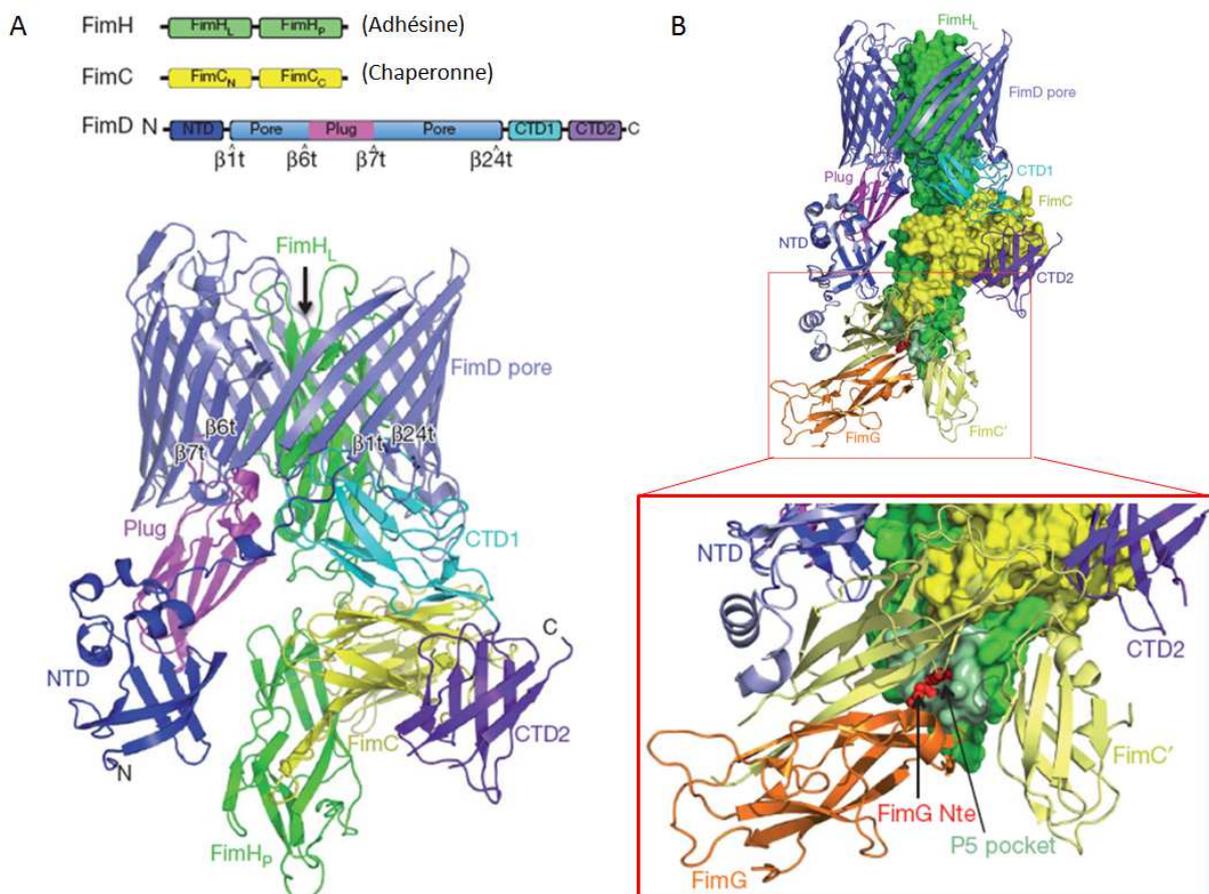


**Figure 33: Structure de la protéine «usher».** A) Structure de la protéine «usher» entière constituée d'un pore de translocation, d'un domaine N-terminal (NTD), d'un domaine Plug et d'un domaine C-terminal composé de deux domaines (CTD1 et CTD2). B) Comparaison de la conformation du pore de translocation inactif (en cyan) et actif (bleu). Le domaine Plug est représenté en violet. L'activation de «usher» se fait par le déplacement du domaine Plug dans le périplasma. Le pore de translocation adopte une forme ovale avec un conduit de 32Å de diamètre, permettant le passage des sous-unités (en vert) (D'après (Phan et al., 2011)).

Le pore transmembranaire (d'environ 800 acides aminés) se compose de 24 brins  $\beta$  antiparallèles qui forment une structure en tonneau bêta. Entre les brins  $\beta 6$  et  $\beta 7$  est inséré un domaine bouchon sur une boucle périplasmique. Ce dernier est constitué de 6 brins  $\beta$  repliés en sandwich bêta. Dans la forme inactive du pore de translocation, le domaine plug est maintenu latéralement et obstrue le pore. L'interaction du complexe chaperonne/adhésine avec la protéine «usher» déclenche son activation. Le mécanisme impliqué à l'activation de «usher» reste encore inconnu mais il induit un changement de conformation par le déplacement du domaine plug dans le périplasma, ouvrant un pore de 32 Å. Les dimensions du pore vont alors passer de 52 à 44 Å pour le diamètre et de 28 à 36 Å pour la hauteur lors de l'activation (Figure 33) (Phan et al., 2011).

Le domaine NTD de la protéine «Usher» est le premier site de fixation des complexes chaperonne/sous unités. Le NTD de PapC a une affinité de l'ordre du nanomolaire et micromolaire avec les complexes chaperonne/sous-unités PapD/PapG et PapD/PapE,

respectivement. Aucune interaction n'a été observée avec les autres complexes PapD/PapK, PapD/PapA et PapD/PapH, alors que le domaine Plug se fixe à tous les complexes.



**Figure 34: Structure des différentes interactions mises en place durant l'assemblage du pilus au niveau de la protéine « usher ».** A) Structure du complexe FimD/FimC/FimH. B) Superposition au niveau du domaine NTD de la structure du complexe FimD/FimC/FimH et de la structure du complexe FimD-NTD/ FimC/FimF. La sous-unité FimF est appelée FimG sur le modèle et est représentée en orange. L'extension N-terminale qui s'insère dans la poche P5 (vert clair) figure en rouge (D'après (Phan et al., 2011).

Le domaine NTD présente cependant une plus grande affinité pour le complexe PapD/PapG que le domaine Plug. Le complexe PapD/PapG est donc recruté par le domaine NTD, avant l'ouverture du pore et la formation du complexe NTD/Plug. Par la suite le recrutement des différents complexes chaperonne/sous-unités va se faire par le complexe NTD/Plug. Le domaine CTD2, qui se fixe avec une affinité de l'ordre du micromolaire avec tous les complexes excepté PapD/PapH, sera le deuxième site de fixation (Volkan et al., 2012). La résolution de la protéine « usher » entière FimD avec le complexe FimC/FimH, prouve que le domaine CTD1 se fixe à la lectine FimH et que le domaine CTD2 se fixe au domaine piline de FimH et à la chaperonne FimC. Par ces différentes interactions, les domaines CTD1 et

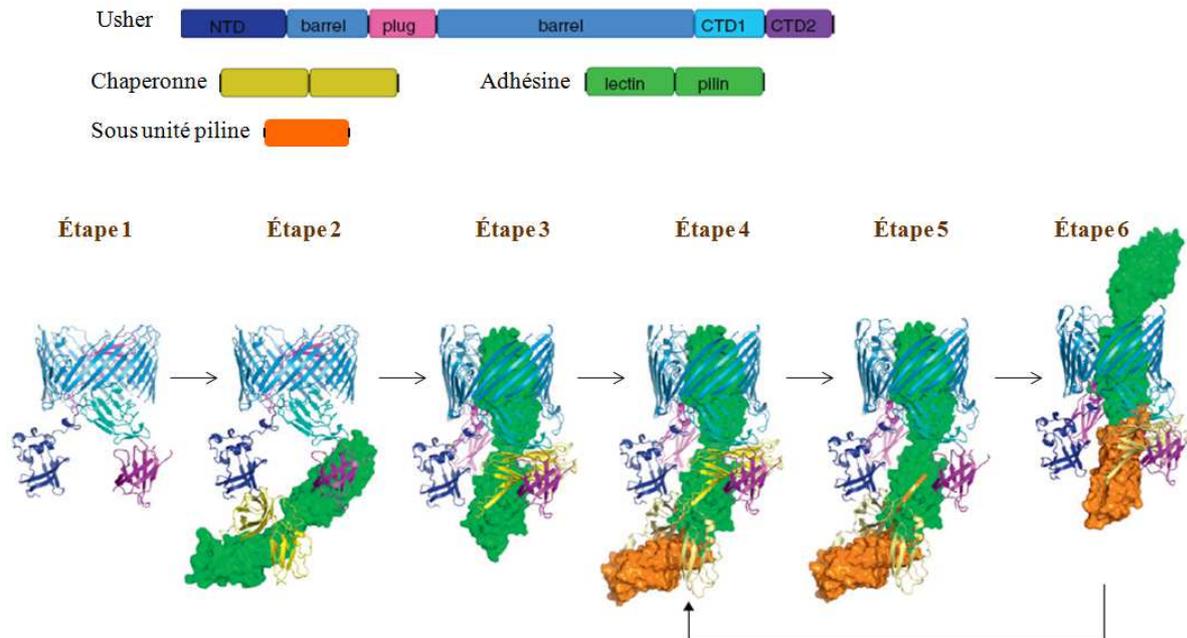
CTD2, vont assurer le bon positionnement de la sous-unité afin de permettre le mécanisme « d'échange de brin donneur ». En effet, par la superposition de la structure FimD entière/FimC/FimH avec FimD-NTD/FimC/FimF, il a été mis en évidence que l'extension N-terminale de la piline FimF tenue par NTD, vient se lier à la poche P5 de la piline FimH, constituant la première étape d'échange par brin donneur (Figure 34) (Phan et al., 2011). Par les domaines NTD/Plug et CTD1/CTD2, la protéine « usher » sous forme monomérique catalyse l'assemblage des sous-unités pilines.

#### 3.2.2.6. Modèle d'assemblage

La structure cristallographique de FimD entière/ FimC/ FimH remet en cause le modèle établi par Remaut (Remaut et al., 2008), où la dimérisation de la protéine « usher » est nécessaire à l'assemblage du pilus. Un nouveau modèle a été établi mais de nombreux mécanismes restent encore inconnus (Figure 35) (Phan et al., 2011).

Durant la formation des complexes chaperonne/sous-unité dans le périplasma, la protéine « usher » est inactivée par le domaine Plug qui obstrue le pore (étape 1). La première étape de la biogenèse du pilus, basée sur la forte affinité de NTD pour le complexe chaperonne/adhésine, se fait par le recrutement du complexe chaperonne/adhésine par NTD (Volkan et al., 2012). Ensuite, par un mécanisme encore inconnu qui implique une étape de dissociation, le complexe chaperonne/adhésine va passer de son premier site de fixation NTD à son second site de fixation qui est le domaine CTD2 (étape 2). Durant ce mécanisme, le complexe chaperonne/adhésine va permettre l'activation de la protéine « usher ». Le domaine plug qui obstruait le pore dans la forme inactive du pore vient interagir avec le domaine NTD dans le périplasma. Cette activation du pore de translocation induit des changements de conformation qui permettent l'entrée du domaine lectine au sein du pore. Le domaine CTD1 se fixe à la lectine de l'adhésine et le domaine CTD2 se fixe au domaine piline de l'adhésine et à la chaperonne FimC. Les domaines CTD permettent le bon positionnement du complexe chaperonne/sous-unité (étape 3). Le complexe NDT/Plug va alors recruter le complexe chaperonne/sous-unité suivant qui va être orienté par le domaine NTD permettant l'échange de brin donneur (étape 4). Lorsque l'extension N-terminale de la piline (n+1) est insérée dans le sillon de la piline (n), la chaperonne de la piline (n) est relarguée dans le périplasma (étape 5). Le complexe chaperonne/piline (n+1) va interagir avec les domaines CTD1 et CTD2 poussant alors la sous-unité assemblée à travers le pore jusqu'à la surface de la bactérie (étape 6). Ces étapes vont se répéter jusqu'à la formation du pilus. Dans le système Pap, le complexe PapD/piline PapH qui ne présente une affinité que pour le domaine Plug va terminer la

formation du pilus. Dans le système Fim aucune sous-unité piline ayant pour rôle de terminer le pilus n'a été identifiée. Le mécanisme qui permet de réguler la longueur de pilus reste à être découvert.



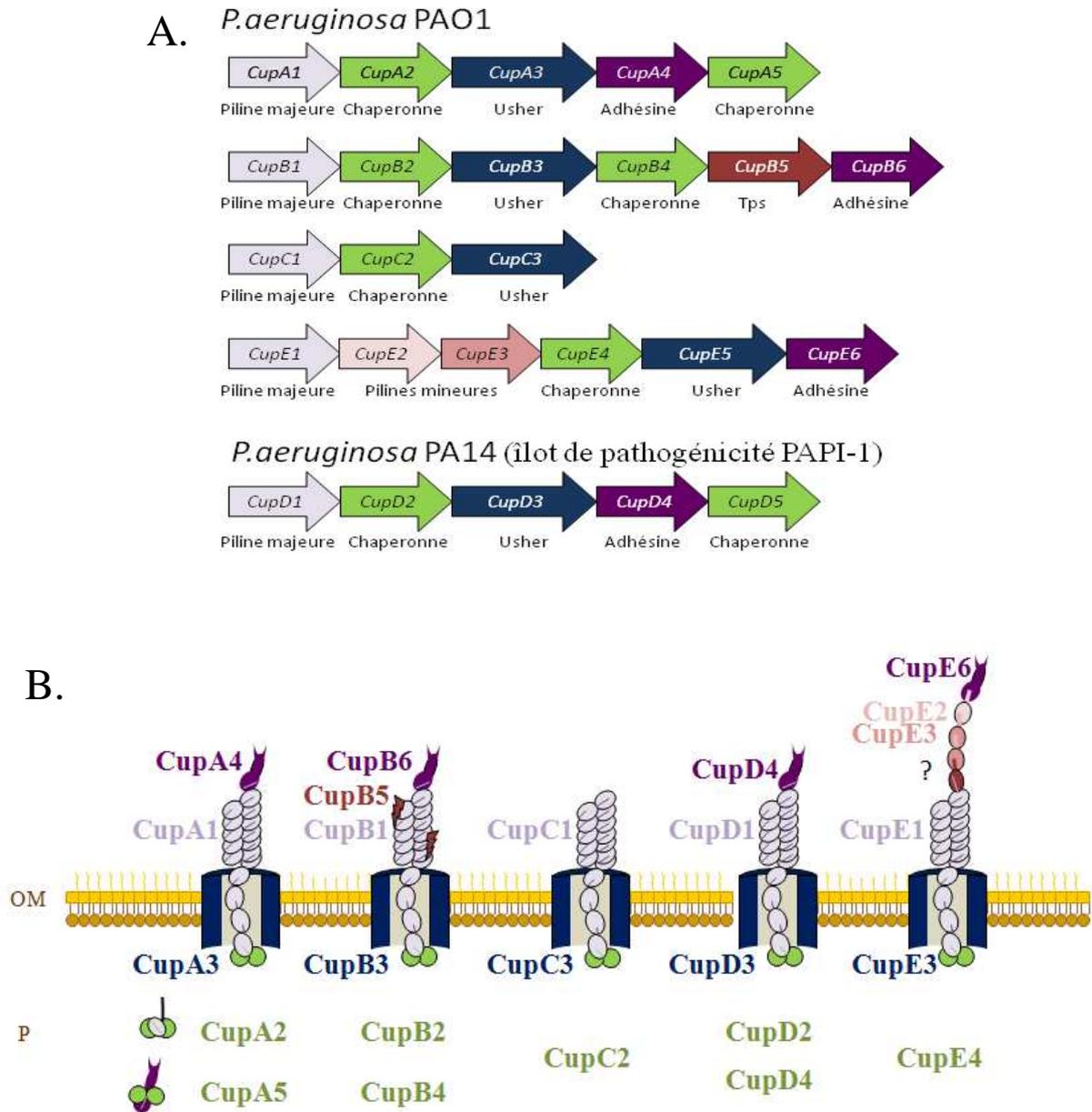
**Figure 35: Modèle de la biogénèse des pili CU au niveau de la protéine « usher ».** A l'étape 1, la protéine « usher » est inactivée par le domaine Plug qui obstrue le pore. A l'étape 2, le complexe chaperonne /adhésine est recruté par le domaine NTD. Par une étape de dissociation le complexe chaperonne/adhésine va interagir avec son deuxième site de fixation : les domaines CTDs. A l'étape 3, la protéine « usher » est activée par le déplacement du domaine Plug dans le périplasma. Le domaine Plug va former un complexe avec le domaine NTD. A l'étape 4, le domaine NTD oriente le complexe chaperonne/sous-unité piline recruté. A l'étape 5, l'échange de brin donneur a lieu et la chaperonne de l'adhésine est relarguée dans le périplasma. A l'étape 6, une interaction entre les domaines CTDs et le complexe chaperonne/sous-unité piline permet le bon positionnement du complexe pour l'initiation à l'échange de brin donneur. L'adhésine récemment assemblée est poussée à la surface de la bactérie. L'étape 4 à 6 se répète jusqu'à l'assemblage complet du pilus. (D'après (Allen et al., 2012)).

### 3.3. Les voies CUP chez *Pseudomonas aeruginosa*

Chez *P. aeruginosa*, plusieurs voies d'assemblage de type CU ont été identifiées et appelées Cup (Chaperone-Usher pathway). Dans le génome de PAO1 sont présents les clusters *cupA* (Vallet et al., 2004), *cupB*, *cupC* (Vallet et al., 2004; Ruer et al., 2008), *cupE*, et *cupF* (He et al., 2004; Giraud et al., 2011). Le cluster *cupD* (He et al., 2004) est présent dans l'îlot de pathogénicité PAPI-I du génome de la souche PA14 (Figure 36).

La particularité des systèmes Cup, à l'exception de *cupE*, est que les clusters *cupA-D* ne possèdent pas de gène qui code pour des pilines mineures et chaque sous-unité possède sa propre chaperonne (Figure 36). Le cluster *cupE* ressemble plus aux systèmes CU d'*E. coli*,

par son organisation et sa composition. Les clusters *cupA-D* appartiennent au clade  $\gamma_4$ -fimbriae et le cluster *cupE* au clade  $\sigma$ -fimbriae. Tous soumis à un système de régulation, les fimbriae CupA-E jouent un rôle bien défini dans la formation du biofilm.



**Figure 36: Les systèmes Cup chez *Pseudomonas aeruginosa*.** A) Représentation de l'organisation des gènes au sein des clusters codant pour les fimbriae de type CU chez *Pseudomonas aeruginosa*. B) Représentation schématique des fimbriae Cup (OM = membrane externe ; P = périplasme).

### 3.3.1. CupA

#### 3.3.1.1. Le système CupA

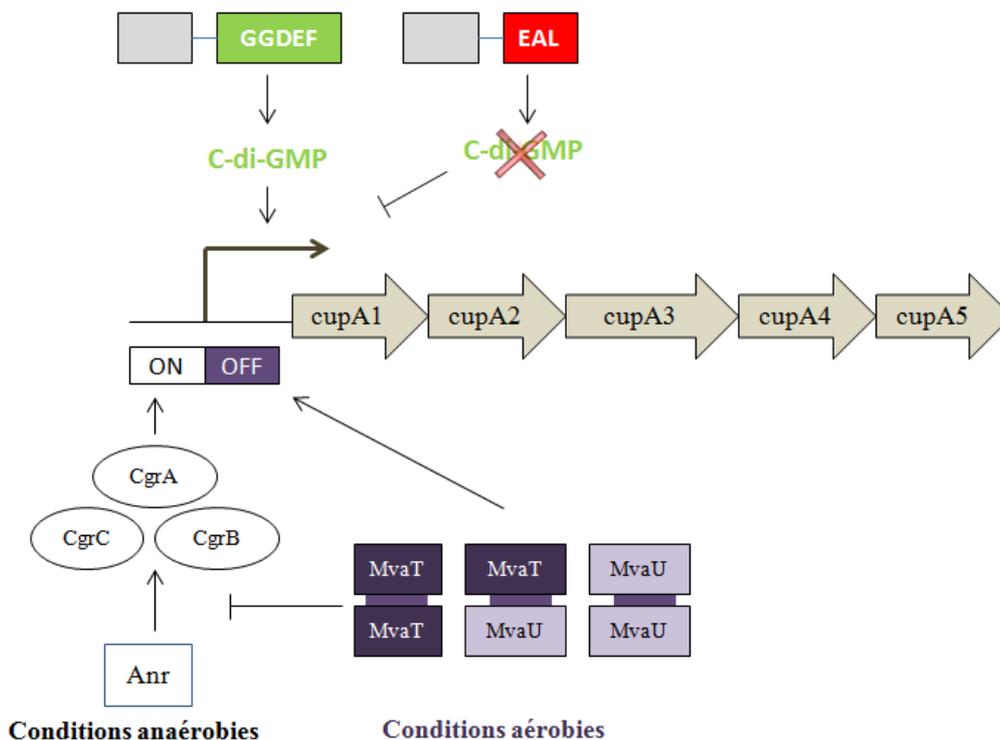
Le cluster *CupA* a été identifié lors d'un criblage d'une banque de transposon Tn construite dans la souche PAK $\Delta$ *pilA* (dépourvue du pili type IV) afin d'identifier de nouveaux acteurs moléculaires impliqués dans la formation du biofilm. Deux mutants ayant perdu la capacité de former un biofilm, présentaient une insertion dans le deuxième gène d'un cluster de cinq gènes. Ce gène appelé *cupA2* codait pour une chaperonne périplasmique de type FGS (Vallet et al., 2001).

Le cluster *cupA* est composé de cinq gènes *cupA1* à *cupA5* (Figure 36) qui code pour les protéines suivantes : la piline majeure CupA1, l'adhésine CupA4, la protéine « usher » CupA3 et deux chaperonnes CupA2 et CupA5. La chaperonne CupA2 se complexe avec la sous-unité CupA1, et la chaperonne CupA5 avec l'adhésine CupA4. L'adhésine CupA4 ne présente aucun domaine classique retrouvé chez les adhésines. Mais sa grande taille et des similarités à d'autres sous-unités pilines retrouvées en C-terminal laissent penser qu'il s'agit d'une adhésine atypique.

Les fimbriae CupA sont impliqués dans l'étape précoce de la formation du biofilm, c'est-à-dire l'étape de l'attachement initial (Vallet-Gely et al., 2007). Ils ont été retrouvés dans les souches SCV (Small Colony Variant) de *P. aeruginosa* cultivés à 28°C, qui sont des petites colonies aux propriétés adhérentes et agrégatives (Meissner et al., 2007). Ils jouent également un rôle dans l'établissement et le maintien de l'intégrité structurale de la pellicule formée grâce à la production de l'exopolysaccharide PEL (Friedman and Kolter, 2004).

#### 3.3.1.2. La régulation du cluster *cupA*

La régulation du cluster *cupA* fait intervenir deux voies principales de régulation (Figure 37). La première voie de régulation permet l'expression du cluster *cupA* selon un mécanisme de variation de phase (Vallet-Gely et al., 2007). Le mécanisme de variation de phase consiste au passage des cellules en phase « on » (qui expriment le cluster *cupA*) à la phase « off » (les cellules n'expriment pas le cluster *cupA*) et inversement. A noter que le passage en phase « on »/ « off » est vingt fois plus fréquent que le passage « off »/ « on ». Ce type de régulation permet à *P. aeruginosa* d'initier la formation du biofilm grâce à la synthèse des fimbriae CupA lorsqu'il est en phase « on ». Alors que le passage en phase « off » permet à *P. aeruginosa* de se détacher du biofilm pour aller coloniser d'autres sites et ce sans être reconnu par le système immunitaire (Vallet-Gely et al., 2007).



**Figure 37: Schéma de la régulation du cluster *cupA*.** (Réalisée d'après (Giraud and De Bentzmann, 2012)).

A une faible concentration d'oxygène, le régulateur transcriptionnel Anr va activer de manière directe la transcription des gènes *cgrA*, *cgrB* et *cgrC* (*cupA* gene regulator) situés en amont du cluster *cupA* (Vallet-Gely et al., 2007). L'expression des gènes *cgrABC* va induire une forte expression du cluster *cupA*. En condition aérobie, les gènes *cgrABC* ainsi que le cluster *cupA* vont être réprimés par la formation d'un complexe hétéromérique ou homomérique des régulateurs « H-NS like », MvaT et MvaU. La propriété des protéines H-NS (heat-stable nucleoid structuring) est de fixer sur l'ADN courbé (comme les régions promotrices des gènes) pour induire une compaction et donc l'inhibition de la transcription des gènes. Il est probable que sous forme de dimère les régulateurs « H-NS like » se fixent directement sur le promoteur des gènes *cgrABC* ainsi que sur le promoteur de *cupA* pour réprimer l'expression de ces gènes (Vallet-Gely et al., 2007). La formation d'homomères MvaT et d'hétéromères MvaT- MvaU domine par rapport à la formation d'homomères MvaU (Vallet-Gely et al., 2005).

La seconde voie de régulation fait intervenir le c-di-GMP. Les protéines qui possèdent un domaine GGDEF vont, en augmentant la concentration intracellulaire du c-di-GMP, induire l'expression du cluster *cupA*. Et à l'inverse les protéines qui possèdent un domaine EAL vont, en diminuant la concentration intracellulaire de ce second messager, diminuer

l'expression de *cupA* (D'Argenio et al., 2002; Meissner et al., 2007; Klebensberger et al., 2009).

Il a également été remarqué que le cluster *cupA* est plus exprimé à 28°C qu'à 37°C (Meissner et al., 2007). La température pourrait donc contrôler l'expression de *cupA* selon un mécanisme qui reste à déterminer.

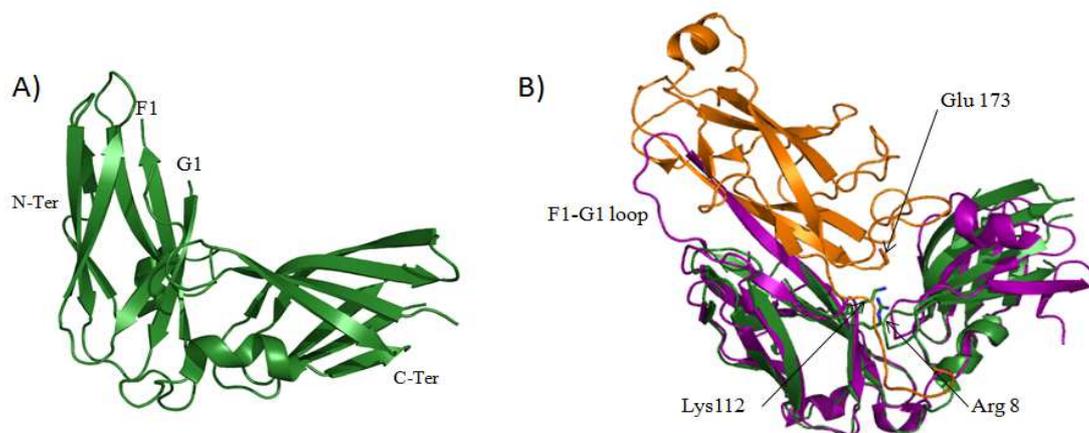
### 3.3.2. CupB et CupC

#### 3.3.2.1. Les systèmes CupB et CupC

Les clusters *cupB* et *cupC* agissent en synergie dans l'agrégation des cellules et la formation de microcolonies, une étape clé à la formation du biofilm.

Le cluster *cupC* est composé de trois gènes : un gène qui code pour une chaperonne périplasmique de la sous-famille FGS appelée CupC2, un gène qui code pour la protéine « Usher », CupC3, et un gène qui code pour la piline majeure CupC1.

Le cluster *cupB* est composé de six gènes. Deux gènes codent pour les chaperonnes CupB2 et CupB6 qui appartiennent également à la sous-famille FGS. La structure de CupB2 a récemment été résolue (Cai et al., 2011).

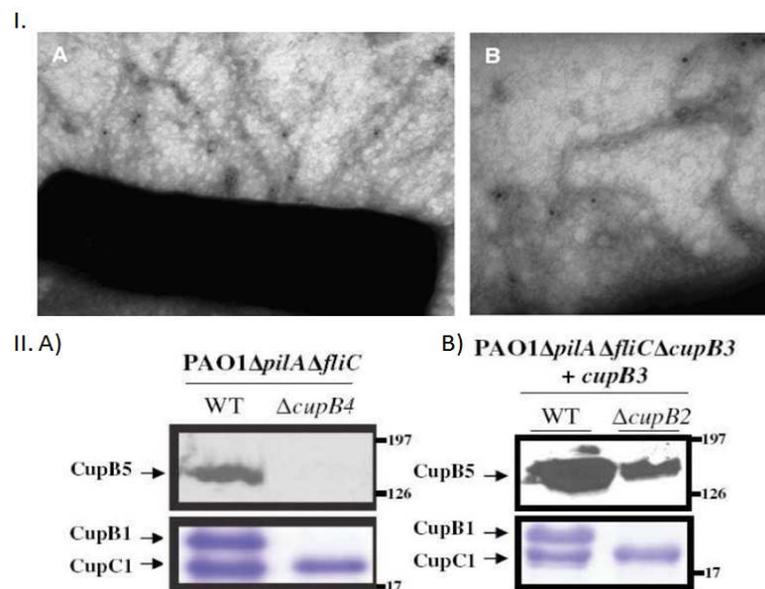


**Figure 38: Structure de la chaperonne CupB2.** A) Structure de CupB2. B) Superposition des structures CupB2 (vert) et PapD (magenta) complexée avec la piline PapH (orange) (Réalisée d'après (Cai et al., 2011)).

La superposition des structures de CupB2 et de PapD complexé avec la piline PapH montre la conservation des deux acides aminés, la lysine K141 et l'arginine R36. Le positionnement de ces deux résidus permettrait une interaction avec le groupement carboxyl C-terminal du brin  $\beta$  F de la piline (Figure 38). Le brin  $\beta$  G de la chaperonne CupB2 est parfaitement superposé au brin  $\beta$  G de PapD. La chaperonne CupB2 qui prend en charge la piline majeure CupB1

présente 32% d'homologie avec la chaperonne CupB4 qui prend en charge l'adhésine CupB6.

La particularité du cluster *cupB* est qu'il possède un gène *cupB5* qui code pour une protéine présentant 44% de similarité à la protéine TpsA4. La protéine TpsA4 appartient au système de sécrétion à deux partenaires TPS où la protéine TpsA va interagir avec la protéine TpsB afin d'être sécrétée. La protéine TpsB forme un tonneau  $\beta$  dans la membrane externe. L'interaction entre les deux protéines se fait par le domaine N-terminal de la protéine TpsA et le domaine POTRA (polypeptide-transport-associated) de la protéine TpsB. Dans l'opéron *cupB*, c'est la protéine « usher » CupB3 qui grâce à l'acquisition du domaine POTRA permet la sécrétion de CupB5. CupB3 appelée « P-usher » (POTRA-like domain-containing Usher) permet l'assemblage du fimbriae mais aussi la sécrétion de CupB5 qui va rester associée à CupB1 (Figure 39). Bien que CupB1 soit prise en charge par la chaperonne CupB2, la présence de la chaperonne CupB4 est nécessaire à la sécrétion de *cupB5* (Ruer et al., 2008) (Figure 39). Le rôle de la protéine CupB5 n'a pas été étudié mais CupB5 comporte, en N-terminal, un domaine à activité hémagglutination (aa 53-167). Ce domaine est trouvé en N-terminal de plusieurs TpsAs caractérisées comme des adhésines, des FHAs (Filamentous hemagglutinin). De plus la localisation de CupB5 près des extrémités des fimbriae laisse croire que la protéine pourrait participer à la formation de microcolonies.



**Figure 39: L'association de CupB5 au fimbriae CupB nécessitant la chaperonne CupB4.** I. Images de microscopie électronique localisant CupB5 sur les fimbriae CupB par un marquage aux billes d'or (Ruer et al., 2008). II. Rôle des chaperonnes CupB2 et CupB4 dans la sécrétion de CupB1 et CupB5. A) Préparation de appendices de la souche PAO1ΔpilAΔfliC (WT) et PAO1ΔpilAΔfliC ΔcupB4 (ΔcupB4). B) Préparation de appendices de la souche PAO1ΔpilAΔfliC (WT) et PAO1ΔpilAΔfliC ΔcupB2

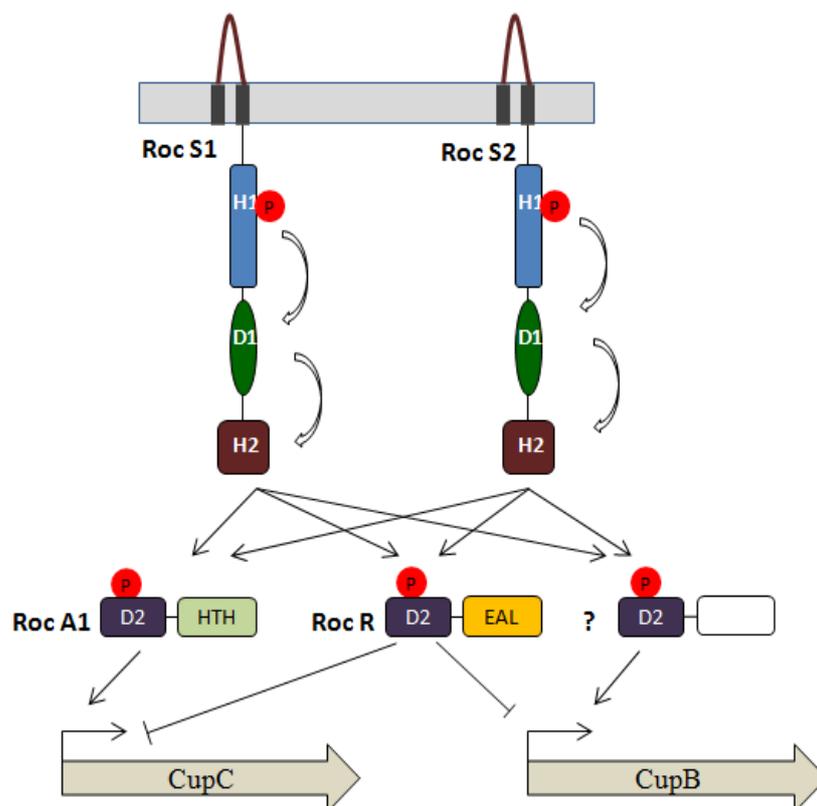
( $\Delta$ cupB2). Le mutant  $\Delta$ cupB2 n'a été obtenu que dans la souche PAO1 $\Delta$ pil $\Delta$ flic $\Delta$ CupB3. Le gène CupB3 est apporté par trans. Des anticorps CupB5 ont été utilisés pour révéler les westerns blot. Les gels SDS page ont été colorés par le bleu de coomassie. (D'après (Ruer et al., 2008), données personnelles prises dans la thèse de Caroline Giraud).

### 3.3.2.2. La régulation des clusters *cupB* et *cupC*

Les clusters *cupB* et *cupC* sont régulés par le système Roc (Regulation of cup) (Figure 40).

Le système Roc1 est un système à trois composants de type non-orthodoxe composé d'une protéine senseur, RocS1, de deux régulateurs de réponse, RocA1 et RocR. La protéine RocR est composée de deux domaines, le domaine receveur D2 et le domaine qui a le motif EAL à activité phosphodiesterase. La protéine RocA1 est composée également d'un domaine receveur D2 mais aussi d'un domaine effecteur de type liaison à ADN appelé HTH (motif hélice-tour-hélice).

Le système RocS2 est un système à deux composants non orthodoxe. Le régulateur de réponse RocA2 est composé d'un domaine receveur D2 mais aussi d'un domaine HTH.



**Figure 40:** Schéma de la régulation des clusters *cupB* et *cupC* par les systèmes Roc. (Réalisée d'après (Giraud and De Bentzmann, 2012)).

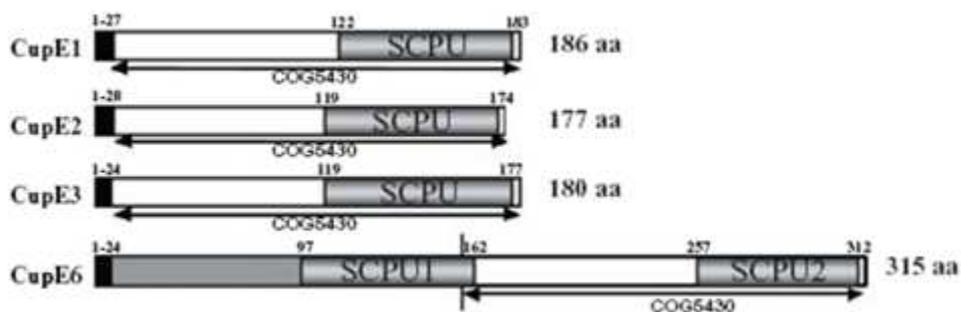
Le senseur RocS1 est capable d'interagir avec les domaines receveurs D2 de RocA1 et RocR (Figure 40) (Kulasekara et al., 2005). RocA1 et RocR sont impliqués dans l'expression de

*cupC* mais RocR agit comme un antagoniste par rapport à l'activité de RocA1. La surproduction de RocS1 et RocA1 active l'expression des gènes *cupC* alors que la surexpression de RocR ne permet pas l'expression des gènes *cupC*. Le senseur RocS2 permet également l'activation des gènes *cupC*. Comme le senseur RocS1, RocS2 interagit avec RocA2 et RocR (Sivaneson et al., 2011). Cependant l'activation de *cupC* ne se fait pas via le régulateur de réponse RocA2 mais par le régulateur RocA1. Par une « cross-régulation » entre le senseur RocS1 et RocS2, les gènes *cupC* sont activés par RocA1 (Sivaneson et al., 2011). Les senseurs RocS1 et RocS2 permettent également l'expression des gènes *cupB*. Mais ce processus ne fait pas intervenir RocA1 et RocA2, il implique un autre facteur encore inconnu. Le régulateur Mvat qui régule *cupA* par variation de phase serait également impliqué mais de manière plus modérée dans la répression des gènes *cupC* et *cupB* (Vallet et al., 2004).

### 3.3.3. CupE

#### 3.3.3.1. Le système CupE

Le système CupE a été identifié sur le génome de PAO1 par homologie aux gènes *cupA2* (codant pour la chaperonne) et *cupA3* (codant pour la protéine « usher »). Le cluster *cupE* est composé de 6 gènes avec le gène *cupE5* qui code pour la protéine « usher » et le gène *cupE4* qui code pour la chaperonne périplasmique appartenant à la sous-famille FGS. Les quatre autres gènes codent pour des sous-unités structurales : une piline majeure CupE1, deux pilines mineures CupE2 et CupE3 et une adhésine CupE6. Toutes les sous-unités possèdent un domaine COG5430 qui est une caractéristique des sous-unités appartenant au clade  $\sigma$ . Le domaine COG5430 de CupE1, CupE2, CupE3 et CupE6 contient un sous-domaine SPCU (Spore coat protein U). L'adhésine CupE6 contient deux domaines SPCU, SPCU1 (aa 97-162) et SPCU2 (aa 257-312) (Figure 41). Un alignement des séquences montre que les sous-unités CupE1, CupE2 et CupE3 ne présentent pas de manière évidente le motif composé de quatre résidus hydrophobes P2-P5 en N-terminal permettant l'échange de brin donneur. La prédiction de la structure secondaire des sous-unités E1, E2 et E3 est identique, ce qui laisse penser que leur repliement sera similaire (Figure 41) (Giraud et al., 2011).



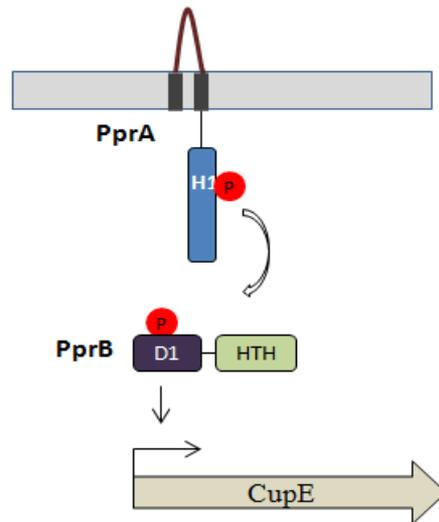
**Figure 41: Le cluster *cupE* chez *P. aeruginosa*.** Représentation schématique des sous-unités. CupE1 représente la piline majeure putative, CupE2 et CupE3 de putatives pilines mineures et CupE6 l'adhésine supposée. Les peptides signal sont représentés en noirs. Les séquences SCPU sont en gris et les domaines COG5430 sont délimités par une double flèche noire. (Giraud et al., 2011).

Une étude phylogénique a montré que *cupE* est un cluster mosaïque car seul le gène qui code pour la protéine « usher » a été acquis par un transfert vertical. Le cluster *cupE*, est présent dans les génomes de toutes les souches séquencées de *P. aeruginosa* (PAO1, PA14, LESB58, PA7). Chez *P. putida*, *P. fluorescens* et *P. entomophila*, il contient un gène en plus qui code pour une protéine avec une séquence SCPU (Giraud et al., 2011).

Les fimbriae CupE sont impliqués dans la formation de microcolonies et dans la maturation du biofilm en forme de champignon (Giraud et al., 2011).

### 3.3.3.2. La régulation du cluster *cupE*

La production des fimbriae CupE ainsi que la formation d'un biofilm ont été observées lors d'un prolongement d'une culture sur milieu solide ou liquide sans agitation. Cependant aucune expression de fimbriae CupE n'a été observée sur milieu liquide sous agitation. Ces résultats montrent que le cluster *cupE* est finement régulé. Des expériences de mutagenèse par transposition ont montré que le cluster *cupE* est régulé selon le système à deux composants PprAB (Figure 42) (Giraud et al., 2011). PprAB régule positivement le cluster *cupE*. PprA est une histidine kinase cytoplasmique classique (Wang et al., 2003) et PprB est le régulateur de réponse qui comporte un domaine receveur (D1) et un domaine de type fixation à ADN (HTH). Lorsque le régulateur de réponse PprB est activé par PprA, il se fixe sur le promoteur de *cupE* et induit son expression (Giraud et al., 2011). L'activation de l'expression de *cupE* se fait en présence d'un faible signal de l'oxygène et/ou l'absence de contraintes mécaniques perçues par PprA.



**Figure 42: Schéma de la régulation du cluster *cupE* par le système à deux composants PprAB.** (Réalisée d'après (Giraud and De Bentzmann, 2012).

Dans l'hypothèse où la concentration du c-di-GMP réprime l'expression de *cupE*, des expériences ont été réalisées. La surexpression de protéines qui contiennent un domaine EAL, comme RocR et PA2133 dans une souche qui surexprime PprB, entraîne une faible diminution de l'expression de *cupE*. D'autres expériences sont en cours afin de montrer si le c-di-GMP est impliqué dans la régulation de *cupE* (Giraud and De Bentzmann, 2012).

---

## 4. Vers la thérapie anti-adhésive

---

*Pseudomonas aeruginosa* représente la cause majeure de la morbidité et de la mortalité chez les patients adultes atteints de la mucoviscidose. Les infections chroniques provoquent un sévère déclin des fonctions respiratoires où seule la transplantation d'organe peut sauver la vie du patient.

Actuellement trois méthodes sont utilisées afin de prévenir les infections chroniques causées par *P. aeruginosa* : (1) La prévention des infections croisées par isolement des patients CF infectés de manière chronique et par l'utilisation de mesures d'hygiène drastiques, (2) l'administration sur une durée de trois semaines ou de trois mois d'une thérapie antibiotique agressive composée soit par la ciprofloxacine et la colistine nébulisée ou soit par la tobramycine nébulisée et (3) l'utilisation de la DNase (Pulmozyme®).

Lorsque les infections chroniques sont diagnostiquées, la thérapie consiste à l'inhalation quotidienne de colistine ou de tobramycine avec de la DNase. En plus de ce traitement, est administrée sur une période de 14-21 jours tous les trois mois, une forte dose d'antibiotiques antipseudomonal par intraveineuse (comme le carbapenem, l'aztreonam, la ceftazidime, la ciprofloxacine ou la piperacilline/tazobactame) (Döring et al., 2000; Cantón et al., 2005). Cette thérapie a pour but de ralentir le déclin des fonctions pulmonaires des patients CF ce qui augmente leur espérance de vie. Elle est adaptée selon l'âge et les stades de l'infection des patients ainsi que l'émergence de souches multi-résistantes au niveau des voies respiratoires (Döring et al., 2000). La thérapie a un effet positif au stade précoce de l'infection c'est-à-dire sur les microcolonies où les souches sont généralement non-mucoïdes, sensibles aux antibiotiques et présentent en faible densité (Rosenfeld et al., 2003) (Taccetti et al., 2005). Une ré-infection est cependant inévitable. Lorsque le biofilm est mature, *P. aeruginosa* devient difficile voire impossible à éradiquer par une antibiothérapie. Face à l'adaptabilité de *P. aeruginosa* vis-à-vis des antibiotiques, l'utilisation de nouvelles stratégies thérapeutiques devient une préoccupation majeure.

### 4.1. Relation entre lectines et pathogénicité de *P. aeruginosa*

Durant les processus d'infection et de formation du biofilm, l'étape cruciale qui permet la fixation du pathogène à la cellule hôte est l'adhésion. Cette étape nécessite la présence de différentes lectines qui vont reconnaître de manière spécifique et se fixer aux glycoconjugués présents à la surface des cellules hôtes. Les lectines peuvent ainsi être considérées comme des

facteurs de virulence et constituent des cibles thérapeutiques dans le développement de nouvelles molécules dites anti-adhésives. Cette stratégie est utilisée dans la nature comme par exemple les oligosaccharides présents dans le lait maternel qui protège les nourrissons contre les infections microbiennes (Newburg et al., 2005).

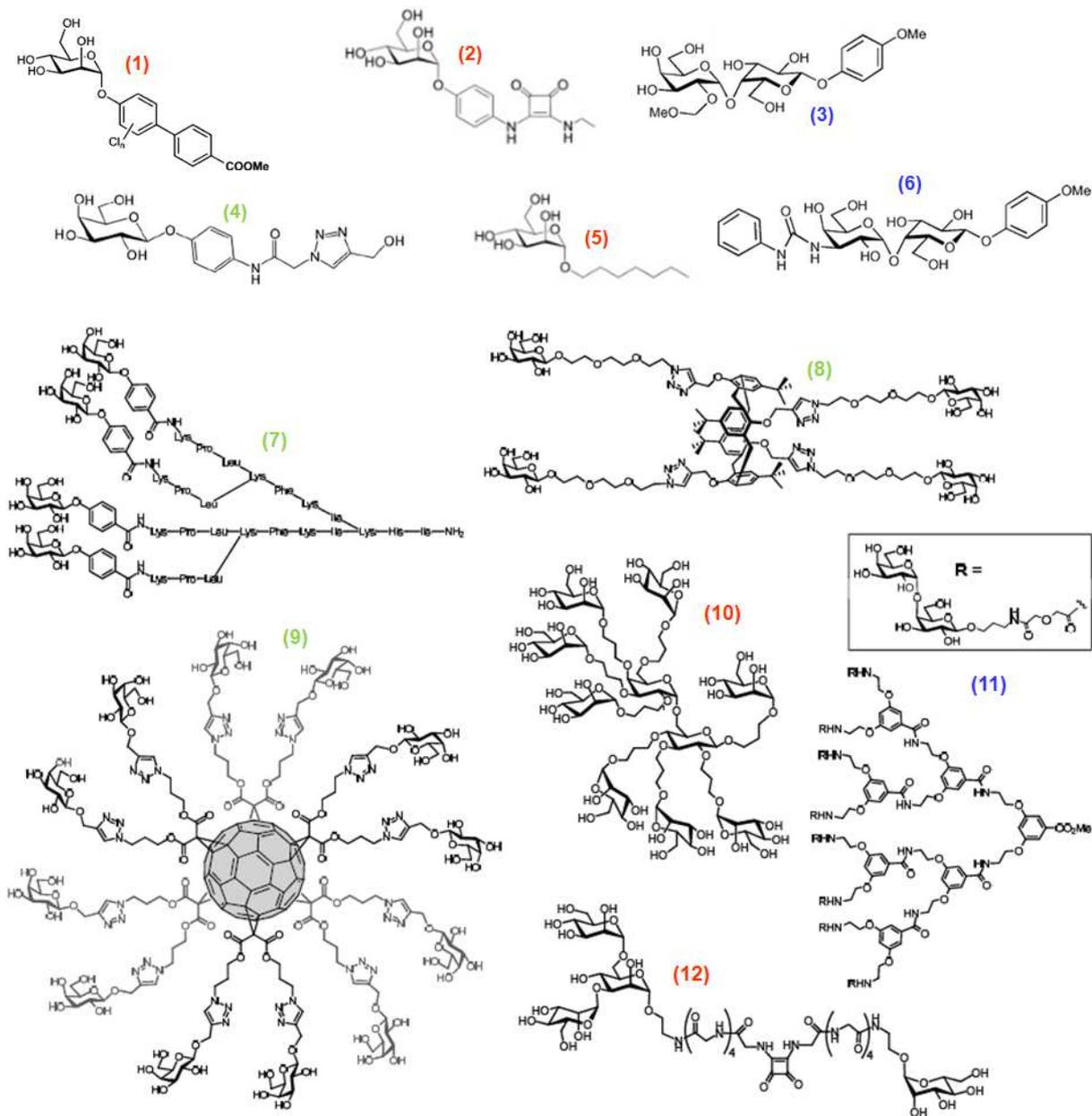
Plusieurs études ont montré que l'approche d'une thérapie anti-adhésive ciblant les lectines de *P. aeruginosa* permettait de réduire de manière considérable sa virulence. L'étude *in vivo* menée par Chemani montre que la simple utilisation de ligands monosaccharidiques permet la diminution du taux de mortalité chez les modèles murins utilisés. Dans cette expérience, les souris sont instillées par un inoculum de *P. aeruginosa* au niveau des poumons. La mortalité des souris est observée après 4 jours d'inoculation du pathogène. L'administration de l' $\alpha$ -méthylgalactoside et du GalNAc, des ligands de la lectine soluble PA-IL avant l'inoculation du pathogène permet un taux de survie de 30% des souris. En bloquant la lectine PA-IL, les monosaccharides administrés ont un rôle protecteur qui permet de réduire la destruction des tissus pulmonaires (Chemani et al., 2009).

En Allemagne, une étude réalisée chez des patients CF a montré que l'inhalation du galactose et du fucose (ligand de PA-IL et PA-III respectivement) permet la diminution de la quantité du pathogène ainsi que la réponse inflammatoire dans les voies respiratoires des patients. Bien que les effets secondaires du traitement sur un long terme n'ont pas pu être étudiés, le traitement a été bien toléré par les patients (Hauber et al., 2008).

Les résultats de cette première étude sont encourageants et montrent l'intérêt de la thérapie anti-adhésive. Il est important de souligner que cette étude a été réalisée avec les ligands naturels des lectines PA-IL et PA-III. L'utilisation de ligands de plus haute affinité appelée également glycomimétiques permettrait d'augmenter l'efficacité de ce type de thérapie.

## **4.2. Les glycomimétiques**

Les glycomimétiques sont des molécules synthétiques qui miment les carbohydrates présents à la surface des cellules mais qui présentent une affinité plus élevée pour la lectine cible (Bernardi and Cheshev, 2008).



**Figure 43: Exemple de différents glycomimétiques monovalents (1 à 6) et multivalents (7 à 12) dirigés contre l'adhésion bactérienne.** Les composés 1 (Klein et al., 2010), 2 et 5 (Hartmann et al., 2012), 10 (Nierengarten et al., 2010), 12 sont spécifiques de l'adhésine FimH d'*E. coli*, les composés 3 et 6 (Ohlsson et al., 2002, 2005), 11 sont spécifiques de l'adhésine PapG d'*E. coli*, et les composés 4 et 8 (Cecioni et al., 2012), 7 (Kadam et al., 2011), 9 (Cecioni et al., 2011) sont spécifiques de la lectine soluble PA-IL.

Les glycomimétiques monovalents sont composés d'un motif glucidique (mono, di ou oligosaccharide) souvent couplés à un aglycone (Figure 43). Leur association à un « cœur » via un bras espaceur permet une présentation multivalente, également retrouvés à la surface des cellules, des épitopes glucidiques (Figure 43). La synthèse de glycomimétiques tels que

les glycoclusters, les glycodendrimères ou encore les glyconanoparticules constitue également un moyen d'augmenter l'affinité d'un motif glucidique envers une lectine cible. Il existe différents types de glycomimétiques qui ont pour objectif d'inhiber l'adhésion des virus, des toxines de bactéries ou des bactéries.

Parmi les différents glycomimétiques dirigés contre l'adhésion bactérienne sont retrouvés ceux qui ciblent les lectines, FimH et PapG, impliquées dans la formation d'infection urinaire chez l'homme (Ohlsson et al., 2005; Klein et al., 2010; Nierengarten et al., 2010; Hartmann et al., 2012; Wellens et al., 2012) et les deux lectines solubles, PA-IL et PA-III, chez *P. aeruginosa*. Pour ces dernières, un grand nombre de glycomimétiques multivalents ont été synthétisés comme des glycodendrimères (Johansson et al., 2008; Chabre et al., 2011; Kadam et al., 2011), des glyconanoparticules (Reynolds et al., 2012) ainsi que des glycoclusters (Cecioni et al., 2012) possédant une affinité envers la lectine de l'ordre du nanomolaire. Récemment, une étude a permis d'identifier un glycomimétique peptide monovalent à une forte affinité pour les deux lectines solubles capable d'inhiber leur cytotoxicité (Gustke et al., 2012). D'autres études ont prouvé la capacité de certains glycomimétiques comme le glycodendrimère FD2 (Johansson et al., 2008) et les glycodendrimères GalAG0/1/2 et GalBG0/1/2 (Kadam et al., 2011) d'inhiber complètement la formation d'un biofilm mais aussi de disperser un biofilm déjà établi. Ces résultats montrent le potentiel de ces molécules dans une thérapie antiadhésive voire antibiofilm.

### **Objectifs de thèse :**

Un rôle dans la formation ou la maturation du biofilm a été démontré pour les lectines solubles et les adhésines Cup de *P. aeruginosa*. La résolution de la structure en trois dimensions de la lectine complexée à son ligand naturel permet d'analyser et de mieux comprendre à un niveau atomique les interactions qui sont mises en place. Elle facilite et aide à la conception de molécules antiadhésives qui est le but final de cette étude.

Parmi les adhésines de type Cup chez *P. aeruginosa*, aucun ligand naturel n'a été identifié et aucune structure cristallographique n'a été résolue. Ces protéines sont difficiles à étudier du fait de leur appartenance à des organelles et aux nombreux problèmes rencontrés afin de les obtenir sous la forme recombinante, soluble et homogène indispensable pour les essais de cristallogenèse et une étude structurale. Une grande partie de mes travaux de thèse consistait à caractériser biochimiquement et structurellement les deux adhésines CupE6 et CupB6.

Un deuxième volet de mes travaux s'inscrivait dans la continuité des travaux de thèse du Dr Bertrand Blanchard et consistait à étudier l'affinité de différents glycomimétiques monovalents pour la lectine PA-IL.

Enfin j'ai eu l'opportunité d'étudier les paramètres thermodynamiques caractérisant les interactions entre la lectine de légumineuse PELa de *Platypodium elegans* et certains de ces ligands puis de résoudre la structure de PELa complexée à un heptasaccharide.

# **Matériels et Méthodes**



---

## 5. Techniques de biologie moléculaire

---

### 5.1. Obtention des différentes constructions plasmidiques

#### 5.1.1. Clonage des gènes *cupE6*, *cupE4*, *cupB6* et *birA*

##### 5.1.1.1. Clonage du domaine adhésine de *cupE6* et *cupB6*

Les vecteurs pProEX-Htb (Life Technology), pET-M30 (EMBL), pBAD-M41 (EMBL), pET32-TEV (Novagen) ont été utilisés pour réaliser les tests d'expression des domaines adhésines CupE6 et CupB6 (Figure 45). Des constructions plasmidiques correspondant aux gènes *cupE6* et *cupB6* ont été fournies par l'équipe de Sophie de Bentzmann à Marseille (LISM-CNRS-IFR88). A partir de ces constructions plasmidiques, les différentes régions choisies pour exprimer le domaine adhésine de CupE6 ou CupB6 ont été amplifiées par la polymérase HotStar Fidelity (Qiagen), selon les indications du fournisseur. Les amorces utilisées figurent dans le tableau 1. Après digestion enzymatique des produits PCR (polymerase chain reaction) et du plasmide, une ligation avec le kit TAKARA (fournisseur) permet d'obtenir la construction souhaitée.

##### 5.1.1.2. Clonage du gène *cupE6* et *cupE4* dans pETDuet

Pour co-exprimer le complexe adhésine CupE6/ chaperonne CupE4, les gènes *cupE6* et *cupE4* ont été clonés dans le vecteur pETDuet (Novagen). La polymérase HotStar Fidelity (Qiagen) est utilisée pour l'amplification des gènes selon les instructions données. Les amorces utilisées figurent dans le tableau 1. Le vecteur pETDuet comporte deux sites multiples de clonage (Figure 45). L'adhésine est clonée de façon à ce qu'une étiquette hexahistidine soit placée à son extrémité N-terminale.

##### 5.1.1.3. Clonage du gène *birA*

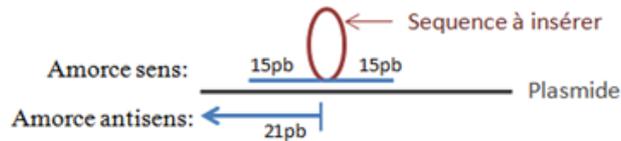
Le gène qui code pour la biotine ligase BirA d'*E. coli* (GenBank: EKI34486.1) est cloné dans le plasmide pCDFDuet (Novagen) (Figure 45) selon le même protocole que précédemment. L'ADN génomique de la souche XL1 d'*E. coli* est utilisé comme matrice pour l'amplification du gène *birA*. Les amorces utilisées figurent dans le tableau 1.

Non de l'amorce	Séquence de l'amorce
<b>Domaine adhésine de CupE6</b>	
CupE6 25 <i>NcoI</i>	5'-GC <b>ACCATGG</b> CCTGCACCACCAGCTCCG-3'
CupE6 37 <i>NcoI</i>	5'-AT <b>GCCATGG</b> CCAGCCTGAGCTCGTTCACCG-3'
CupE6 146 <i>HindIII</i>	5'-CAG <b>AAGCTTT</b> CAGGTGGCGGTGCGCAGGTACAG-3'
CupE6 151 <i>HindIII</i>	5'-CT <b>AAGCTTT</b> CACGGCAGGGTCACCCCGGT-3'
CupE6 170 <i>HindIII</i>	5'-GT <b>CAAGCTTT</b> CAGCCGCGCACAGGTG-3'
<b>Domaine adhésine de CupB6</b>	
CupB6 39 <i>NcoI</i>	5'-ACT <b>CCATGG</b> CCCTGAACGAGGTCG-3'
CupB6 49 <i>NcoI</i>	5'-CAT <b>CCATGG</b> CCGACAATGGCAAAGG-3'
CupB6 226 <i>HindIII</i>	5'-CG <b>AAGCTTT</b> CAGCCGATCAGCGAATAG-3'
CupB6 233 <i>HindIII</i>	5'-CG <b>AAGCTTT</b> CAGCCGATAGCGATGG-3'
<b>Adhésine CupE6</b>	
CupE6 25 <i>BamHI</i>	5'-CT <b>GGGATCC</b> GGCCTGCACCACCAGCTCC-3'
CupE6 37 <i>BamHI</i>	5'-CT <b>GGGATCC</b> GAGCCTGAGCTCGTTCACC-3'
CupE6 315 <i>HindIII</i>	5'-CG <b>AAGCTTT</b> CAGTAGGTGAGGATCACCCGCACGG-3'
<b>Chaperonne CupE4</b>	
CupE4 <i>NdeI</i>	5'-GCAGAC <b>CATATG</b> GCCAGTGCGGTGCTGATCT-3'
CupE4 <i>XhoI</i>	5'-CG <b>CTCGAGT</b> CAGTACCCGGGAACGCTG-3'
<b>BirA</b>	
BirA <i>NcoI</i>	5'-ACT <b>CCATGG</b> CCCGGGCGGCTATTA-3'
BirA <i>XhoI</i>	5'-CG <b>CTCGAGT</b> TATTTTCTGCACTACGCAG-3'

**Tableau 1: Séquences des amorces utilisées dans cette étude.** Les sites de restriction sont en gras.

### 5.1.2. Le plasmide pETDuet-TEV-Biot

Le design des amorces permettant l'insertion d'une séquence avec un nombre important de bases nucléotidiques est réalisé selon la méthode GeneTailor™ Site-Directed Mutagenesis System développée par Invitrogen. Les amorces nécessaires à l'obtention du plasmide pETDuet-TEV-Biot sont indiquées dans la (Figure 44). L'amplification par PCR se fait avec la polymérase PfuUltraII fusion (Agilent) selon les instructions fournies. Après les cycles d'amplification, l'ADN matriciel méthylié est digéré par l'enzyme *DpnI* (Promega) et le plasmide comportant la séquence insérée est transformé dans la souche d'*E.coli* XL1-blue. La séquence qui correspond au peptide reconnu par la protéase TEV (tabacco etch virus) est d'abord insérée dans le plasmide pETDuet. Puis après l'obtention du plasmide pETDuet-TEV la séquence du peptide reconnue par la biotine est insérée.



<p><b>Amorces pour obtenir pETDuet-TEV</b></p> <p>5'-CACAGCCCAACGACCGAAAACCTGTATTTTCAGGGCCAGGATCCGAATTCG-3'</p> <p>5'-GGTCGTTGGGCTGTGGTGATG-3'</p>
<p><b>Amorces pour obtenir pETDuet-TEV-Biot</b></p> <p>5'-CTGTATTTTCAGGGCGCCGGCCTGAACGACATCTTCGAGGCTCAGAAAATCG</p> <p>AATGGCACGAACAGGATCCGAATTCGAGCTCG-3'</p> <p>5'-CTGAAAATACAGGTTTTTCG-3'</p>

**Figure 44: Séquences des amorces utilisées pour obtenir le plasmide pETDuet-TEV-Biot.** La séquence à insérer est en rouge. Les bases nucléiques qui permettent l'hybridation au plasmide sont en bleu. Les bases nucléiques en vert indiquent une erreur qui s'est produite dans le design des amorces pETDuet-TEV. Ces bases nucléiques sont à insérer hors elles ont été considérées comme appartenant au vecteur pETDuet.

### 5.1.3. Tests d'expression

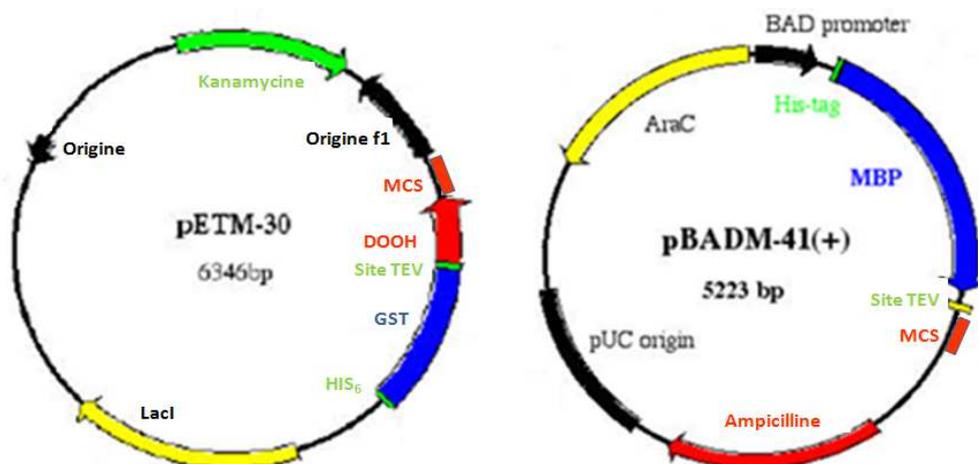
Toutes les constructions réalisées ont été séquencées puis transformées dans différentes souches d'expression. Les caractéristiques des souches d'*E. coli* utilisées sont résumées dans le tableau 2. L'expression de protéines peut être influencée par différents paramètres tels que la température de culture, la durée de l'expression, la densité optique ( $DO_{600}$ ) au moment de l'induction, la concentration de la molécule utilisée pour induire (IPTG (Iso-Propyl- $\beta$ -D-Thiogalactopyranoside) ou arabinose), le milieu utilisé ainsi que le type de vecteur. Afin de tester ces différents paramètres, des tests d'expression ont été réalisés suivant ce type de protocole. Une préculture liquide en milieu Luria Bertani (LB) ou Terrific Broth (TB) contenant les antibiotiques appropriés est réalisée à 37°C pendant une nuit sous agitation (160 rpm). La culture estensemencée à partir de la préculture jusqu'à une  $DO_{600}$  d'environ 0,1. Selon le vecteur d'expression, l'induction se fait à une  $DO_{600}$  de 0,6- 1,2 avec de l'IPTG (de 0,2 mM à 1 mM) ou 0,2% d'arabinose. L'expression de la protéine peut se faire soit à 30°C pendant 3h soit à 16°C sur la nuit. Avec le milieu autoinductible, la culture est égalementensemencée à partir de la préculture. Elle est réalisée sur la journée à 37°C puis sur la nuit à 16°C. A la fin de la culture, les cellules sont centrifugées pendant 15 min à 5000g à 4°C. Les cellules peuvent être directement cassées pour une purification ou elles sont congelées à -20°C.

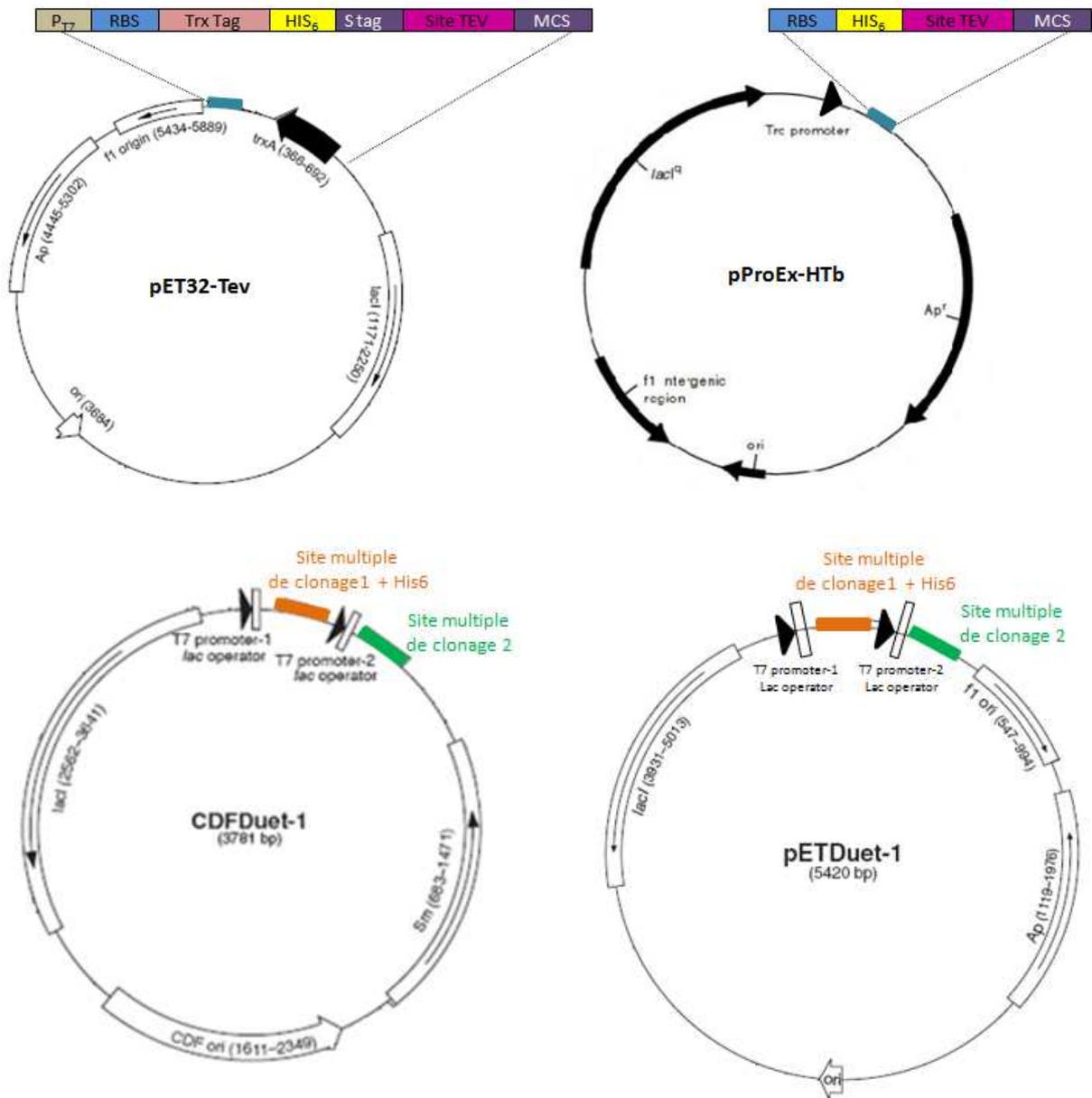
Pour le test d'expression, 1ml de culture est prélevé avant et après induction des protéines. Les cellules sont centrifugées 2 min à 13000g et resuspendues dans 25  $\mu$ l de tampon et 25  $\mu$ l

de tampon dénaturant 2X. Pour analyser la solubilité de la protéine, l'échantillon prélevé après induction est repris dans 1ml de tampon puis soniqué. Le culot qui correspond à la fraction insoluble est repris dans 25 µl de tampon et 25 µl de tampon dénaturant 2X. Le surnageant qui correspond à la fraction soluble peut être concentré jusqu'à 25 µl avant d'être dénaturé par le tampon 2X. Les échantillons sont chauffés à 100°C pendant 5 min avant d'être déposés sur gel de polyacrylamide en conditions dénaturantes. Des gels de 15 % sont réalisés pour l'analyse des protéines de 10 à 20 kDa. Pour l'analyse des protéines de 60 à 20 kDa, des gels de 12 % sont utilisés.

Souches	Caractéristiques	Résistance	Référence
BL21 (DE3) RIL	Souche déficiente des protéases OmpT et Lon. L'appellation (DE3) signifie que la souche possède une copie chromosomale du gène codant pour l'ARN polymérase T7 sous contrôle du promoteur <i>lacUV5</i> et donc inducible à l'IPTG. La souche possède un plasmide qui code pour les ARNt des codons rares chez <i>E.coli</i> : AGA, AGG, AUA et CUA.	Chloramphenicol	Stratagene
BL21 (DE3) pLysS	Souche qui possède un plasmide qui code pour les ARNt AGA, AGG, AUA, CCC, GGA, CGG et CUA.	Chloramphenicol	Stratagene
Artic Express (DE3) RIL	Cette souche permet l'expression de deux chaperonnes, Cpn10 et Cpn60 qui aide au repliement de la protéine à basse température. La souche possède un plasmide qui code pour les ARNt des codons AGA, AGG, AUA et CUA.	Gentamycine Streptomycine	Stratagene
Artic Express (DE3) pLysS	Souche qui possède un plasmide qui code pour les ARNt AGA, AGG, AUA, CCC, GGA, CGG et CUA.	Chloramphenicol Gentamycine	Stratagene
Rosetta gami2 (DE3) pLysS	Cette souche possède des mutations sur les gènes des enzymes thioredoxin reductase ( <i>trxB</i> ) et glutathione reductase ( <i>gor</i> ), ce qui favorise la formation des ponts disulfures dans le cytoplasme. Souche qui possède un plasmide qui code pour les ARNt AGA, AGG, AUA, CCC, GGA, CGG et CUA.	Chloramphenicol	Novagen
AVB100	Cette souche comporte dans son chromosome le gène qui code pour BirA inducible à l'arabinose. Cette souche convient pour les plasmides inducibles à l'IPTG.	Pas de résistance	Avidity

Tableau 2: Différentes souches d'*E.coli* utilisées pour l'expression de protéines recombinantes





**Figure 45: Carte des plasmides utilisés pour l'obtention des différentes constructions.** Le vecteur pET32a a été modifié par le Dr Annabelle Varrot pour donner le vecteur pET32-TEV. Le site de clivage pour la protéase entérokinase a été remplacé par un site de clivage pour la protéase TEV.

---

## 6. Caractérisation biochimique et biophysique

---

### 6.1. Purification des protéines recombinantes

#### 6.1.1. Préparation des extraits bruts

Le culot bactérien obtenu après la culture est repris dans cinq fois le volume de tampon d'équilibration. Des inhibiteurs de protéases (Complete<sup>TM</sup> Protease Inhibitor) sont ajoutés avant la lyse mécanique des cellules par pression à 1,8 kbars avec un One shot cell Disrupter (Constant System Ltd). Avant cette étape, les cellules peuvent être lysées par 1 mg.ml<sup>-1</sup> de lysozyme pendant 30 min dans la glace en agitant de temps en temps. Pour les protéines peu solubles, 1 % de triton X100 peut être ajouté après la lyse pendant 30 min sous agitation avant l'étape de purification. Le lysat bactérien est ensuite centrifugé à 50 000g pendant 30 min à 4°C.

#### 6.1.2. Purification du domaine adhésine CupB6

##### 6.1.2.1. Chromatographie sur résine de Nickel ou de Zinc

Le surnageant contenant les protéines solubles est chargé sur une colonne HisTrap ou sur une résine de zinc (GE Healthcare) préalablement équilibrée avec du tampon de lyse. Le tampon est souvent composé de 20 mM HEPES pH 7,5- 500 mM NaCl- 20 mM Imidazole. Si les protéines sont solubles mais agrègent, 1M d'urée et 10% de glycérol peuvent être rajoutés. L'ajout de 50 mM d'arginine et de lysine peut aider à une meilleure purification dans notre cas. Après un lavage avec le tampon d'équilibration, la protéine est éluée soit par un gradient de 20- 40 min à 1 ml.min<sup>-1</sup> soit par pallier à différentes concentrations d'imidazole. Le tampon d'éluion est composé 20 mM HEPES pH 7,5- 500 mM NaCl- 500 mM Imidazole. Les colonnes HisTrap sont régénérées après cinq utilisations selon les instructions. La colonne chargée par les ions zinc nécessite une régénération à chaque utilisation.

##### 6.1.2.2. Chromatographie sur des billes de glutathion

La purification des protéines recombinantes qui comportent une étiquette GST (glutathione S-transferase) sont purifiées sur des billes glutathion Sepharose 4B (GE Healthcare). Environ 1 ml de billes pour 100-500 mg d'extrait cellulaire sont équilibrées avec du PBS à pH 7,4 (Phosphate Buffer Saline) avant d'être incubées 4 h à 4°C sous agitation avec le surnageant.

Les billes sont lavées avec 40 volumes de billes avec le tampon d'équilibration. L'élution se fait avec le tampon suivant : 20 mM HEPES pH 8- 10 mM glutathion.

#### 6.1.2.3. Chromatographie sur Amylose

Les protéines qui comportent une fusion avec la maltose binding protein (MBP) à l'extrémité N-terminale sont purifiées sur une colonne d'amylose (GE Healthcare) pré-équilibrée avec le tampon 20 mM HEPES pH 7,5- 200 mM NaCl. L'injection de la protéine se fait à 0,3 ml.min<sup>-1</sup>. Après un lavage avec le tampon ci-dessus, l'élution est réalisée avec 20 mM HEPES pH 7,5- 200 mM NaCl- 10 mM de maltose.

#### 6.1.2.4. Clivage de l'étiquette par la TEV

Avant le clivage de l'étiquette, la protéine recombinante est dialysée contre 20 mM HEPES pH 7,5- 150 mM NaCl puis incubée toute la nuit à 10°C avec la protéase TEV diluée au 1/20 et 1 mM de DTT. Une étape de purification sur colonne de nickel est réalisée pour obtenir la protéine clivée pure. La protéase et l'étiquette sont retenues par la résine alors que la protéine clivée n'est pas retenue.

### 6.1.3. Purification du complexe CupB6/CupB4

La purification du complexe CupB6/B4 se fait en deux étapes. La première étape de purification est réalisée sur une résine de nickel (GE Healthcare). Le tampon d'équilibration est composé de 20 mM HEPES pH 7,5- 500 mM NaCl- 20 mM d'imidazole et 50 mM d'arginine et de lysine. Après une étape de lavage, le complexe est élué par un gradient de 20 à 500 mM d'imidazole. Après l'analyse sur gel SDS-PAGE 12%, les fractions sont concentrées par filtration sur Vivaspinn 30,000 MWCO, (Sartorius). La seconde étape de purification consiste en une filtration sur superdex 75 Hiload 16/60, (GE healthcare). Le tampon utilisé est 20 mM MES pH 6- 100 mM NaCl. Des injections de 500 µl à une concentration maximale de 8 mg.ml<sup>-1</sup> et à un débit de 0,5 ml.min<sup>-1</sup> permettent d'obtenir le complexe pur.

### 6.1.4. Purification de la lectine PA-IL

Le gène *lecA* qui code pour la protéine PA-IL est cloné dans le vecteur pET 25b (Novagen). La lectine PA-IL est exprimée dans la souche BL21 (DE3) d'*E. coli*. Après une culture en LB où les cellules ont été induites à une DO<sub>600</sub> de 0,6 avec 0,5 mM d'IPTG, la protéine est purifiée par chromatographie d'affinité sur la résine Sépharose 4B (GE Healthcare) dans le

tampon 20 mM Tris-HCl pH 7,5 et 100  $\mu$ M de CaCl<sub>2</sub>. L'élution se fait avec 20 mM Tris-HCl pH 7,5- 100  $\mu$ M de CaCl<sub>2</sub>- 1 M de NaCl. Avant d'être lyophilisée et conservée à 4°C, PA-IL est dialysée contre de l'H<sub>2</sub>O MilliQ avec 5  $\mu$ M de CaCl<sub>2</sub> pendant 4 jours.

### **6.1.5. Purification de la lectine PELa**

L'expression de PELa se fait dans la souche BL21 (DE3) d'*E. coli* contenant le plasmide pET-29-PELa. L'induction se fait avec 0,5 mM d'IPTG à une DO<sub>600</sub> de 0,6. Après une expression à 30°C pendant 4 h, les cellules sont centrifugées et reprises dans 50 mM Tris-HCl pH 7,5- 150 mM NaCl- 5 mM CaCl<sub>2</sub>- 5 mM MnCl<sub>2</sub>. Une étape de renaturation des corps d'inclusions est nécessaire avant la purification de PELa. Après deux lavages du culot avec 50 mM Tris-HCl pH 8- 4 M Urée- 0,5 M NaCl- 1 mM EDTA- 0,5 % de triton X100, PELa est dénaturée et solubilisée dans 50 mM Tris-HCl pH 8- 8 M Urée- 10 mM DTT. L'étape de renaturation se fait progressivement en diluant la protéine à 1 mg.ml<sup>-1</sup> 20 fois dans 50 mM Tris-HCl pH 7,5- 240 mM NaCl- 2 mM CaCl<sub>2</sub>- 2 mM MnCl<sub>2</sub>- 10 mM KCl- 1 mM DTT. Après 1 h à 10°C suivie d'une centrifugation, PELa est purifiée sur colonne d'affinité D-mannose-Sepharose (Sigma-Aldrich) préalablement équilibrée avec 50 mM Tris-HCl pH 7,5- 150 mM NaCl- 5 mM CaCl<sub>2</sub>- 5 mM MnCl<sub>2</sub>. L'élution se fait par 200 mM de mannose additionné au tampon d'équilibration.

## **6.2. Détermination des domaines protéiques stables de CupB6/B4**

La digestion par des protéases de manière ménagée sur une protéine permet de définir des parties flexibles digérées rapidement et les parties plus structurées qui sont moins accessibles à la digestion. Différentes protéases ont été testées sur le complexe CupB6/B4 à une concentration de 4 mg.ml<sup>-1</sup> dans 20 mM HEPES pH 7,5- 100 mM NaCl. Les différentes conditions sont regroupées dans le tableau 3. La réaction a été arrêtée par ajout de tampon dénaturant 2X et analysée sur gel SDS-PAGE. Les profils de digestion permettant d'obtenir des parties stables du complexe sont ensuite envoyés à la plateforme de spectrométrie de masse de l'Institut de Biologie Structurale (IBS). La technique de LC-ESI-TOF est employée sur les échantillons protéolysés liquides afin de déterminer la masse et de donner une estimation de la séquence en acides aminés des fragments présents.

Protéases	Sites de coupures spécifiques	Ratio	Temps
Chymotrypsine	Au niveau de la liaison carboxylique de la tyrosine, du tryptophane, de la phénylalanine ou de la leucine	1/100 1/250 1/500 1/750	15 min 1h
Trypsine	Au niveau de la liaison carboxylique de la lysine et de l'arginine		
Endopeptidase Glu-C	Au niveau de la liaison carboxylique de l'acide glutamique		
Protéase K	Au niveau du carboxyle d'un acide aminé à chaîne latérale hydrophobe ou aromatique	1/500 1/1000 1/1500 1/10000	10 min 15 min 30 min 1h

**Tableau 3: Conditions testées pour la protéolyse ménagée de CupB6/B4 réalisée à 37°C.**

### 6.3. Diffusion dynamique de la lumière (DLS)

La technique de DLS mesure des fluctuations d'intensité en fonction du temps qui apparaissent lorsque des particules sont soumises au mouvement Brownien. Plus une molécule est volumineuse, moins son déplacement en solution est rapide. Les particules en mouvement vont interagir et disperser le faisceau dans toutes les directions. L'analyse de la lumière diffusée va être recueillie par un détecteur et permet de déduire un coefficient de diffusion des particules qui est converti en distribution de taille puis en rayon hydrodynamique. Il est également possible d'en déduire le taux de polydispersité de la molécule, qui représente une déviation standard de la taille de la particule reportée à son rayon. Cette valeur permet d'apprécier le taux d'agrégation et d'oligomérisation de la protéine.

Les mesures ont été effectuées avec 50 µL de protéine pure dans son tampon natif dans des cuves en quartz (Hellma) grâce à un système « Zetasizer nano series » (Malvern Instrument).

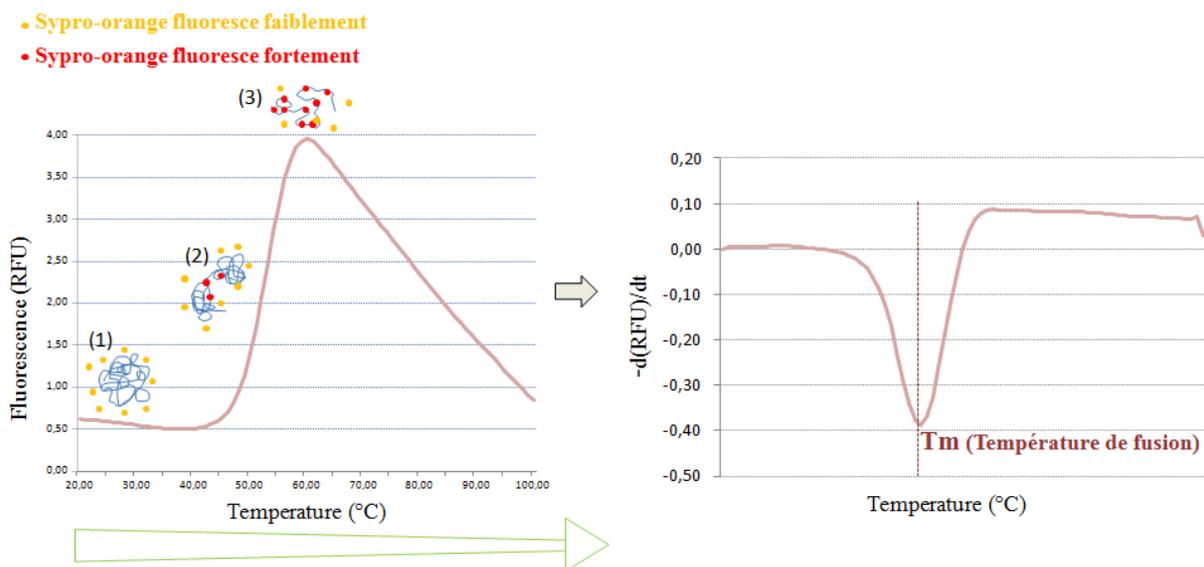
### 6.4. Thermal Shift Assay (TSA)

#### 6.4.1. Principe de la méthode Thermal Shift Assay

L'étude de la destruction thermique des protéines (TSA) permet d'obtenir des informations sur la stabilité des protéines grâce à la mesure de la température de fusion ( $T_m$ ). La température de fusion d'une protéine correspond à la température autour de laquelle la protéine va se dénaturer. Plus la température de fusion est élevée, plus la protéine est considérée comme stable.

Durant l'expérience de TSA, un gradient de température de 20 à 100°C à raison de 1 degré par minute est induit sur un échantillon protéique mélangé à un fluorophore (Figure 46). L'utilisation d'un fluorophore capable de fluorescer uniquement dans un environnement hydrophobe va permettre d'obtenir une cinétique de dénaturation de la protéine ainsi que de déterminer le  $T_m$  de la protéine. En effet tant que la protéine est correctement repliée, les acides aminés hydrophobes ne sont pas accessibles au fluorophore. La fluorescence mesurée reste faible. Puis lorsque la température augmente, la dénaturation progressive de la protéine rend les acides aminés hydrophobes accessibles au fluorophore. Une augmentation de la fluorescence est alors observée jusqu'à l'obtention d'un pic qui correspond au  $T_m$  de la protéine. Le calcul de la dérivée de la mesure de la fluorescence en fonction de la température facilite la lecture du  $T_m$  de la protéine d'intérêt.

Les études de TSA peuvent être utilisées pour optimiser le tampon optimal de la protéine et les conditions de cristallisation d'une protéine (Phillips and De la Peña, 2011) mais également aider à choisir la température de cristallisation la plus adéquate (Dupeux et al., 2011).



**Figure 46: Cinétique de la fluorescence qui permet de définir la température de fusion ( $T_m$ ) par une étude de dénaturation thermique (TSA).** (1) A basse température, le fluorophore spécifique des résidus hydrophobes émet une faible fluorescence car la protéine est correctement repliée. (2) L'augmentation de la fluorescence est due à l'accès du fluorophore aux résidus hydrophobes de la protéine déstructurée. (3) Obtention d'un pic qui correspond au  $T_m$  de la protéine.

### 6.4.2. Protocole permettant l'optimisation des conditions de cristallisation par TSA

Les mesures de TSA sont réalisées sur un appareil de PCR en temps réel de type IQ5 commercialisé par Bio-Rad.

Pour définir le tampon optimal du complexe, le mélange est composé de 12,5 µl de solution mère 2X du tampon à tester, de 5 µl de CupB6/B4 à une concentration finale de 0,3 mg.ml<sup>-1</sup> mélangé à du Sypro Orange 50X (Molecular Probes), puis 7,5 µL d'H<sub>2</sub>O.

Une fois le tampon optimal déterminé, différents additifs (Hampton Research) sont testés. Le mélange comporte 5 µl de CupB6/B4 à une concentration finale de 0,3 mg.ml<sup>-1</sup> mélangé à 50X Sypro Orange (Molecular Probes), 2,5 µL des additifs 10X puis 17,5 µL du tampon optimal.

Toutes les mesures sont réalisées en triplicate sur des plaques 96 puits.

## 6.5. La diffusion des rayons X aux petits angles (SAXS pour Small-angle X-ray scattering)

### 6.5.1. Introduction au SAXS

La diffusion aux petits angles est une méthode puissante pour analyser la conformation globale à basse résolution de macromolécules biologiques en solution.

Les expériences de SAXS requièrent seulement quelques milligrammes de protéines mais l'échantillon doit être très pur, monodisperse et rester soluble à haute concentration. Lorsque l'échantillon ou le tampon est exposé aux rayons X, les modulations de l'intensité diffusée sont enregistrées par un détecteur. L'agitation thermique des protéines en solution conduit à une intensité diffusée isotrope qui dépend du module du vecteur de diffusion  $q$  ( $q=4\pi\sin(\theta)/\lambda$ ), où  $2\theta$  est l'angle de diffusion et  $\lambda$  la longueur d'onde du faisceau incident) (Figure A) :

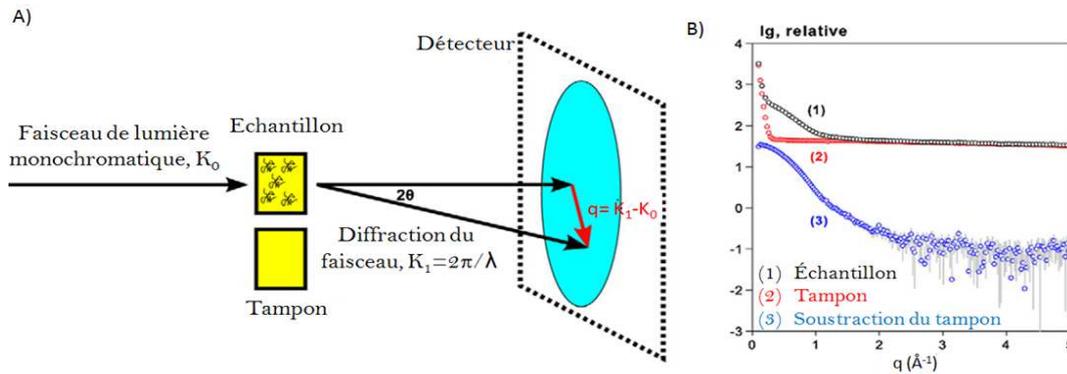
$$I(q) = [I(q)]_{\Omega} = [A(q).A(q)]_{\Omega} \quad (\text{équation 1})$$

avec l'amplitude diffusée  $A(q)$  qui est la transformée de Fourier de l'excès de densité électronique.

$$A(q) = \zeta[\rho(r)] = \int \Delta\rho(r) \exp(iqr) dr \quad (\text{équation 2})$$

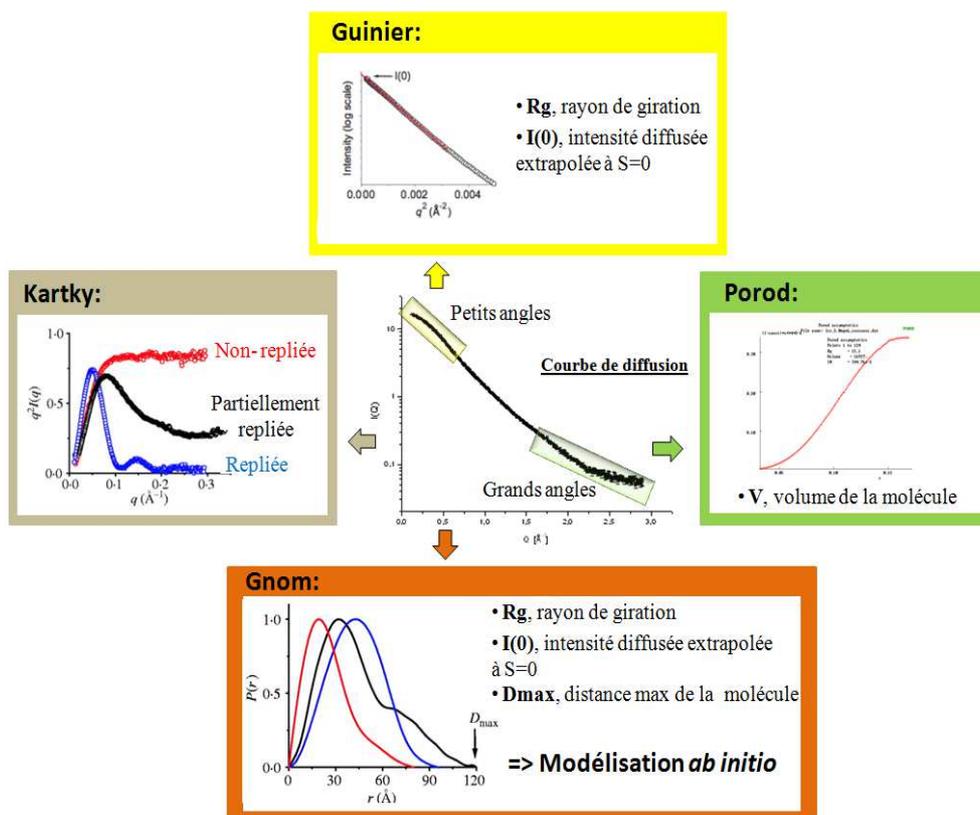
où  $\Delta\rho(r) = \rho(r) - \rho_s$ ,  $\rho(r)$  et  $\rho_s$  sont les densités électroniques de la particule et du solvant, respectivement et  $[I(q)]_{\Omega}$  la position de la moyenne sphérique.

Le profil de diffusion est représenté typiquement comme la courbe radialement moyennée  $I(q)$  en une dimension. Après la soustraction de la diffusion du tampon, l'intensité diffusée corrigée du bruit de fond  $I(q)$  est proportionnelle à la diffusion d'une particule moyennée dans toutes les orientations (Figure 47 B). Ainsi la courbe de diffusion est une moyenne de toutes les conformations qu'une protéine adopte en solution.



**Figure 47: A) Représentation schématique d'une expérience de SAXS. B) Exemple d'une courbe de diffusion obtenue après soustraction du tampon. (D'après (Mertens and Svergun, 2010)).**

A partir de la courbe de diffusion plusieurs informations peuvent être déterminées telles que la taille, la forme, l'état de repliement, l'existence de flexibilité entre des domaines, l'agrégation de la protéine (Figure 48).



**Figure 48: Différentes informations que la courbe de diffusion permet de fournir.**

Cette étude permet une meilleure caractérisation de la protéine ou du complexe étudié qui constitue des informations utiles à la cristallisation ou au comportement de la protéine dans la solution.

### 6.5.2. Paramètres structuraux globaux

Tous les programmes utilisés au traitement des données de SAXS, sont inclus dans la suite ATSAS 2.4 et peuvent être téléchargés à l'adresse suivante : <http://www.embl-hamburg.de/biosaxs/software.html>.

Les premières analyses des données se font par le programme PRIMUS (Konarev et al. 2003) qui permet de soustraire les spectres des tampons des spectres de protéines, de normaliser les spectres par la concentration, de superposer et de moyenniser les spectres ainsi que de déterminer le rayon de giration et de  $I(0)$  à partir de la représentation de Guinier.

#### 6.5.2.1. La relation de Guinier et le rayon de giration

Pour une solution monodisperse de macromolécules globulaires, l'équation de Guinier est définie par:

$$\ln I(q) = \ln I(0) - \frac{1}{3} Rg^2 q^2 \quad (\text{équation 3})$$

L'ordonnée à l'origine est  $\ln I(0)$  et la pente  $Rg^2/3$ . Le paramètre  $I(0)$  permet de calculer la masse moléculaire de la protéine et d'estimer la concentration de l'échantillon.

$Rg$  est le rayon de giration de la particule, c'est-à-dire la distance quadratique moyenne de tous les points de la particule au centre de gravité pondéré par la masse diffusante (le contraste de densité électronique) en chaque point :

$$Rg^2 = \frac{\int_V r^2 (\rho(r) - \rho(s)) dr}{\int_V (\rho(r) - \rho(s)) dr} \quad (\text{équation 4})$$

Le rayon de giration mesure le degré de (non)sphéricité de la particule en solution. Pour un volume donné et de masse moléculaire donnée, une protéine qui a un rayon de giration très petit correspond à une sphère de même volume. Plus le rayon de giration est grand, à volume constant, moins la particule est sphérique.

#### 6.5.2.2. Détermination de la masse moléculaire

L'intensité  $I(0)$  diffusée à  $q=0$  normalisée par la concentration  $c$  de la solution est proportionnelle à la masse molaire  $M$  de la particule étudiée :

$$\frac{I(0)}{c} = \frac{MN_A}{\mu^2} (1 - \rho(s)\psi) \text{ (équation 5)}$$

Avec  $N_A$  le nombre d'Avogadro,  $\rho(s)$  la densité électronique du solvant,  $\Psi$  le volume électronique spécifique moyen et  $\mu = M/n$  avec  $n$  le nombre d'électron ( $\mu_{\text{protéine}} = 1,87$  g/mol/électron).

Pour toutes les protéines,  $N_A(I - \rho(s)\Psi)^2 / \mu^2$  peut être considéré comme une constante  $k$ .

L'expérience est calibrée avec une protéine de masse moléculaire connue comme la BSA (Albumine de sérum bovin), ce qui permet de déterminer la masse moléculaire de la protéine étudiée grâce à l'équation :  $M = I(0)/kc$ .

### 6.5.2.3. Fonction de distribution des distances $P(r)$

La fonction  $P(r)$  représente l'histogramme des distances entre toutes les paires de points de la particule. La fonction  $P(r)$  est nulle lorsque la distance entre deux points est supérieure à  $D_{\text{max}}$ , la dimension maximale de la particule. La fonction  $P(r)$  est calculée à partir de  $I(s)$  en utilisant la méthode de transformée de Fourier indirecte (Glatter 1977).

$$P(r) = \frac{1}{\pi} \int_0^{\infty} r \frac{q}{2\pi} I(q) \sin rq. dq \text{ (équation 6)}$$

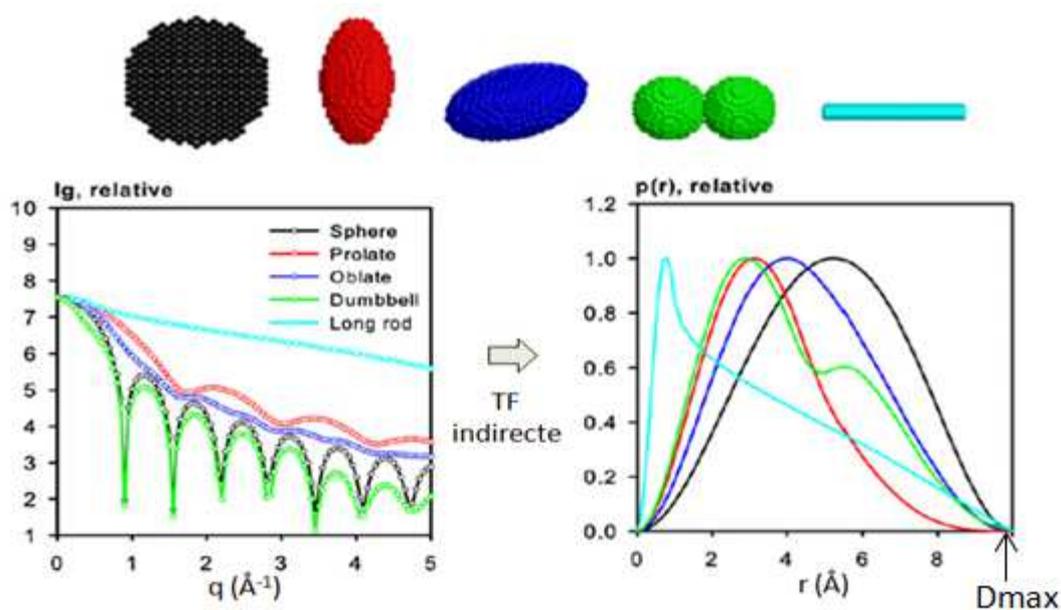
Le programme GNOM (Svergun 1992) est utilisé pour calculer la fonction de distribution des distances interatomiques  $P(r)$ . Les valeurs  $I(0)$  et  $Rg$  peuvent être calculées par la fonction  $P(r)$  selon :

$$I(0) = 4\pi \int_0^{D_{\text{max}}} p(r) dr \text{ (équation 7)}$$

$$Rg^2 = \frac{\int_0^{D_{\text{max}}} r^2 p(r) dr}{2 \int_0^{D_{\text{max}}} r^2 p(r) dr} \text{ (équation 8)}$$

La détermination des paramètres  $I(0)$  et  $Rg$  utilise l'ensemble de la courbe de diffusion et non pas seulement les intensités aux petits angles comme dans la méthode de Guinier. Elle est donc moins sensible aux interactions entre les particules ou à la présence d'oligomères en solution.

L'allure de la fonction  $p(r)$  donne également une indication sur la forme globale de la particule. Les molécules globulaires ont une  $p(r)$  en forme de cloche, dont le maximum correspond à  $D_{\text{max}}/2$  ; les particules allongées ont un maximum de  $p(r)$  pour  $r < D_{\text{max}}/2$  ; la fonction  $p(r)$  des particules aplaties présente un maximum élargi pour  $r < D_{\text{max}}/2$  tandis que celle de particules trouées présente un maximum pour  $r > D_{\text{max}}/2$  (Figure 49).



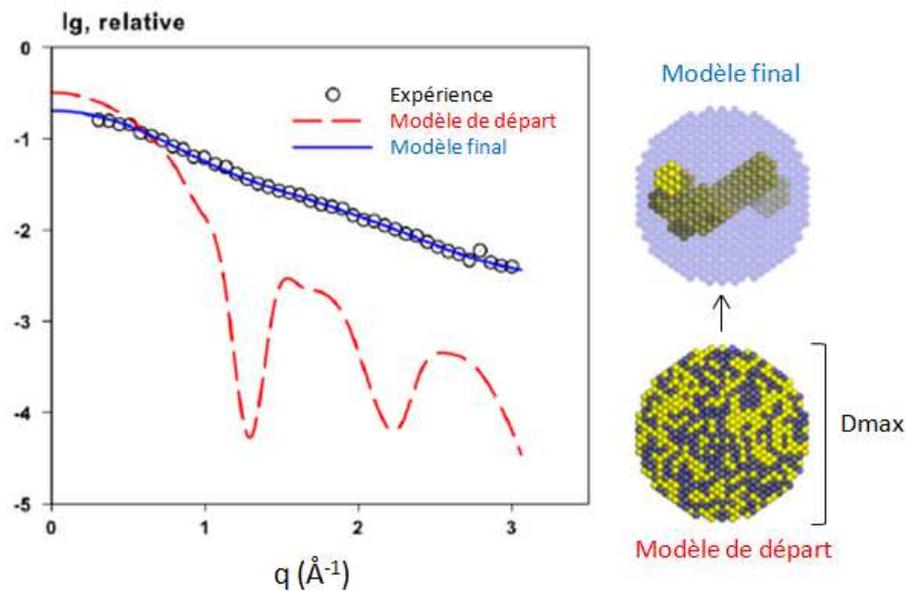
**Figure 49: Différentes formes géométriques définies par l'aspect de la fonction de distribution  $P(r)$  calculée par une transformée de Fourier indirecte de la courbe de diffusion.** (D'après (Mertens and Svergun, 2010)).

### 6.5.3. Modélisation ab initio

La fonction de distribution des distances qui détermine la forme générale de la protéine ainsi que sa taille maximale ( $D_{\max}$ ) permet de définir un modèle de départ. Pour calculer un modèle qui peut correspondre à l'enveloppe de la protéine, le modèle de départ est rempli de petites sphères (pseudo-atomes) qui sont attribuées soit à la protéine, soit au solvant. Par recuit simulé, la fonction d'accord est minimisée entre les courbes théorique et expérimentale. La fonction d'accord est de la forme  $f(x) = \chi^2 + \alpha P(x)$  où  $\chi^2$  mesure l'écart entre les données expérimentales et calculées,  $P(x)$  correspond à un terme de pénalité (compacité et connectivité de la molécule) et  $\alpha$  est le poids de la pénalité. Cette procédure réalisée par le programme DAMMIN (Svergun 1999) permet de déterminer une configuration compacte de pseudo-atomes qui correspondent à la courbe de diffusion (Figure 50).

Plus de dix modèles sont calculés indépendamment les uns des autres par DAMMIF. Afin de générer le modèle le plus probable, les programmes SUPERCOMB/DAMAVAR (Volkov et al. 2003) vont aligner les différents modèles deux à deux pour déterminer une pseudo-distance NSD (Normalized Spatial Discrepancy) entre les différents modèles et un modèle pris comme référence. Les modèles de faible NSD (inférieur à 0.7) par rapport au modèle de référence sont superposés pour faire apparaître les caractéristiques les plus représentées dans tous les

modèles. L'enveloppe ainsi générée et qui contient les modèles superposés va servir comme volume initial pour un ultime calcul par DAMMINF.



**Figure 50: Modélisation *ab initio* utilisant DAMMIN.** Les petites sphères en bleu du modèle de départ sont assignées au solvant et en jaune à la protéine. (D'après (Mertens and Svergun, 2010)).

L'existence d'une structure atomique similaire à la protéine étudiée, permet de vérifier grâce au programme CRY SOL l'exactitude de la reconstruction de l'enveloppe. Ce programme va d'abord calculer la courbe de diffusion théorique correspondant à la structure cristallographique. Puis il va la comparer avec la courbe de diffusion déterminée par les mesures de SAXS. Si les courbes ne se superposent pas, la présence d'une flexibilité multi-domaine est envisageable. L'utilisation de la représentation Kratky ( $I(q).q^2$  en fonction de  $q$ ) est utilisée pour identifier l'état de désordre d'une protéine et faire la distinction avec une particule globulaire. En effet la représentation de Kratky amplifie les particularités d'un profil de diffraction ce qui permet d'identifier le degré de compacité d'une protéine.

#### 6.5.4. Etude de la flexibilité par la méthode EOM (Ensemble Optimization Method)

La méthode EOM récemment développée par Bernadó fournit une approche utile pour évaluer la flexibilité entre des domaines d'une protéine mais uniquement si la structure à haute résolution des domaines est disponible (Bernadó et al., 2007). Cette méthode nécessite deux étapes : (1) La génération de 1000 conformations de manière aléatoire mais qui couvrent l'espace configurationnel de la protéine étudiée définie expérimentalement. Cet ensemble de

conformations est produit à partir des structures cristallographiques des domaines et de la séquence de la protéine par le programme Pre\_BUNCH. (2) Puis l'utilisation d'un algorithme appelé génétique permet par différentes méthodes d'optimisation la sélection d'un sous ensemble de conformations  $N$  qui décrivent collectivement la courbe de diffusion expérimentale. L'algorithme calcule pour  $N$  conformations choisies, la distribution de  $R_g$  et la valeur de NSD correspondant. Le fonctionnement de l'algorithme génétique est expliqué en détail dans la revue (Bernadó and Svergun, 2012).

#### **6.5.5. Collecte des données pour le complexe CupB6/B4**

Les mesures ont été réalisées sur la ligne de lumière ID29 de l'ESRF, Grenoble. Trois concentrations de CupB6/B4 ont été testées (3,5- 1,5 et 0,5  $\text{mg}\cdot\text{ml}^{-1}$ ) dans 20 mM MES pH 6- 100 mM NaCl- 50 mM d'arginine et de lysine et 5 % de glycérol. 10 mesures successives de 10 sec de temps d'exposition ont été enregistrées pour chaque échantillon, avec un flux continu de solvant permettant l'homogénéité des mesures. Une mesure du tampon seul a été réalisée avant et après chaque mesure de protéine afin d'éliminer la diffusion de base due au solvant et au capillaire. Une solution de BSA à concentration déterminée a été mesurée comme référence et a permis la calibration des mesures. Le détecteur utilisé était un Pilatus 1M (Dectris). Toutes les expériences ont été menées à 20°C.

---

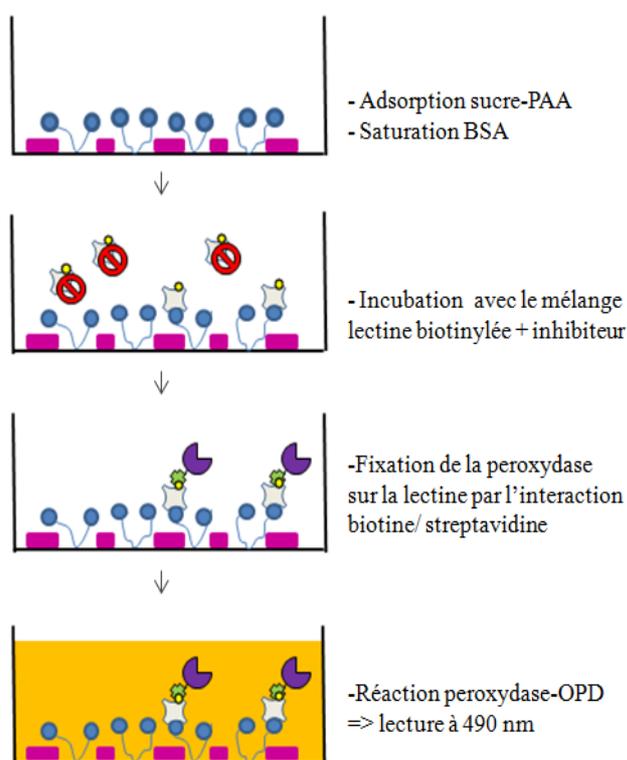
## 7. Outils pour l'étude des interactions protéines-sucre

---

### 7.1. Enzyme-Linked Lectin Assay : ELLA

#### 7.1.1. Principe du test ELLA

La mesure utilisant une lectine liée à une enzyme (ELLA) est basée sur le test ELISA (Enzyme Linked Immunosorbent Assay). Cette technique est utilisée pour cribler par compétition les meilleurs ligands saccharidiques d'une lectine. Elle se réalise en quatre étapes (Figure 51).



**Figure 51: Principe du test ELLA en compétition.**

Par adsorption, un polymère glycosylé (sucre-PAA) est immobilisé afin de recouvrir le fond des plaques 96 puits. Après une étape de saturation par BSA (Bovine Serum Albumin), le ligand étudié est dilué en cascade et la lectine biotinylée est ajoutée. La concentration de la lectine à utiliser est définie par une courbe d'étalonnage au préalable. Par des étapes d'incubation et de lavage la lectine biotinylée restante va être liée à un conjugué streptavidine-peroxydase. L'étape de révélation utilise le substrat OPD (*ortho*-phénylènediamine). Sa conversion en 2,3-diaminophénazine par la peroxydase donne une coloration orange mesurée à 490 nm.

L'intensité de cette coloration est proportionnelle à la quantité de lectine liée au glycopolymère et donc à la capacité du ligand testé à inhiber l'adhésion de la lectine. Les valeurs d'absorbance mesurées pour chaque concentration d'un ligand sont exprimées en pourcentage d'inhibition grâce aux valeurs de références mesurées (0 % avec le contrôle sans lectine et 100 % avec le contrôle sans inhibiteurs). Une courbe d'inhibition est alors générée

ce qui permet d'en extraire une valeur d'IC<sub>50</sub> qui indique la concentration en inhibiteur nécessaire pour provoquer 50 % d'inhibition.

### 7.1.2. Test ELLA avec la lectine PA-IL

L' $\alpha$ -Galactose-polyacrylamide à 5  $\mu\text{g}\cdot\text{ml}^{-1}$  (Lectinity) dans du tampon carbonate 50 mM pH 9.6 est fixé sur le fond des plaques 96 puits (Nunc Maxisorb). Après une étape de saturation au PBS-BSA 3 %, les inhibiteurs sont déposés dans le premier puits à 2  $\text{mg}\cdot\text{ml}^{-1}$  puis dilués en cascade au 1/3 dans le tampon PBS-BSA 0,3 %. La lectine PA-IL biotinylée (Sigma-Aldrich) à 1,5  $\mu\text{g}\cdot\text{ml}^{-1}$  est ajoutée. Après une incubation et des étapes de lavage, le substrat OPD (*ortho*-phénylènediamine) (Sigma-Aldrich) est ajouté pour l'étape de révélation. La réaction est stoppée par 30 % d'acide sulfurique et la mesure est faite à 490 nm sur le lecteur de plaques de microtitration (Bio-Rad ; model 680).

## 7.2. Microcalorimétrie de titration isotherme (ITC)

L'ITC permet de mesurer l'interaction entre deux partenaires qui sont libres en solution. Les paramètres tels que la constante d'affinité  $K_a$ , la stœchiométrie de l'interaction  $n$ , puis les contributions enthalpique  $\Delta H$  et entropique  $T\Delta S$  sont définies durant une mesure d'ITC. La gamme d'affinité mesurée par l'ITC s'étend de  $10^3$  à  $10^8 \text{ M}^{-1}$ .

### 7.2.1. Notion de thermodynamique

L'interaction non covalente entre une macromolécule  $M$  et son ligand  $L$  conduit à la formation réversible d'un complexe  $ML$ . La force de l'interaction est représentée par la constante d'association  $K_a$  en  $\text{M}^{-1}$ . Plus cette constante est grande, plus l'affinité entre les deux partenaires est forte. Le  $K_a$  correspond au rapport de la concentration du complexe  $ML$  sur celles de la macromolécule  $M$  et du ligand  $L$ . L'affinité est souvent exprimée par la constante de dissociation  $K_d$  en  $M$ .

$$K_a = [ML]/[M][L] = 1/K_d \text{ (équation 9)}$$

D'un point de vue thermodynamique, la constante d'affinité  $K_a$  est fonction de l'énergie libre de l'interaction. La variation d'énergie libre  $\Delta G$  (énergie de Gibbs) reliée à l'affinité de l'interaction s'écrit de la manière suivante :

$$\Delta G = -RT \ln K_a \text{ (équation 10)}$$

avec  $T$  la température absolue en Kelvins ( $^{\circ}\text{K}$ ),  $R$  la constante des gaz parfaits ( $R=8,314 \text{ J}\cdot\text{mol}^{-1}\cdot^{\circ}\text{K}^{-1}$ ) et  $K_a$  la constante d'affinité en  $\text{M}^{-1}$ .

Lorsque  $\Delta G$  est négatif, la réaction est spontanée et susceptible de fournir de l'énergie. Elle est également dite exergonique. Cette énergie est la résultante de deux contributions énergétiques : l'enthalpie  $\Delta H$  et l'entropie  $\Delta S$ .

$$\Delta G = \Delta H - T\Delta S \text{ (équation 11)}$$

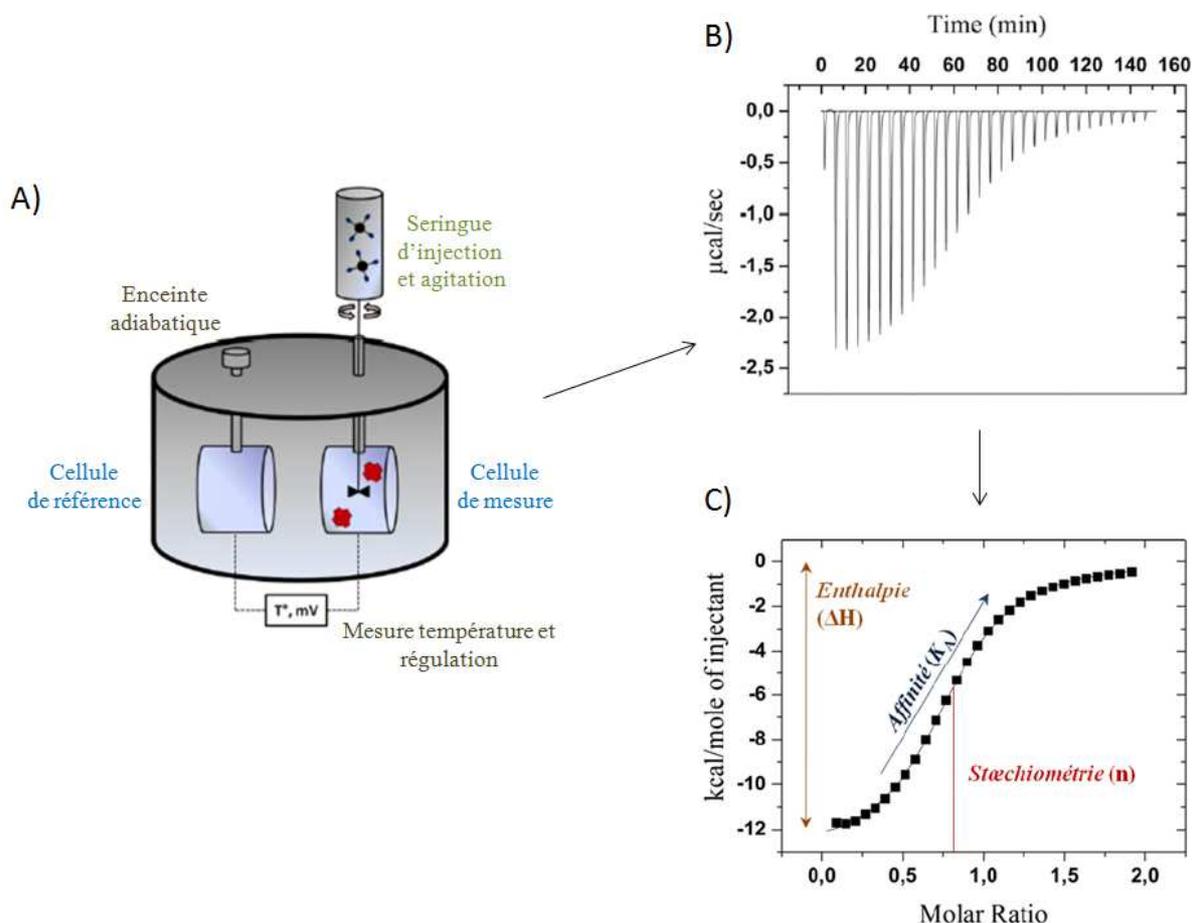
L'enthalpie  $\Delta H$  reflète l'énergie échangée par un système. Comme tout système a tendance à avoir l'énergie la plus basse possible,  $\Delta H$  tend à être négatif. L'entropie représente le degré de désordre d'un système. Son augmentation est due aux mouvements Browniens et augmente avec la température.  $T\Delta S$  tend à être positif.

### 7.2.2. Principe de l'ITC

Le principe d'un microcalorimètre repose sur la régulation de la température entre deux cellules, la cellule de référence et la cellule de mesure, isolées de l'environnement externe par une enceinte adiabatique (Figure 52). La cellule de référence contient uniquement du tampon et la cellule de mesure contient un des deux partenaires de l'interaction. Les deux cellules sont toujours maintenues à la même température. Durant l'expérience, de multiples injections du ligand contenu dans la seringue sont réalisées à des intervalles de temps réguliers dans la cellule de mesure. L'interaction entre la lectine et le ligand peut émettre ou absorber de la chaleur selon que l'enthalpie de l'interaction est exothermique ou endothermique. Après chaque injection, la température dans la cellule de mesure est ramenée à la même température que celle de référence grâce à un signal électrique. La puissance de ce signal en  $\mu\text{cal}/\text{sec}$  est enregistrée pour chaque injection et correspond à la chaleur émise ou absorbée par l'interaction de la lectine avec le ligand injecté. Au cours des injections, l'évolution de la différence de puissance en fonction du temps permet d'obtenir le thermogramme où chaque pic correspond à une injection. Au début de la titration, tous les sites de la protéine sont libres et permettent au ligand de s'associer. Puis à la fin de la titration, tous les sites sont occupés et aucune chaleur d'association n'est observée.

L'intégration de ce thermogramme permet d'obtenir une courbe isotherme d'association où chaque point représente une injection. L'ajustement d'une courbe théorique adéquate sur la courbe expérimentale permet grâce au logiciel « Origin » (Microcal) de calculer la constante d'association, la stoechiométrie ainsi que l'énergie libre. L'enthalpie globale de l'interaction est extraite de l'amplitude de la sigmoïde tandis que la stoechiométrie  $n$  est déterminée par le

point d'inflexion et la constante d'association de la pente de la courbure (Figure 52). D'après les équations 10 et 11, l'énergie libre de l'interaction ainsi que l'entropie sont calculées à partir des données expérimentales de l'enthalpie, de la stœchiométrie et de l'affinité.



**Figure 52: Principe du microcalorimètre isotherme de titration.** A) Représentation de l'instrument microcalorimètre. B) Exemple d'un thermogramme expérimental où chaque pic correspond à la puissance du signal électrique produit par l'appareil pour ramener la température de la cellule de mesure à celle de référence. C) Courbe de titration de l'interaction isotherme. Cette courbe est générée après l'intégration des données et permet de déterminer différentes constantes thermodynamiques.

La forme générale de la courbe dépend des concentrations en macromolécule et en ligand, du volume et du nombre d'injections ainsi que du temps de repos entre chaque injection. L'allure sigmoïde est très importante pour la précision dans la détermination des paramètres thermodynamiques. Elle est définie par le paramètre de Wiseman. La valeur de  $c$  est calculée par l'équation :

$$c = n[M_0]K_a \quad (\text{équation 12})$$

avec  $c$  le paramètre de Wiseman,  $n[M_0]$  la concentration initiale de sites de liaison en solution et  $K_a$  la constante d'association en  $M^{-1}$ .

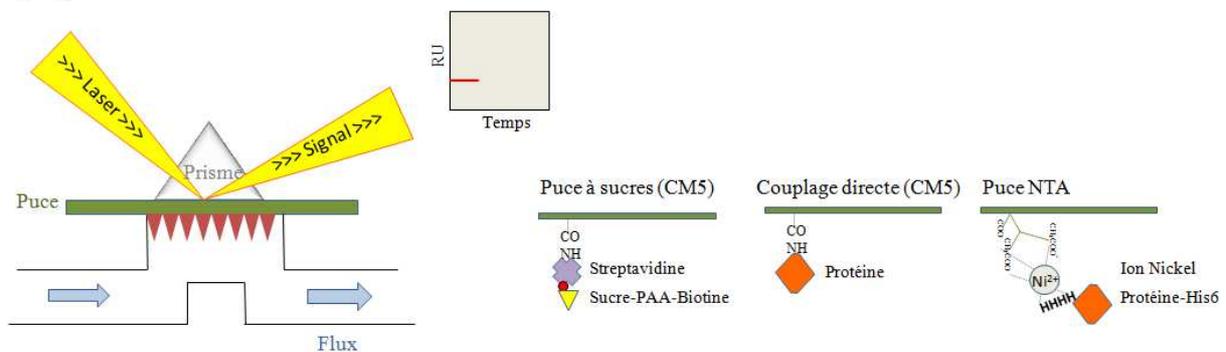
Le paramètre  $c$  doit être compris entre 10 et 100. Plus  $c$  est élevé, plus la pente de la sigmoïde est grande et moins le  $K_d$  est fiable. Lorsque la constante d'association et la stœchiométrie d'un système sont inconnues, la concentration en ligand dans la seringue est 10 fois plus élevée que celle de la macromolécule dans la cellule.

## 7.3. La résonance plasmonique de surface (SPR)

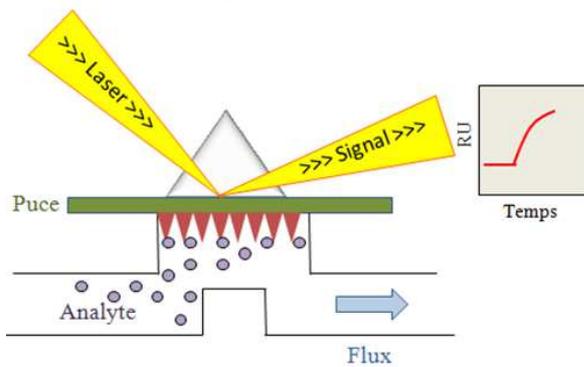
### 7.3.1. Le principe de la résonance plasmonique de surface

La technique de résonance plasmonique de surface (SPR) permet par la mesure en temps réel d'identifier une interaction entre deux partenaires ainsi que d'obtenir le  $K_d$  et les constantes de vitesse des réactions d'association/dissociation qui caractérisent l'interaction.

#### 1) Ligne de base



#### 2) Injection de l'analyte: Association



#### 3) Dissociation

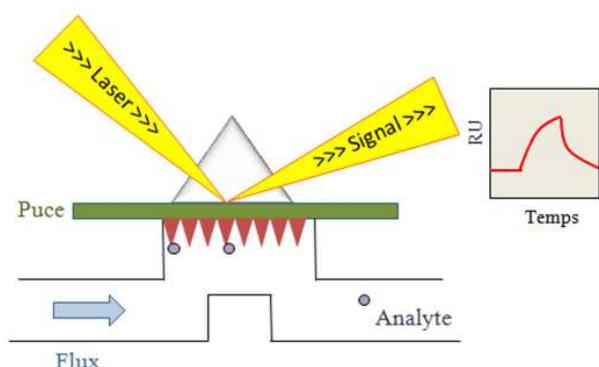


Figure 53: Principe du système de Résonance Plasmonique de surface.

Le dispositif comporte une puce soumise à un flux continu de tampon sur laquelle est fixé un des deux partenaires de l'interaction appelé le ligand (Figure 53). L'autre partenaire qui est l'analyte est injecté au contact de la puce. L'interaction est mesurée par le biais d'un faisceau laser qui se réfléchit en surface de la puce. L'angle de réflexion du laser change lorsque la surface est modifiée due à l'interaction de l'analyte avec le ligand. Ces variations sont

enregistrées sur un sensorgramme et quantifiées en unités de résonance (RU). 1000 RU correspondent à un déplacement de 0,1° de l'angle, ce qui équivaut à la fixation de 1ng d'analyte par mm<sup>2</sup>.

Un signal qui correspond à la ligne de base est enregistré tant que l'analyte n'est pas injecté ou s'il n'interagit pas avec le ligand. Pendant toute la durée de l'injection de l'analyte, l'interaction avec le ligand provoque une déviation progressive du laser. Tant que l'analyte s'accumule sur la puce une augmentation du signal est observée. En fin d'injection l'analyte se dissocie progressivement et le signal revient à sa valeur de départ.

Toutes les expériences ont été réalisées avec l'instrument Biacore X100 (GE Healthcare).

### 7.3.2. Etude de la spécificité de CupB6/B4 avec les puces CM5

Le tampon de course utilisé est composé de 20 mM HEPES pH 7,5 ; 150 mM NaCl, 0,005% Tween 20.

Le biocapteur CM5 (GE Healthcare) utilisé comporte deux canaux appelés FC1 et FC2 (pour Flow Cell 1 et 2) qui sont recouverts d'une matrice de carboxyméthyl-dextran. L'activation des groupements carboxyles par le mélange de EDC (1-éthyl-3-(3-diméthylaminopropyl) carbodiimide) et NHS (*N*-hydroxysuccinimide) (GE Healthcare) permet soit la fixation de la streptavidine (Sigma) pour les puces à sucres, soit la fixation de la protéine par un couplage amine. Pour les puces à sucres, la répétition de deux cycles « activation EDC/NHS » et « fixation de la streptavidine » dans le tampon de couplage permet la fixation en grande quantité de la streptavidine. Puis sur la cellule FC2, grâce à la forte interaction biotine-streptavidine, est fixé le ligand sucre-polyacrylamide-biotine (PAA-sucre-biotine) (Lectinity Holdings, Inc., Moscou). Aucun sucre n'est fixé sur la cellule de référence FC1.

Pour la fixation de la protéine de manière covalente sur la puce, deux cycles « activation EDC/NHS » et « fixation de la protéine » dans le tampon de couplage sont également réalisés sur la cellule FC2. Les quantités de ligand à immobiliser ( $RU_{\text{ligand}}$ ) sont estimées sur la base de la formule :

$$RU_{\text{ligand}} = 4.R_{\text{max}}.(MW_{\text{analyte}}/MW_{\text{ligand}}).valence_{\text{ligand}} \text{ (équation 13)}$$

Les groupements carboxyles libres sont désactivés par de l'éthanolamine-HCl (GE Healthcare) avant toutes injections de l'analyte. L'étape de régénération est adaptée en fonction de l'interaction ligand-analyte.

### 7.3.3. Etude de la spécificité de CupB6/B4 avec les puces NTA

Avant la capture des ions nickel sur la puce NTA, une première étape de régénération est réalisée avec 350 mM d'EDTA dans le tampon de course. Une étape de capture des ions nickel suivit d'une étape de lavage est réalisée avant l'injection de la protéine qui contient une étiquette hexahistidine. Les quantités de ligand à immobiliser ( $RU_{\text{ligand}}$ ) sont estimées sur la base de la formule (équation 13). La concentration de la protéine peut varier bien qu'une meilleure fixation et stabilisation est observée à faible concentration comme 0,2  $\mu\text{M}$ .

Après chaque injection d'analyte, deux étapes de régénération sont nécessaires. Tous les analytes testés ont été fournis par Elicityl. Le tampon de course utilisé est composé de 20 mM HEPES pH 7,5 ; 150 mM NaCl, 50 mM EDTA et 0,005 % Tween 20. Pour une meilleure fixation de la protéine, le pH peut varier de 7 à 8,5 ainsi que la concentration de NaCl.

## 7.4. La cristallographie aux rayons X

Pour résoudre la structure tridimensionnelle des macromolécules biologiques à l'échelle atomique, la cristallographie aux rayons X est particulièrement bien adaptée. Le rayonnement X dont la longueur d'onde est de l'ordre de l'Angström ( $10^{-10}$  m) est utilisé. L'interaction entre ces rayons et une molécule génère une diffusion de faible intensité. L'utilisation d'un cristal composé de molécules empilées de façon régulière et périodique dans l'espace permet d'augmenter cette intensité.

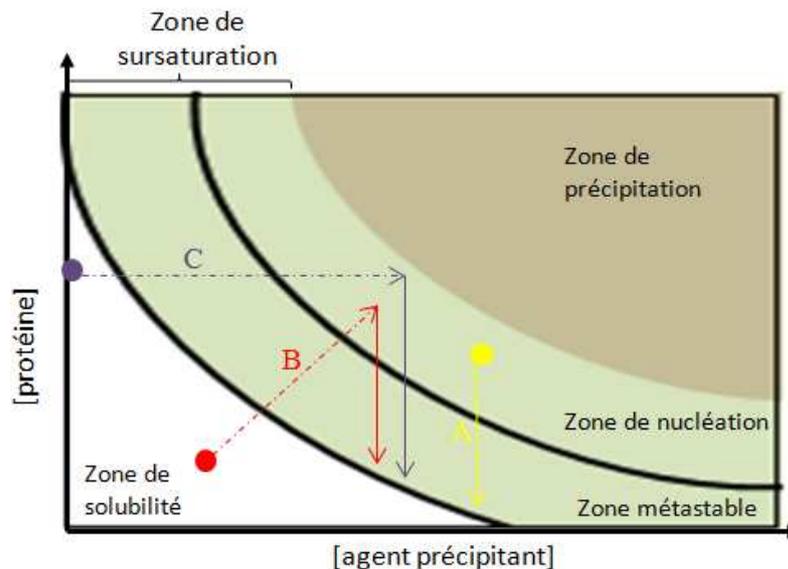
La résolution d'une structure tridimensionnelle par cristallographie nécessite trois étapes : (1) l'obtention d'un échantillon très pur et monodisperse, (2) l'obtention d'un monocristal qui diffracte et (3) la résolution du problème des phases.

### 7.4.1. La cristallogénèse

Un cristal est l'arrangement périodique et ordonné d'un motif se répétant à l'identique dans toutes les directions de l'espace. Le principe de la cristallisation est donc d'induire un changement d'état de la protéine de sa phase soluble et désordonnée à une phase solide et ordonnée, ces deux états étant en équilibre. Ce phénomène de transition de phase dépend de nombreux paramètres et peut se diviser en deux étapes : la germination ou plus communément appelé nucléation et la croissance cristalline. Pour induire la nucléation, la protéine est amenée dans un état soluble instable de sursaturation en diminuant lentement sa solubilité (Figure 54). Ce qui revient à destabiliser les interactions protéines-solvant au profit d'interactions protéines-protéines. Des interactions non-spécifiques conduisent à la

précipitation (phase solide amorphe). Au contraire des interactions spécifiques entre les molécules de protéines mènent à la formation de germes cristallins. Le retour à l'équilibre thermodynamique s'effectue alors plus ou moins rapidement par la croissance de la phase solide.

Le diagramme de phase décrit le comportement des macromolécules au cours du processus de cristallisation (Figure 54).



**Figure 54: Représentation schématique du diagramme de phase d'une protéine.** La courbe de solubilité définit la limite entre la phase soluble et solide. Sous l'influence de la concentration de la protéine et du précipitant, le système atteint la zone de nucléation. La sursaturation élevée conduit à la nucléation du cristal. Puis la diminution de la concentration de la protéine amène le système dans la zone métastable où la sursaturation est plus faible et permet la croissance du cristal. Représentation du processus de cristallisation avec la méthode a) en batch, b) par diffusion de vapeur, c) par microdialyse.

#### 7.4.1.1. Les paramètres physico-chimiques et les principaux agents précipitants

Les paramètres qui influencent la solubilité des protéines sont très nombreux et variables d'une protéine à l'autre. L'état de sursaturation qui conduit à la cristallisation, dépend de plusieurs facteurs tels que la pureté et la concentration de la protéine, les agents qui agissent sur la solubilité, la stabilité et/ou la conformation de la protéine, le pH dont dépend la répartition des charges à la surface de la protéine, la nature et la concentration des agents précipitants et des additifs, la température affectant la solubilité, la stabilité de la protéine et la cinétique de cristallisation.

L'agent précipitant tient un rôle majeur dans le processus de cristallisation puisqu'il affecte directement et de façon majoritaire la solubilité de la protéine. Les différents agents les plus

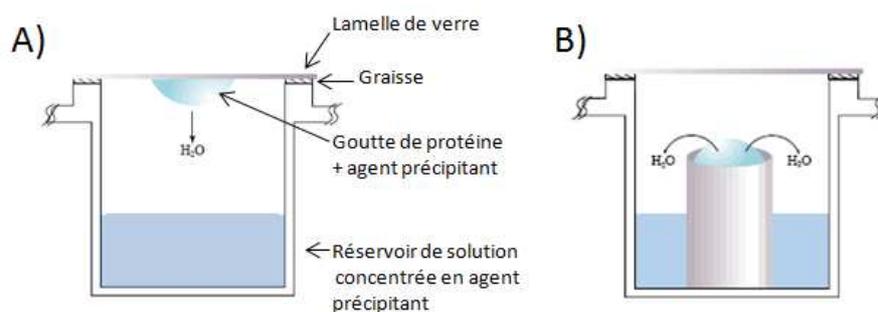
couramment employés peuvent être classés en trois familles distinctes suivant leurs effets sur les interactions protéines-solvant et protéines-protéines :

- les polymères qui agissent par exclusion du solvant au voisinage de la protéine,
- les sels non chaotropiques (non dénaturants pour les protéines) qui modifient la force ionique et augmentent les interactions hydrophobes en excluant le solvant,
- les solvants organiques qui diminuent la constante diélectrique du milieu.

Afin de maîtriser pleinement la solubilité de la protéine, il est souvent nécessaire de combiner les agents précipitants ou d'y associer un ou plusieurs additifs.

#### 7.4.1.2. La cristallisation par diffusion de vapeur

Il existe de différentes méthodes pour obtenir des cristaux de protéines. Chacune permet d'atteindre les zones de germination et de croissance cristalline mais d'une manière différente (Figure 54). La méthode la plus utilisée est la diffusion de vapeur qui peut être réalisée soit en goutte suspendue ou soit en goutte assise (Figure 55).



**Figure 55: Technique par diffusion de vapeur A) en goutte pendue et B) en goutte assise.**

Une goutte de protéine est mélangée avec une goutte du tampon situé dans le réservoir qui contient un agent précipitant. Le mélange peut se faire avec différents ratio protéine/ agent précipitant. La goutte est ensuite déposée soit sur une lamelle qui est retournée au dessus du réservoir soit posée sur un rebord proche du réservoir. La cellule est fermée hermétiquement afin de permettre un équilibre des concentrations entre la goutte et le réservoir par diffusion des espèces volatiles. La diminution du volume de la goutte entraîne une augmentation de la concentration de la protéine et de l'agent précipitant qui provoquent l'état de sursaturation.

#### 7.4.1.3. Optimisation

L'obtention de microcristaux ou de cristaux dont la diffraction n'est pas de bonne qualité nécessite une étape d'optimisation. Le but de cette étape est d'ajuster finement les conditions

afin de diminuer la germination, d'augmenter la taille des cristaux, d'améliorer leur aspect, d'éliminer les problèmes de macle ou de cristaux collés et surtout d'augmenter l'ordre cristallin. Différents paramètres peuvent être utilisés afin de limiter le taux de nucléation ou de ralentir la croissance cristalline:

- le scale-up : dans de nombreux cas des cristaux plus gros peuvent être obtenus, moyennant quelques réajustements des conditions, en augmentant simplement la quantité des produits utilisés (volumes des gouttes).

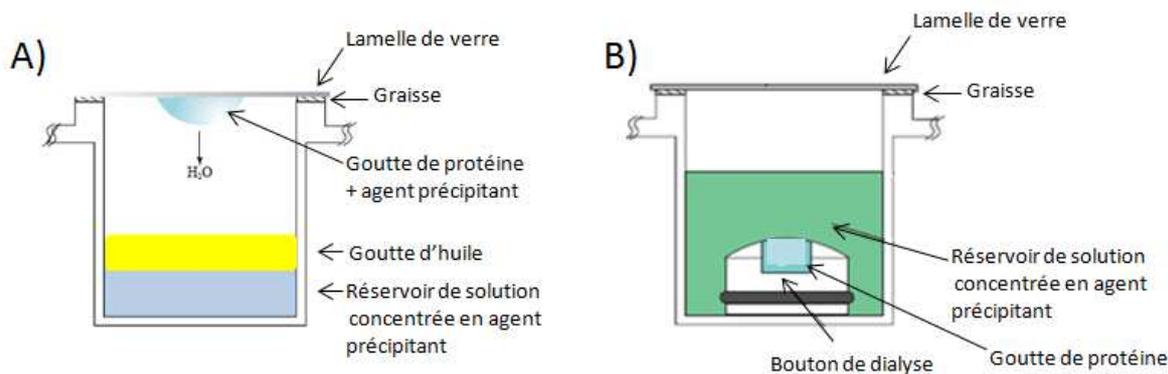
- la modification du rapport volumique entre la solution de protéine et le réservoir permet de modifier la vitesse de concentration et d'atteindre un niveau de sursaturation différent.

- la méthode de perturbation : cette approche consiste à modifier les conditions initiales de cristallisation afin de les déplacer vers des conditions plus favorables. Cela peut être réalisé en ajoutant des additifs chimiques de divers types. Ces additifs peuvent être des agents de solubilisation (les détergents non ioniques, le dioxane...), des sels chaotropiques (l'urée, le thiocyanate de potassium...), des sels métalliques ( $Zn^{2+}$ ,  $Ni^{2+}$ ...), des composés agissant sur les potentiels d'oxydoréduction (DTT, ferricyanure...).

- l'ensemencement : cette méthode consiste à utiliser les germes obtenus en forte sursaturation pour ensemer en faible sursaturation. Les germes peuvent être soit des monocristaux (macroensemencement) soit un broyat de cristaux (micro-ensemencement). En permettant de maîtriser le nombre de germes ensemencés, cette méthode permet de favoriser la croissance de quelques grands cristaux au lieu d'une multitude de petits.

- la cristallisation en gel : la cristallisation dans les gels de silice ou d'agarose semble efficace pour, à la fois, réduire le taux de germination et augmenter la taille et la qualité de diffraction des cristaux. Les gels permettent d'éviter les effets de convection qui perturbent la formation des cristaux.

- le contrôle de la diffusion de vapeur : différentes techniques permettent de modifier les échanges de vapeur entre la goutte et le réservoir. L'une consiste à placer une couche d'huile à la surface du réservoir créant ainsi une barrière (Figure 56). Les vapeurs d'eau diffusent plus lentement ce qui permet d'agir directement sur la cinétique de croissance des cristaux. Cette technique permet de réduire le nombre de cristaux en croissance et d'augmenter leur taille. Une autre technique qui permet d'augmenter la taille des cristaux est la microdialyse. La protéine est placée dans un bouton de dialyse qui est par la suite immergé dans un réservoir contenant une solution concentrée en agent précipitant (Figure 56).



**Figure 56: Techniques utilisées pour augmenter la taille des cristaux, en A) ajoutant une goutte d'huile entre la goutte et le réservoir ou B) en utilisant la technique de microdialyse.**

#### 7.4.1.4. Collecte des données expérimentales

La diffraction des cristaux est testée au synchrotron de l'ESRF (European Synchrotron Radiation Facility, Grenoble). Lors de la collecte, le rayonnement peut provoquer la formation de radicaux libres à l'intérieur du cristal et endommager les molécules. Mais l'intensité du faisceau peut également provoquer un échauffement local du cristal provoquant sa destruction complète. Afin d'éviter ces dommages, les cristaux sont congelés dans l'azote liquide.

Les cristaux des macromolécules biologiques contiennent beaucoup d'eau (de 30 à 80 %), la congélation dans l'azote liquide peut entraîner la formation de glace cristalline. Cette glace peut endommager ou détruire le cristal. Lorsque l'eau autour du cristal cristallise, l'apparition de ce qui est appelé communément des anneaux de glace perturbent les clichés de diffraction. Pour éviter cela, une méthode consiste à remplacer une partie de l'eau par un cryoprotectant selon une technique de trempage.

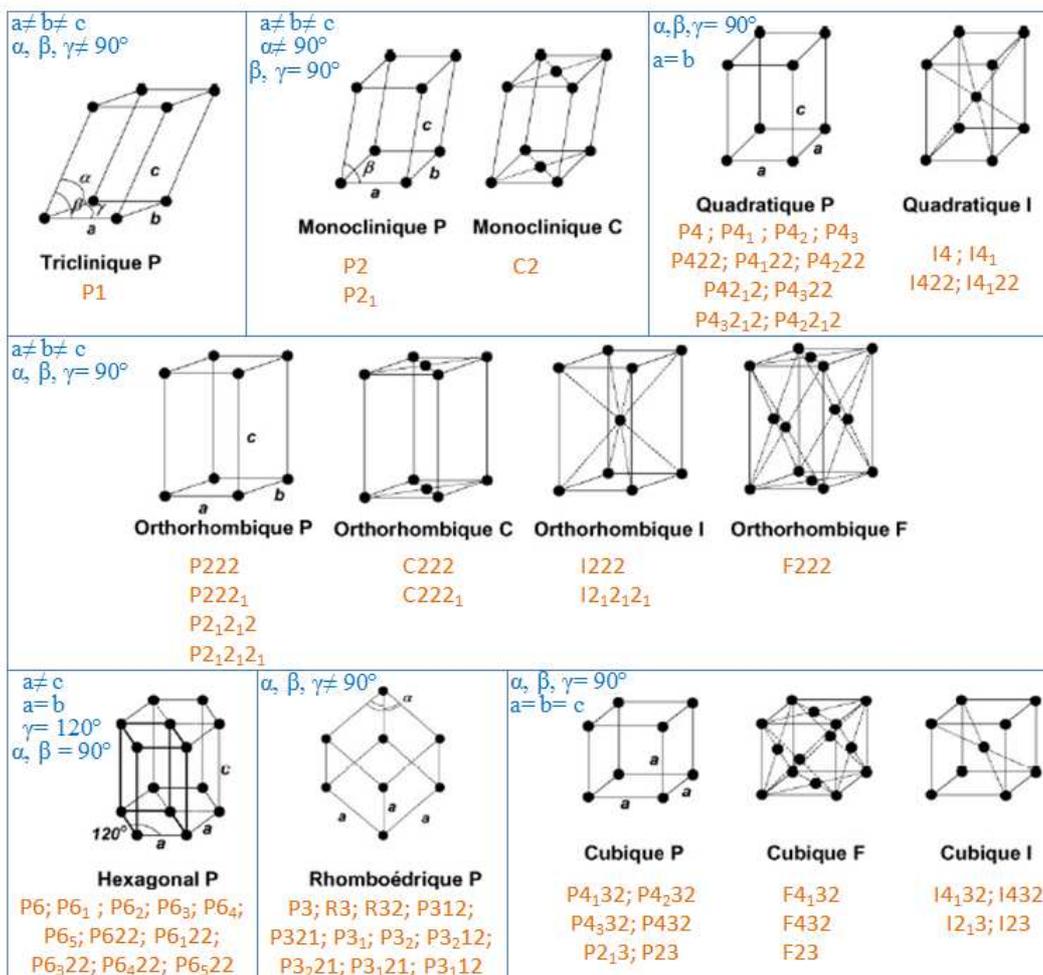
La collecte de données est effectuée sous un flux d'azote continu. Pendant la collecte, le cristal tourne par paliers de 0 à plusieurs degrés autour d'un axe perpendiculaire au faisceau de rayons X. Les rayons diffractés sont collectés par des détecteurs de type CCD (couple charge device) et ceci génère des clichés de diffraction. De nos jours, l'enregistrement d'un jeu de données complet est réalisé avec des temps d'acquisition très court de l'ordre de quelques minutes sur certaines lignes de lumières.

## 7.4.2. La cristallographie

### 7.4.2.1. Définition d'un cristal

Un cristal est la résultante de la translation dans les trois directions de l'espace de la maille élémentaire. Cette entité est le plus petit volume qui contient toutes les informations du cristal et est définie par trois vecteurs qui génèrent les six paramètres de la maille tels que les trois longueurs des vecteurs  $a$ ,  $b$ ,  $c$ , les trois angles interaxiaux  $\alpha$ ,  $\beta$  et  $\gamma$ .

La maille élémentaire est délimitée par un ensemble de points appelés nœuds du réseau. L'ensemble de la maille est engendré par l'unité asymétrique à laquelle sont appliqués les opérateurs de symétrie du groupe d'espace (Figure 57). Elle peut contenir plusieurs molécules qui sont alors reliées entre elles par des opérateurs de rotation-translation définissant une symétrie non cristallographique (SNC).



**Figure 57: Systèmes cristallins et groupes d'espaces.** Il existe 14 types différents de mailles élémentaires appelés les 14 réseaux de Bravais. Pour les cristaux de macromolécules biologiques, il existe 64 types de symétries distinctes appelées groupes d'espaces.

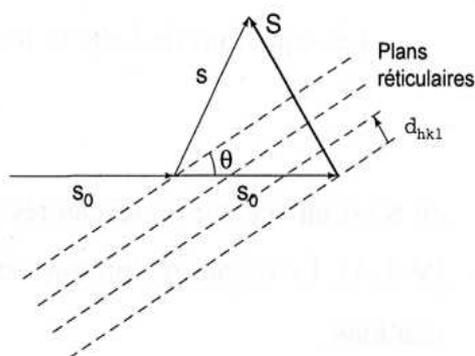
### 7.4.2.2. Principe de la diffraction aux rayons X

La diffraction met en évidence le caractère ondulatoire de la lumière. Cela s'explique par les interférences destructives ou constructives d'ondes cohérentes diffusées par les atomes d'un réseau lorsqu'ils sont exposés aux rayons X. L'arrangement périodique d'un cristal fait qu'un faisceau incident peut produire plusieurs faisceaux diffractés, appelés réflexions, ayant chacune une direction et une intensité déterminée. L'ensemble des directions de diffraction dépend uniquement de l'orientation des trois vecteurs de base  $a$ ,  $b$  et  $c$  qui définissent le réseau cristallin, par rapport au faisceau incident. L'expérience de diffraction peut être décrite par les équations de Bragg.

D'après la loi de Bragg, la diffraction des rayons X est assimilée à une réflexion du faisceau incident sur une famille de plans réticulaires parallèles (perpendiculaires à  $S$ ) d'indice  $h, k, l$  et de distance interréticulaire  $d_{hkl}$ , se produisant aux angles  $2\theta$  (angle entre les faisceaux incidents et diffractés, donc entre  $s_0$  et  $s$ ) (Figure 58) vérifiant l'équation :

$$2 d_{hkl} \sin \theta = n\lambda \text{ (avec } n \text{ un nombre entier)}$$

Avec  $\lambda$  la longueur d'onde du faisceau incident,  $\theta$  l'angle sous lequel le rayon traverse le cristal et  $d_{hkl}$  la distance inter-planaire entre les plans cristallins parallèles produisant un rayon diffracté.



**Figure 58: Schéma décrivant le phénomène de la diffraction selon la loi de Bragg.**

Si cette équation est respectée, alors les ondes vont s'ajouter et former une tache de diffraction également appelée réflexion. La position de chaque tache observée sur le cliché de diffraction est reliée à une famille de plans réticulaires définie par les indices de Miller ( $h, k, l$ ). Le cliché de diffraction correspond à l'espace réciproque par opposition à l'espace réel du cristal.

### 7.4.2.3. Indexation, intégration et mise à l'échelle du jeu de données

Le protocole de traitement des clichés de diffraction enregistrés lors de la rotation du cristal comporte quatre étapes essentielles : l'indexation, l'intégration des intensités, la mise à l'échelle et la moyennation.

L'indexation consiste, à l'aide de la position des taches de diffraction couplée à l'angle d'oscillation, à déterminer l'orientation du cristal ainsi que les paramètres de maille. Les programmes d'indexation cherchent trois vecteurs de base permettant de relier toutes les réflexions, puis la matrice d'orientation ainsi que la maille qui expliquent toutes les réflexions observées. Il est alors possible de prédire la position de l'ensemble des taches de diffraction sur les images et d'en mesurer l'intensité.

L'intégration des intensités consiste à établir des profils moyens des taches et de calculer l'intensité de chaque réflexion par intégration numérique suivant les profils calculés. Cette étape permet de corriger l'étalement des taches sur un diamètre de quelques millimètres mais également de corriger les effets d'absorption et de polarisation du faisceau incident. L'ensemble des images intégrées sont ensuite mises à l'échelle entre autre afin de mettre au même niveau l'intensité des différentes réflexions collectées sur différentes images de diffraction. Les réflexions équivalentes sont comparées pour mettre en évidence la présence de réflexions mal mesurées (hors-la-loi: outliers).

Finalement, une intensité moyenne est calculée pour chaque HKL unique et le module des facteurs de structure est calculé. C'est la réduction des données qui permet d'obtenir l'ensemble des réflexions uniques nécessaires pour calculer l'image de la densité électronique par transformée de Fourier.

La qualité d'un jeu de données est déterminée par différents critères statistiques calculés au cours du traitement. Les paramètres les plus significatifs et utilisés sont :

- la résolution : qui déterminera la finesse des détails dans l'image de la densité électronique.
- le rapport signal sur bruit ( $I/\sigma$ ) moyen : plus il est grand et meilleure est l'estimation des facteurs de structure.
- la complétude : qui mesure le rapport entre le nombre de réflexions uniques mesurées et le nombre théorique de ces réflexions pour la résolution des données. Elle doit être la plus grande possible et supérieure à 90 %.
- la multiplicité: qui correspond au nombre de fois où l'intensité d'une réflexion unique est mesurée. Ce qui est équivalent au nombre moyen de réflexions symétriques mesurées. Plus la

valeur numérique de ce critère est grande, meilleure sera l'estimation de l'intensité moyenne d'une réflexion unique.

- le facteur  $R_{sym}$  : (ou  $R_{merge}$ ) qui représente le désaccord entre les intensités des réflexions équivalentes par symétrie. Plus ce facteur est bas, meilleure est la cohérence du jeu de données.

#### 7.4.2.4. Facteurs de structure et densité électronique

Chaque tache observée correspond à une onde diffusée par une maille élémentaire du cristal qui est décrite par une amplitude et une phase reliées au facteur de structure  $F_{hkl}$  défini par la relation :

$$F_{hkl} = \iiint \rho(x,y,z) e^{2\pi i(hx+ky+lz)} dV \quad (\text{équation 14})$$

Le facteur de structure  $F_{hkl}$  correspond à la transformée de Fourier de la densité électronique dans le cristal.

En appliquant la transformée de Fourier, on peut avoir accès à la densité électronique ( $\rho$ ) pour un point de coordonnées  $(x, y, z)$  dans le volume de la maille  $V$  comme la transformée de Fourier inverse du facteur de structure  $F_{hkl}$  :

$$\rho(x,y,z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx+ky+lz)} \quad (\text{équation 15})$$

Lors d'une expérience de diffraction, les clichés obtenus permettent de déterminer l'intensité de l'onde diffractée dans une direction  $h k l$ , qui est proportionnelle au carré du module du facteur de structure. En revanche l'information concernant la phase n'est pas connue. On ne peut donc pas calculer directement la densité électronique à partir des clichés de diffraction. Il existe plusieurs méthodes (MIR (Multiwavelength anomalous diffraction), SAD (Single anomalous diffraction), MAD (Multiwavelength anomalous diffraction) et le remplacement moléculaire) permettant d'obtenir les phases nécessaires au calcul de la carte de densité électronique.

#### 7.4.2.5. Le remplacement moléculaire

On utilise la méthode de phasage par remplacement moléculaire lorsque l'on dispose d'une structure cristallographique d'une protéine possédant typiquement plus de 30 % d'identité de séquence en acides aminés avec notre protéine d'intérêt.

Le positionnement d'une ou plusieurs molécules modèles dans la maille du cristal étudié permet de pouvoir calculer une première estimation des phases pour les facteurs de structure.

La méthode consiste à calculer les fonctions de Patterson représentant l'ensemble des vecteurs interatomiques d'une molécule, pour le modèle cristallographique et pour la structure recherchée. Elle est indépendante des phases.

La première étape de la méthode du remplacement moléculaire est la recherche de l'orientation du modèle initial. Cette étape consiste à déterminer les trois angles d'Euler ( $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ ) pour lesquels la fonction de corrélation entre la fonction de Patterson du modèle orienté et celle issue du jeu de vecteurs observés est maximale.

Durant la seconde étape du remplacement moléculaire, le modèle correctement orienté doit être positionné en translation dans la nouvelle maille pour donner le meilleur accord entre les facteurs de structure calculés et mesurés. Pour ce faire, les vecteurs entre molécules symétriques de la fonction de Patterson sont utilisés. Le programme utilisé dans ce travail est PHASER (CCP4) (McCoy et al., 2007).

#### 7.4.2.6. Affinement des structures

Grâce aux phases obtenues, une première carte de densité électronique est calculée et visualisée à l'aide du logiciel graphique COOT (Emsley and Cowtan, 2004). Cette première carte contient des erreurs et n'est pas toujours de qualité suffisamment bonne pour déterminer la position exacte des atomes. La procédure d'affinement permet de corriger le modèle initial afin d'obtenir un accord optimal entre les modules des facteurs de structure observés et ceux calculés à partir du modèle. En pratique, on procède en alternant des étapes de reconstruction manuelle du modèle initial avec des étapes d'affinement numérique des paramètres du modèle (position et agitation de chaque atome).

Les étapes de reconstruction manuelle du modèle sont effectuées sur COOT. Les cartes de densité utilisées sont essentiellement des cartes dites différences ou cartes résiduelles  $F_o - F_c$ . Elles illustrent le désaccord entre le modèle et les facteurs de structure observés. Un pic positif s'interprète comme des atomes manquant dans le modèle. Un pic négatif sur un atome du modèle indique que celui-ci est sans doute mal placé. A la fin de l'affinement, cette carte résiduelle doit contenir le moins possible de pics négatifs ou positifs. La carte  $2F_o - F_c$  représente une combinaison linéaire de la carte de densité électronique et de la carte résiduelle. Elle permet de visualiser la densité électronique tout en tenant compte des imperfections du modèle.

Les étapes d'affinement numériques sont réalisées avec le programme REFMAC (CCP4). Ce programme permet d'affiner les structures par des méthodes de minimisation des coordonnées atomiques du modèle en tenant compte des contraintes structurales (liaisons atomiques, angles

de liaison...) et en calculant les facteurs de structure en module et en phase puis en les mettant à l'échelle par rapport aux facteurs de structure observés. L'affinement permet donc de maximiser une fonction de vraisemblance contenant les paramètres à affiner décrits par des distributions de probabilité.

Durant les différentes étapes d'affinement, la correspondance entre le modèle et les données expérimentales est mesurée par le facteur d'accord cristallographique  $R$ :

$$R = \frac{\sum_{hkl} ||F_o(hkl)| - |F_c(hkl)||}{\sum_{hkl} |F_o(hkl)|} \quad (\text{équation 16})$$

Le facteur  $R$  fait partie de la fonction à minimiser. Plus sa valeur est basse, plus le modèle est en accord avec les données. Ce facteur n'étant pas complètement objectif, un second facteur est utilisé, le facteur  $R_{free}$ . Il est calculé à partir de 5 % de réflexions qui ne sont pas affinées. Une diminution du facteur  $R$  sans diminution corrélée du facteur  $R_{free}$  indique un affinement inefficace qui consiste à faire entrer artificiellement dans le modèle le bruit des cartes de densités électroniques.

En fin d'affinement, les paramètres structuraux du modèle final sont vérifiés et comparés aux valeurs de structures protéiques avec le logiciel PROCHECK (Laskowski et al., 1996). La structure est ensuite déposée dans la base de données internationale RCSB Protein Data Bank.

# **Résultats et discussion**



---

## 8. Mise en application des méthodes qui permettent d'étudier les interactions lectine-sucre

---

### 8.1. Introduction

Les lectines jouent un rôle important dans de multiples processus de reconnaissance moléculaire et de signalisation cellulaire. Chez les légumineuses, il existe cinq classes de lectines définies en fonction de leur spécificité : les lectines à fucose, les lectines à GlcNAc/ Glc, les lectines à glucose/ mannose, les lectines à Gal/ GalNAc, et enfin les lectines à NeuAc. Leur large spectre de ligands en fait un modèle de choix pour comprendre les interactions protéine/ sucres, mais elles sont également utilisées à des fins biotechnologiques. A ce titre, trois lectines ont été récemment identifiées comme ayant une activité anti-virale contre le virus human parainfluenza de type 2 (Uematsu et al., 2012).

Malgré une large variété de sites de reconnaissance aux carbohydrates (CRDs: Carbohydrate-Recognition Domain), les lectines de légumineuses présentent une forte similarité de séquence protéique (Sharon and Lis, 2004) ainsi que le même type de repliement (Loris et al., 1998). Les lectines de légumineuses affichent un poids moléculaire compris entre 25 et 30 kDa et présentent un repliement appelé « legume lectin fold » ou « jelly roll » caractérisé par un sandwich de brins  $\beta$  antiparallèles. Les lectines de légumineuses ont une structure quaternaire de type dimérique ou tétramérique. Leur CRD est constitué de cinq boucles (A, B, C, D, E) qui forment une cavité. La taille et la séquence de ces boucles varient entre les différentes lectines et définit ainsi leur spécificité (Young and Oomen, 1992). Deux acides aminés très conservés, Asp et Asn, sont cruciaux pour la reconnaissance du sucre. La présence de résidus hydrophobes (Phe, Tyr, Trp, Leu) vient stabiliser cette interaction par un empilement. La disposition spatiale identique de ces résidus est assurée par la coordination proche des deux cations  $\text{Ca}^{2+}$  et  $\text{Mn}^{2+}$  (Sharma and Surolia, 1997). La plupart des lectines de légumineuses ont été isolées et caractérisées dans les tribus *Phaseoleae* et *Vicieae* de la sous-famille des Légumineuses *Papilionoideae*. Au sein de la tribu des *Dalbergieae*, seule la structure atomique de la lectine glucose/mannose PAL de *Pterocarpus angolensis* a été résolue (Loris et al., 2003).

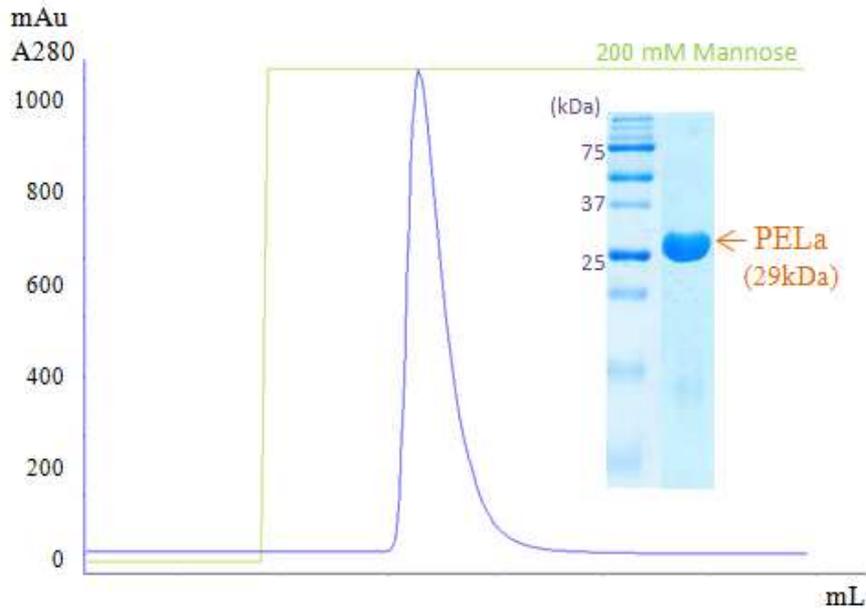
Durant sa thèse, le Dr Raquel Benevides, a identifié une lectine à mannose/ glucose présente dans les graines de *Platypodium elegans* Vogel qui appartient également à la tribu des

*Dalbergiaceae. Platypodium elegans* Vogel est une espèce naturelle (arbre) présente au Brésil. Des travaux préliminaires avaient permis de démontrer que des extraits de graines de *Platypodium elegans* étaient capables d'agglutiner les érythrocytes de lapins mais pas les érythrocytes humains. Le Dr Benavides a isolé un gène de lectine en se basant sur la présence de motifs conservés chez les légumineuses (motifs de fixation de métaux). Elle a ensuite déterminé sa séquence puis l'a cloné pour produire la protéine sous forme recombinante chez *E. coli*. La protéine ne s'exprimant que sous forme de corps d'inclusion, un protocole de renaturation a été mis au point. La lectine PELa ainsi obtenue sous forme soluble et pure a été marquée puis envoyée au « Consortium for Functional Glycomics » afin d'étudier sa spécificité par des puces à sucres.

Dans le cadre de cette étude, mon travail a consisté à exprimer et purifier PELa afin d'étudier par ITC les paramètres thermodynamiques qui caractérisent son interaction avec des ligands sélectionnés. Afin de mieux comprendre les interactions mises en place au sein du complexe PELa/ligand, des essais de cristallisation ont été entrepris. Les structures cristallographiques des complexes PELa/trimannoside et PELa/heptasaccharide ont été résolues par remplacement moléculaire. J'ai résolu et affiné la structure du complexe PELa/heptasaccharide. Ce travail a donné lieu à une publication, où je suis 2<sup>ème</sup> auteure, qui est jointe en annexe.

## **8.2. Production de PELa sous forme recombinante**

La protéine PELa est produite dans les corps d'inclusion dans la souche BL21 (DE3) d'*E. coli*. La renaturation des corps d'inclusions se fait par une étape de solubilisation avec 8 M d'urée suivie d'une renaturation de 1 mg.ml<sup>-1</sup> de PELa dans 50 mM Tris-HCl pH 7,5- 240 mM NaCl- 2 mM MnCl<sub>2</sub>- 2 mM de CaCl<sub>2</sub>- 10 mM KCl- 1 mM de DTT. Son affinité pour le mannose permet grâce à une chromatographie d'affinité sur D-mannose-Sepharose (Sigma-Aldrich) d'obtenir la protéine PELa sous forme pure et correctement repliée (Figure 59).



**Figure 59: Purification de PELa par chromatographie d'affinité sur colonne D-Mannose après renaturation des corps d'inclusions.** Le pic obtenu après élution est concentré puis analysé sur gel SDS-PAGE 15 %.

### 8.3. Etude de la spécificité de PELa

#### 8.3.1. Résultat de l'analyse par puces à sucres

Les résultats issus de l'analyse des puces à sucres développées par le « Consortium for Functional Glycomics » mettent en évidence que la lectine PELa possède une spécificité pour les *N*-glycanes et notamment les *N*-glycanes de type oligomannosiques qui comportent un  $\alpha$ -mannose comme sucre terminal non-réducteur (composés C et D). PELa présente cependant une nette préférence pour les *N*-glycanes de type complexe et asymétrique avec une courte branche 1-6 qui comporte un résidu  $\alpha$ -mannose en sucre non-réducteur et une branche 1-3 qui possède un LacNAc (*N*-acetyl-lactosamine) ou un LacNAc sialylé sur le résidu  $\alpha$ -mannose (composés E et F). La présence d'un GlcNAc sur la branche 1-6 diminue de manière significative l'affinité de la lectine (Figure 60).

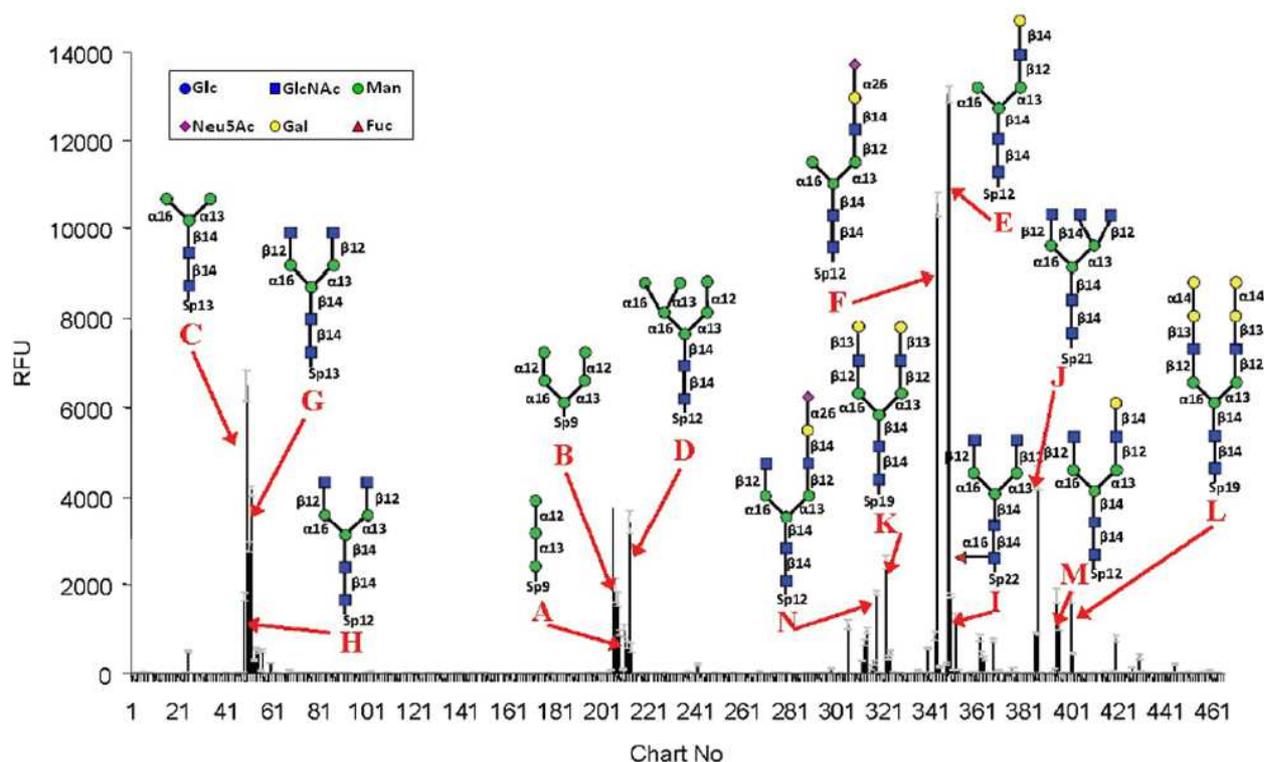


Figure 60: Résultats des puces à sucres (version 4.1) pour PELa à 10 µg.ml<sup>-1</sup>.

### 8.3.2. Etude de la spécificité de PELa par ITC

Une étude par microcalorimétrie isotherme de titration a été entreprise avec les glycopeptides correspondant aux composés C, H et E afin de vérifier les résultats obtenus par les puces à sucres et d'obtenir les paramètres thermodynamiques de l'interaction. Ces molécules présentant un Asn en position réductrice, ont été synthétisées par le Prof. Carlo Unverzagt (Bayreuth). D'après les thermogrammes obtenus (Figure 61), les pics montrent une réaction de type exothermique qui indique une réaction spontanée et susceptible de fournir de l'énergie. Ce type de réaction est caractéristique des interactions lectines-sucres.

Le composé C qui correspond au noyau d'un *N*-glycane présente une faible affinité pour PELa ( $K_d = 185 \mu\text{M}$ ). Pour le composé H, un meilleur accord est observé en utilisant un modèle qui comporte deux sites de reconnaissance aux sucres ( $n = 2$ ). Ce résultat indique que chaque branche de l'heptasaccharide interagit de manière différente avec PELa. Comme chaque monomère de PELa ne présente qu'un seul site de reconnaissance aux sucres, les données sont mises en accord avec un modèle  $n = 1$ . Les deux *N*-glycanes complexes E (asymétrique) et H (symétrique) possèdent une affinité similaire de  $K_d \sim 5 \mu\text{M}$ . Cette affinité est forte pour une lectine à mannose/glucose de légumineuse. A titre d'exemple, la lectine à mannose/glucose PAL qui appartient également à la tribu des *Dalbergieae* possède une affinité de l'ordre de 16 µM pour le  $\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-3)[\text{GlcNAc}\beta(1-2)\text{Man}\alpha(1-6)]\text{Man}$

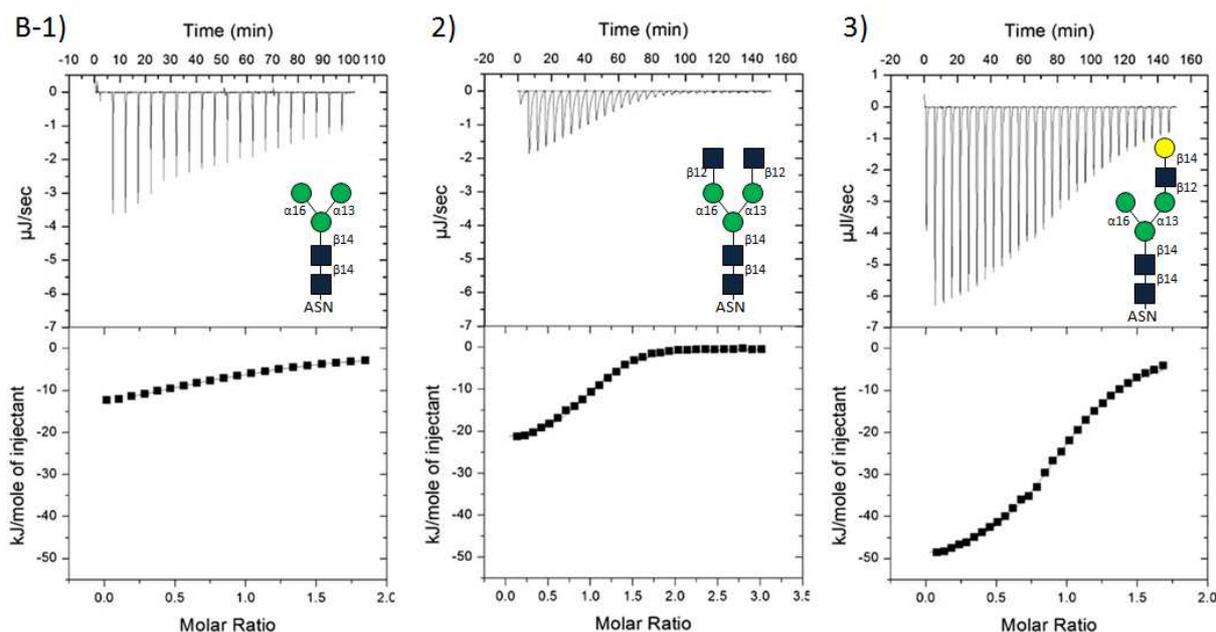
(Buts et al., 2006) soit trois fois inférieure à celle entre PELa et le *N*-glycane de type complexe et asymétrique.

Les résultats obtenus par ITC sont en désaccord avec ceux des puces à sucres. La différence d'affinité observée entre les résultats obtenus avec les puces à sucres et l'ITC peut provenir du fait que dans un cas le ligand est attaché à un support alors que dans l'autre cas le ligand est en solution. Cependant l'allure des pics obtenus dans les thermogrammes indiquent que les deux composés se fixent différemment à PELa. En effet les pics pour le composé H sont plus larges ce qui montre que le retour à l'équilibre est plus lent. De plus, bien que la même énergie libre soit observée ( $-30 \text{ kJ}\cdot\text{mol}^{-1}$ ), les contributions enthalpiques  $\Delta H$  et entropiques  $\Delta S$  sont différentes. Le composé H présente une enthalpie moyenne mais une contribution entropique favorable alors que le composé E présente une enthalpie plus forte et une contribution entropique défavorable (Figure 61).

Les liaisons mises en jeu durant les interactions lectines-sucres sont principalement des liaisons hydrogènes ainsi que des interactions de Van der Waals. Celles-ci mènent à une enthalpie souvent forte qui reflète la force des interactions entre le ligand et la protéine comparée à celles entretenues avec le solvant. Une enthalpie d'interaction favorable indique que la disposition entre les groupes donneurs et accepteurs des liaisons hydrogènes du site de fixation est correcte. La contribution entropique est généralement défavorable en raison d'un réaménagement du solvant dans le site de fixation ainsi que de la formation d'un complexe protéine-sucre qui contraint le ligand et les chaînes latérales des acides aminés impliqués dans l'interaction dans certaines conformations suite à la diminution des degrés de liberté. La différence du comportement des deux composés H et E peut être due à l'ajout du résidu GlcNAc sur la branche 1-6.

**A)**

Ligand	$K_d$ ( $\mu\text{M}$ )	$n$	$-\Delta G$ ( $\text{kJ/mol}$ )	$-\Delta H$ ( $\text{kJ/mol}$ )	$T\Delta S$ ( $\text{kJ/mol}$ )
Composé H	4,6 (0,3)	1,06 (0,06)	30,4	23,2 (0,3)	-7,2
Composé E	4,5 (0,2)	0,98 (0,06)	30,5	53,9 (1,8)	23,4
Composé C	133	1,03	22,1	18,7	-3,4



**Figure 61: Données de microcalorimétrie pour l'interaction des composés C, H et E avec PELA.** A) Données d'ITC caractérisant l'interaction entre PELA et les composés C, H et E. B) Mesure ITC obtenue par la titration de PELA (60  $\mu\text{M}$ ) par des *N*-glycanes 1) C, 2) H et 3) E. En haut figurent les thermogrammes et en bas les courbes de titration et d'ajustement par rapport à un modèle théorique qui comporte un seul site. Excepté le composé C, les valeurs sont des moyennes d'au moins deux expériences réalisées à 298 K.

## 8.4. Etudes par cristallographie des interactions mises en place entre PELA et son ligand

### 8.4.1. Structure du complexe PELA/composé H

Des cristaux de PELA obtenus après quelques jours de croissance ont permis de résoudre les structures cristallographiques de différents complexes PELA /ligand. La structure de PELA complexée avec le trimannose (PDB 3ZVX) a été résolue par le Dr Annabelle Varrot.

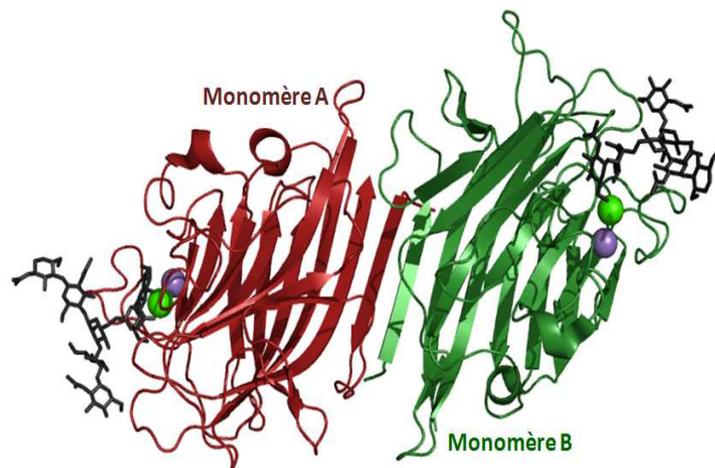
Ma contribution a été de résoudre et d'affiner la structure de PELA complexée avec l'heptasaccharide symétrique H. Les statistiques relatives aux données collectées et à l'affinement de la structure figurent dans le tableau 4.

Les cristaux contenant le complexe ont diffracté à 1,6 Å de résolution et présentent le groupe d'espace  $P2_12_12_1$  ( $a=51,40$  ;  $b=76,98$  et  $c= 125,59$  Å,  $\alpha=\beta=\gamma=90^\circ$ ). Le remplacement moléculaire a été utilisé pour résoudre le problème des phases à partir des coordonnées de la structure PELA/trimannose (PDB 3ZVX).

Ligne de Lumière (longueur d'onde)	BM30A (0,9797)
Groupe d'espace	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
Maille a, b, c (Å)	51,40; 76, 98; 125, 59
$\alpha, \beta, \gamma$ (°)	90
Reflections mesurées/unique	186416/55022
Résolution (Å)	47,57- 1,65 (1,74-1,65)
Multiplicité moyenne	3,4 (2,6)
R <sub>merge</sub>	0,054 (0,250)
Complétude (%)	91,4 (65,8)
Moyenne I/ $\sigma$ I	13,6 (3,9)
Facteur B de Wilson	14,00
R <sub>cryst</sub> /R <sub>free</sub>	15,5/18,8
R <sub>msd</sub> bonds (Å)	0,015
R <sub>msd</sub> angles (°)	1,59
R <sub>msd</sub> chiral (Å <sup>3</sup> )	0,098
Nombre d'atomes: protéine A/B	2004/1986
Bfac (Å <sup>2</sup> )	12,4/14,0
Nombre d'atomes: solvant A/B	295/223
Bfac (Å <sup>2</sup> )	25,7/26,6
Nombre d'atomes: ligand A/B	89/92
Bfac (Å <sup>2</sup> )	14,2/16,6
Nombre d'atomes: hetero A/B	30/12
Bfac (Å <sup>2</sup> )	24,7/25,7
Code PDB deposition	3ZYR

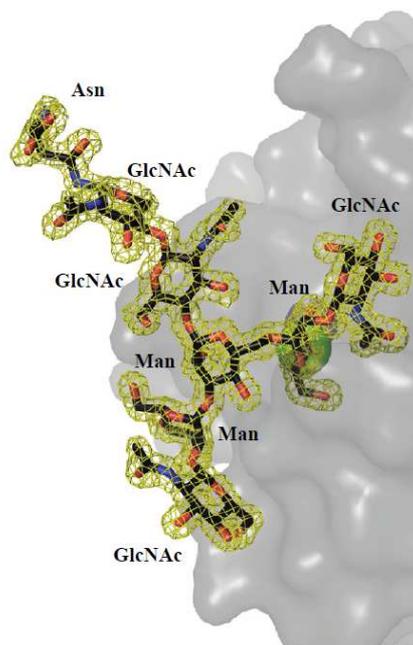
**Tableau 4: Statistiques pour les données et le raffinement de la structure PELa/Heptasaccharide.**

Dans les deux structures atomiques résolues, il a été observé que chaque monomère de PELa adopte le repliement classique des lectines de légumineuses c'est-à-dire un repliement de type sandwich  $\beta$  avec la présence des deux ions  $\text{Ca}^{2+}$  et  $\text{Mn}^{2+}$  proches du site de fixation des sucres. Les deux monomères s'assemblent par un arrangement antiparallèle de leur premier brin  $\beta$  ce qui permet la formation du dimère canonique des lectines de légumineuses (Loris et al., 1998) (Figure 62).



**Figure 62: Structure cristallographique d'un dimère de PELa complexée avec un N-glycane complexe et symétrique (composé H).** L'heptasaccharide H est représenté en bâtons noir. Les sphères vertes et violettes représentent respectivement les ions  $\text{Ca}^{2+}$  et  $\text{Mn}^{2+}$ .

La carte de densité obtenue permet d'identifier la totalité du composé H dans le CRD des deux monomères de PELa. Dans le monomère A, le résidu asparagine lié au *N*-glycane a été localisé grâce à sa stabilisation par un pont salin avec l'Arg221 qui provient d'une molécule symétrique de PELa (Figure 63). Il n'a pas pu être clairement localisé dans le monomère B. Mis à part cette différence, l'oligosaccharide est très similaire dans les monomères A et B en terme de conformation et de liaisons hydrogènes donc toutes les descriptions se réfèrent au monomère A.



**Figure 63:** Représentation de la densité électronique  $2mF_o-DF_c$  qui entoure à  $1\sigma$  le symétrique *N*-glycane de type complexe (0,47 électron par Å) dans le monomère A. La densité électronique est en jaune et la surface du monomère est en gris.

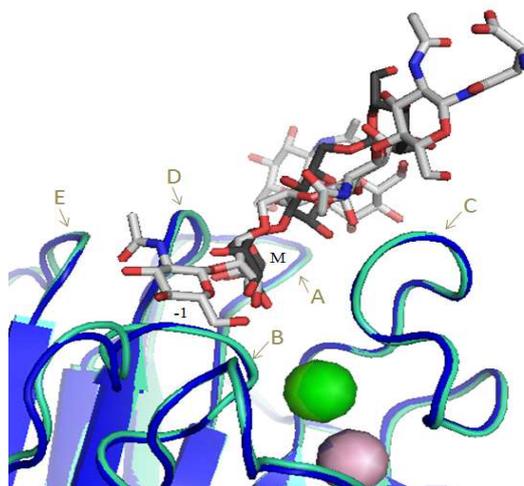
Tous les résidus, à l'exception de l'asparagine et du premier GlcNAc, interagissent avec la surface de la protéine PELa ce qui génère une surface de contact étendue (Figure 64). Au final, 13 résidus sont impliqués dans l'interaction de PELa/heptasaccharide avec la formation de 15 liaisons hydrogène directes et 6 pontés par une molécule d'eau. Dans l'interaction de PELa –composé H sont retrouvés les résidus Asp95, Ala113, Gly115, Asn145, Ser146, Asn147 et Gln231 généralement impliqués dans les interactions ligand-lectine à glucose/mannose des légumineuses mais également les résidus Ser54, Ser91, Asn92, Gly111, Gly229 et Gln231.

Le deuxième GlcNAc du cœur chitobiose est appuyé contre la boucle C et établit un pont hydrogène avec l'oxygène de la chaîne principale de l'Asn145. Le mannose de la branche 1-6 est localisé dans le site primaire d'interaction M avec la même orientation et le même réseau



hydrogènes entre le ligand et le monomère A réalisée par Ligplot. Les molécules d'H<sub>2</sub>O sont représentées en bleu et les liaisons hydrogènes sont représentées par des traits en pointillés vert.

La formation d'une liaison hydrogène entre la Ser54 et le GlcNAc au site -1 implique la boucle E (Figure 65). Elle va constituer un sous-site et permet de stabiliser l'extrémité du bras 1-6 de l'heptasaccharide. Cette boucle est rarement impliquée dans les interactions avec les sucres, excepté au sein du complexe ConA/Man ( $\alpha$ 1-2) Man et PAL/Man ( $\alpha$ 1-2) Man (Moothoo et al., 1999; Loris et al., 2003).



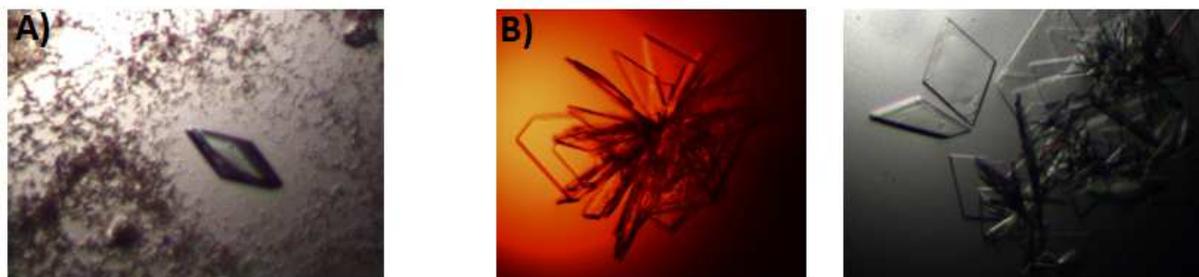
**Figure 65: Superposition du complexe PELa (cyan)/trimannose (gris foncé) et PELa (bleu)/heptasaccharide (gris clair) par Pymol.** Les deux ligands sont représentés par des bâtons, les sphères verte et violette représentent les ions Ca<sup>2+</sup> et Mn<sup>2+</sup>. Les boucles (A à E) qui interviennent dans l'interaction lectine/ligand sont indiquées.

Une étude sur la conformation de l'oligosaccharide en complexe avec PELa, faite par le Dr Anne Imberty, a mise en évidence que le  $\beta$ GlcNAc1-2 Man de la branche 1-6 présentait une conformation de haute énergie avec une distorsion. L'angle dièdre  $\psi$  de 105°C est en dessous de la valeur minimum qui est de 150°C (cf en annexe J. Biol. Chem- Figure 6-C) . Les conformations du  $\beta$ GlcNAc1-2Man de la branche 1-3, du  $\alpha$ Man1-3Man et du  $\alpha$ Man1-6Man sont de faible énergie. Cette étude infirme donc les résultats obtenus par les puces à sucre mais également les résultats d'ITC qui indiquaient que la présence du GlcNAc à la branche 1-6 n'était pas favorable.

#### 8.4.2. Essais de cristallogénèse du complexe PELa/composé E

Des essais de co-cristallisation de PELa avec le composé E ont été réalisés avec les conditions 42 (1.5 M sulfate d'ammonium, 0.1 M Tris-HCl, pH 8.5 et 12 % v/v glycérol) et 48 (100 mM Bicine, pH 9, 10 % PEG 20K et 2 % dioxane) du kit Crystal Screen II (Hampton Research). Ces derniers qui forment des amas de plaques ont été cassés pour obtenir des plaques uniques

(Figures 66). Un morceau de plaque a été placé dans une boucle et a été congelé dans l'azote liquide après transfert dans une solution où 30 % de glycérol ont été ajoutés à la solution de cristallisation. Des données ont été collectées à une résolution de 2.45 Å sur la ligne de lumière ID23-1 de l'ESRF, Grenoble. La carte de densité générée a montré que le sucre ne s'était pas fixé à la protéine et que nous avons une structure native. Nous n'avons pas poursuivi avec ces données. Des essais de trempage ont été réalisés sans succès.



**Figure 66: Cristaux obtenus avec 10mg/ml de PELa et 1mM du composé E dans les conditions A) 42 (1.5 M sulfate d'ammonium, 0.1 M Tris-HCl, pH 8.5 et 12 % v/v glycérol) et B) 48 (100 mM Bicine, pH 9, 10 % PEG 20K et 2 % dioxane) du kit Crystal Screen II (Hampton Research).**

## 8.5. Conclusion

L'étude de la spécificité de PELa réalisée par les puces à sucres ont permis d'identifier une forte affinité de PELa pour les *N*-glycanes de type complexes bi-antennés et asymétriques. L'étude menée par microcalorimétrie n'a pas permis de confirmer ces résultats. En effet la même affinité a été observée avec un *N*-glycane complexe symétrique (composé H) ou asymétrique (composé E). Cependant la différence de contribution enthalpique pour les deux composés avec une plus forte enthalpie pour le composé E permet d'émettre l'hypothèse que la présence du GlcNAc à la branche 1-6 n'est pas favorable.

Aucun cristal comportant le complexe PELa/composé E n'a été obtenu. Cependant la structure du complexe PELa/composé H a permis de montrer que le GlcNAc de la branche 1-3 établit des contacts avec les sites +2 et +3 dans une conformation de basse énergie. Alors que la fixation du GlcNAc de la branche 1-6 induit une conformation à haute énergie due à une forte distorsion. Ce dernier résultat permet de confirmer que la présence du GlcNAc à la branche 1-6 est défavorable.

Cette étude a mis en évidence la forte affinité que PELa possède pour les *N*-glycanes asymétriques. Elle est inhabituelle chez les lectines à glucose/mannose de légumineuses et pourrait être mise à profit dans d'éventuelles applications biotechnologiques.

---

## 9. A la recherche d'inhibiteurs de PA-IL dérivés du galactose

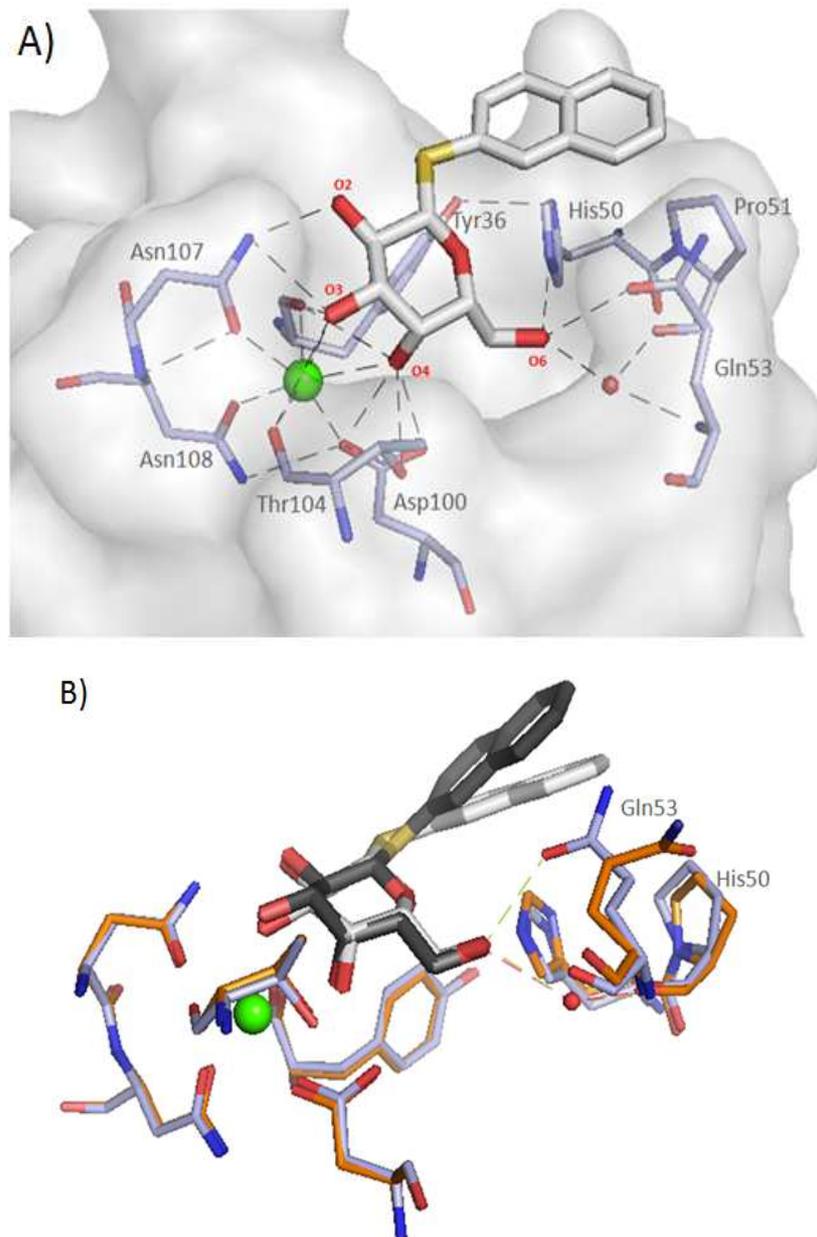
---

### 9.1. Introduction

La lectine PA-IL est l'une des deux lectines solubles produites par *Pseudomonas aeruginosa*. Elle est tétramérique, calcium dépendante et spécifique du D-galactose et les disaccharides présentant un résidu  $\alpha$ -D-galactosyl en position non-réductrice avec une affinité de l'ordre du millimolaire. Elle est à l'étude dans notre équipe depuis plus de 10 ans. Plusieurs structures cristallographiques ont permis de cartographier son site de fixation (Cioci et al., 2003; Blanchard et al., 2008) et plusieurs glycomimétiques ciblant cette lectine ont été synthétisés (Cecioni et al., 2011, 2012; Kadam et al., 2011; Reynolds et al., 2012). Bien que spécifique pour l' $\alpha$ -D-galactose, il a été montré que PA-IL était capable de fixer avec une forte affinité des  $\beta$ -galactosides présentant un groupement hydrophobe ou aromatique en aglycone tel le *p*-nitrophenyl- $\beta$ -galactoside (PNP- $\beta$ -Gal) (Garber et al., 1992).

Durant ses travaux de thèse, le Dr Bertrand Blanchard avait testé plusieurs molécules composées d'un galactose relié par une liaison  $\beta$  à un cycle aromatique. Ces sucres ont été synthétisés par l'équipe du Pr R. Roy de l'université du Québec à Montréal. Parmi tous les composés testés, le thio-naphtalene- $\beta$ -galactoside DEG 144 possédait la meilleure affinité pour PA-IL ( $K_d$  de 6  $\mu$ M). La structure atomique du complexe PA-IL/DEG 144 a été résolue par Dr Annabelle Varrot et Dr Bertrand Blanchard (PDB 4A6S, manuscrit en préparation) (Figure 67). Les interactions entre le galactose, le calcium et PA-IL sont identiques à celles observées entre PA-IL et le D-galactose (Cioci et al., 2003). Les hydroxyles O3 et O4 du résidu galactose participent à la coordination du calcium. Une molécule d'eau renforce la position du sucre en créant un pont hydrogène entre l'hydroxyle O6 et les groupements carbonyle et amine de la Pro51 et de la Gln53 respectivement (Figure 67 A). Si le galactose se fixe de la même manière dans les quatre sites du tétramère PA-IL, la conformation du groupement naphtalène est différente entre les monomères A et D et les monomères B et C. Une variation d'angle  $\varphi$  et  $\psi$  de 35° et 25°, respectivement, est observée au niveau de la liaison osidique. Ce changement de conformation est corrélé avec le changement de conformation de la chaîne latérale du résidu Gln53. Dans les chaînes B-C, la Gln53 forme une liaison hydrogène avec l'hydroxyle O6, alors que dans les chaînes A-D le résidu Gln53 est tourné vers le solvant (Figure 67 B). Le résidu His50 ne subit pas de changement de

conformation et une interaction de type « empilement en forme de T » est observée avec le groupement naphthalène (McGaughey et al., 1998). Ce type d'interaction est également observé au sein du complexe PA-IL/PNP- $\beta$ -Gal (Kadam et al., 2011).



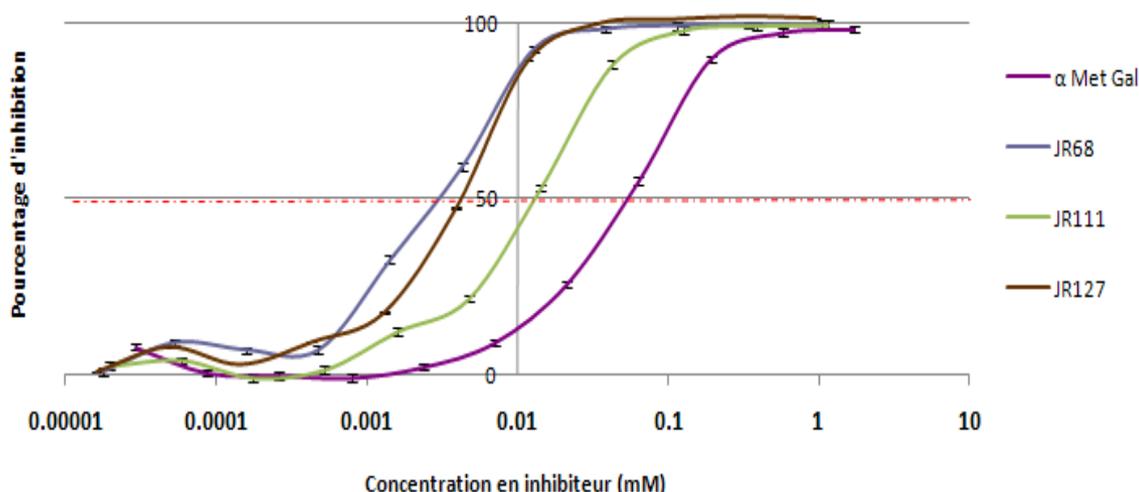
**Figure 67: Structures cristallographiques du complexe PA-IL/DEG144.** A) Représentation des interactions mises en place entre le DEG144 et le site de reconnaissance C de PA-IL. L'ion calcium et la molécule d'eau sont représentés par une sphère verte et rouge respectivement. Les liaisons hydrogènes sont indiquées par des traits en pointillés noirs. B) Superposition du site D et C de PA-IL complexé avec le DEG144. Les acides aminés du site D et C sont représentés en orange et en bleu clair, respectivement. Le DEG144 du site D et C est représenté en noir et blanc, respectivement. Les liaisons hydrogènes entre le Gln53 et le DEG144 formées dans le site C sont en vert et en rose dans le site D.

Dans cette étude mon travail consistait à tester de nouvelles molécules synthétisées utilisant comme base la structure du DEG144. Les molécules ont d'abord été criblées par test ELLA, puis une étude par microcalorimétrie a été réalisée sur les meilleurs composés afin de déterminer leurs paramètres thermodynamiques. Des essais de cristallogenèse ont également été réalisés avec un très bon inhibiteur, le JR127. Ces travaux font l'objet d'une publication en cours de rédaction.

## 9.2. Criblage par test ELLA d'inhibiteurs dérivés du galactose

Le test ELLA permet de déterminer la capacité du ligand à inhiber l'adhésion de la lectine à une surface de glycopolymère ( $\alpha$ -Gal-PAA). Cette technique est utilisée pour cribler rapidement les meilleurs inhibiteurs parmi les molécules synthétisées.

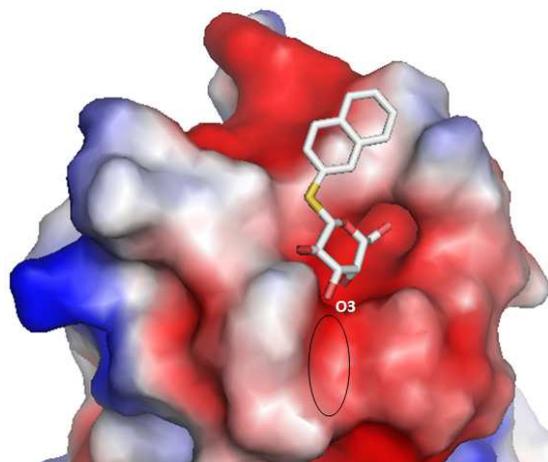
Une gamme étalon est dans un premier temps réalisée pour déterminer la concentration finale de PA-IL à utiliser pour cribler les inhibiteurs. Une concentration de  $3 \mu\text{g}\cdot\text{ml}^{-1}$  permet d'obtenir une réponse dans la zone linéaire (absorbance en fonction de la concentration en lectine). L' $\text{IC}_{50}$  ainsi que le potentiel d'inhibition (I %), déterminés par ELLA pour chaque composé, figurent dans le tableau 5. L' $\text{IC}_{50}$  indique la concentration en inhibiteur nécessaire à provoquer 50 % d'inhibition. Cette valeur est directement lue sur la courbe d'inhibition (Figure 68). Le potentiel d'inhibition est calculé en divisant l' $\text{IC}_{50}$  de la molécule inhibitrice à celui de l' $\alpha$ MeGal ( $\alpha$ -D-méthyl-galactoside), utilisé comme référence.



**Figure 68: Courbes d'inhibition ELLA des molécules dérivées du galactose.** Dans toutes les expériences ELLA menées avec PA-IL, l' $\alpha$ MeGal est utilisé comme inhibiteur de référence. Dans cet exemple, les courbes sélectionnées correspondent à de meilleurs inhibiteurs que l' $\alpha$ MeGal car leur  $\text{IC}_{50}$  est plus faible. Le JR68 est le DEG144.

La plupart des composés testés dans cette série ont pour structure de base la structure du DEG144 où d'autres résidus ont été greffés principalement à l'extrémité non-réductrice dans le but d'augmenter l'affinité pour PA-IL et donc d'accroître le pouvoir d'inhibition du composé.

Une grande partie des composés possèdent un cycle benzénique couplé par chimie « click » en O3 du  $\beta$ Gal (JR83, JR76, JR77, JR85, JR84) (Tableau 5). Le cycle benzénique pourrait se plaquer contre la zone hydrophobe et ainsi augmenter l'affinité du composé (Figure 69).



**Figure 69: Représentation de la surface du complexe PA-IL/DEG144.** Les résidus chargés négativement figurent en rouge et les résidus positifs en bleus. Les résidus hydrophobes sont en blanc. Le cercle représente le potentiel emplacement du cycle benzénique que possèdent certains composés.

D'autres composés possèdent l' $\alpha$ Gal1-4Gal à la place du  $\beta$ Gal (composé JR99, JR103) (Tableau 5). Les glycolipides présentant le disaccharide  $\alpha$ Gal1-4Gal sont les ligands naturels de PA-IL sur les tissus humains, et l'interaction a été caractérisée par ITC et modélisation moléculaire (Blanchard et al., 2008). Le mélibiose (Gal $\alpha$ 1-6Glc) qui est également un bon ligand de PA-IL a été utilisé pour la synthèse des composés JR133 et JR131 (Tableau 5).

Les composés JR127 et JR133 ont un groupement 4-méthylcoumarine et le composé JR111 a un groupement fluorobenzène à la place du groupement naphthalène (Tableau 5).

Inhibiteur	Structure de l'inhibiteur	IC <sub>50</sub> (mM)	I%	Inhibiteur	Structure de l'inhibiteur	IC <sub>50</sub> (mM)	I%
JR 68 DEG 144		3.10 <sup>-3</sup>	20	JR 84		8,1.10 <sup>-1</sup>	0,1
JR 127		5.10 <sup>-3</sup>	17	JR 85		1,45	0,04
JR 133		5.10 <sup>-2</sup>	4,17	JR 77		1,49	0,05
JR 111		2.10 <sup>-2</sup>	2,78	JR 98		1,5	0,05
JR 131		1,7.10 <sup>-1</sup>	1,23	JR 90		1,51	0,03
α Met Gal		7.10 <sup>-2</sup>	1	JR99		1,49	0,05
JR 150		1,7.10 <sup>-1</sup>	0,44	JR 76		1,52	0,05
JCC 12		1,9.10 <sup>-1</sup>	0,28	JR 103		1,67	0,04
JCC 11		3,8.10 <sup>-1</sup>	0,13	JR 83		1,67	0,04

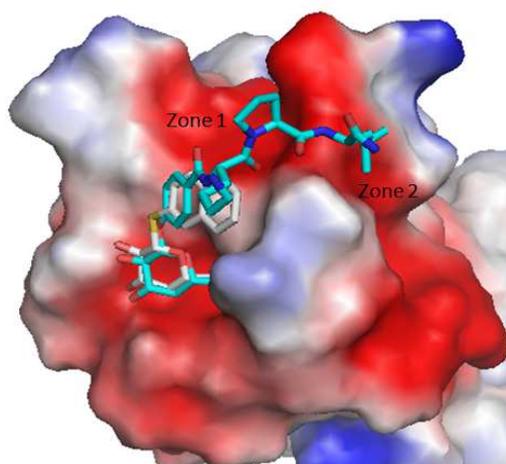
**Tableau 5: Structure, IC<sub>50</sub> et potentiel d'inhibition des différents inhibiteurs testés par la technique ELLA. Le JR68 est le DEG144.**

D'après les résultats du test ELLA, le meilleur inhibiteur de PA-IL reste le DEG144. Cependant le composé JR127 présente également une faible IC<sub>50</sub> et un bon potentiel d'inhibition (Tableau). Il est clair que la présence d'un résidu greffé au niveau de l'hydroxyle O3 du βGal diminue considérablement le potentiel d'inhibition du composé. Ce résultat n'est pas surprenant, car il est connu que l'hydroxyle O3 participe à la coordination du calcium et forme de nombreuses liaisons hydrogènes avec les acides aminés de PA-IL. Une substitution sur l'hydroxyle O3 doit déstabiliser l'interaction du βGal avec le site de reconnaissance de PA-IL.

Il a également été observé que la présence du Galα1-6Glc ou du Galα1-4Gal diminue le potentiel d'inhibition. La comparaison du composé JR131 et JR99 montre que Galα1-6Glc présente une meilleure affinité pour PA-IL que le Galα1-4Gal. Pour le composé JR131 et JR133, la substitution du groupement naphthalène par un groupement 4-méthylcoumarine permet une faible augmentation du potentiel d'inhibition. Il est fort possible que l'ajout d'un

résidu Glc ou Gal déplace le groupement naphthalène et déstabilise les interactions qu'il créait avec les résidus hydrophobes : Pro51, His50 et Gln53 (Blanchard et al., 2008).

En se basant sur l'étude de Kadam et al., il serait intéressant d'étudier la possibilité de greffer des résidus sur un aglycone de façon à atteindre les zones 1 et/ou 2 (Figure 70) (Kadam et al., 2011).



**Figure 70: Superposition du DEG144 et du GalA-KPL-NH2 au niveau du site actif de PA-IL.** Les composés DEG144 et GalA-KPL-NH2 sont représentés par des bâtons blancs et cyan respectivement. Les résidus chargés négativement figurent en rouge et les résidus positifs en bleu. Les résidus hydrophobes sont en blanc.

### 9.3. Etude par ITC

La détermination des paramètres thermodynamiques de l'interaction du DEG144, JR127 et JR133 avec PA-IL a été faite par ITC. Le composé PNP- $\beta$ -Gal est connu pour être un bon inhibiteur de PA-IL (Garber et al., 1992; Kadam et al., 2011). L'étude de l'affinité de ce composé pour PA-IL par ITC permet d'avoir une valeur de comparaison. Le composé JR133 possède le même  $K_d$  ( $\sim 20 \mu\text{M}$ ) que PNP- $\beta$ -Gal (Tableau 6). Les deux composés, DEG144 et JR127, ont le même  $K_d$  qui est de 5-6  $\mu\text{M}$ . Cette affinité est forte pour PA-IL, elle est 4 et 10 fois supérieure à celle obtenue avec le PNP- $\beta$ -Gal et le  $\beta$ MeGal, respectivement. Actuellement, dans la littérature deux autres molécules monovalentes possèdent une affinité similaire pour PA-IL : le GalA-KPL-NH2 ( $K_d \sim 4,2 \mu\text{M}$ ) (Kadam et al., 2011) et le GalPhNAz-triaz ( $K_d \sim 5,8 \mu\text{M}$ ) (Cecioni et al., 2012).

Si les deux composés, DEG144 et JR127, possèdent la même énergie libre ( $\sim 30 \text{ kJ/mol}$ ), JR127 possède une plus forte contribution entropique ( $T\Delta S$  de -17,1) que le DEG144 ( $T\Delta S$  de -3,8) (Tableau 6). Cette différence peut s'expliquer par la présence des deux atomes

d'oxygène supplémentaires qui peuvent constituer des accepteurs pour des liaisons hydrogènes stabilisant alors la conformation de certaines chaînes latérales de PA-IL.

Ligand	$K_d$ ( $\mu M$ )	$n$	$-\Delta G$ (kJ/mol)	$-\Delta H$ (kJ/mol)	$-T\Delta S$ (kJ/mol)
$\beta$ MeGal	55,7+/- 3,1	1	24,3	19,0	5,3
PNP- $\beta$ -Gal	21,2	1	26,7	45	18,3
DEG 144	6,3 +/- 0,4	1	29,7	33,9	3,8
JR 127	5,4 +/- 0,3	1	30,1	47,2	17,1
JR 133	19,6 +/- 0,3	1	26,9	33,4	6,5
GalA-KPL-NH2 <sup>1</sup>	4,2 +/- 0,9	1	7,4	10,8	3,4
GalPhNAz-Triaz <sup>2</sup>	5,8 +/- 0,9	0,8	30	53	23

**Tableau 6: Synthèse des données thermodynamiques obtenues par microcalorimétrie isotherme de titration pour les  $\beta$ MeGal, PNP- $\beta$ -Gal, DEG144, JR127 et JR133.** Toutes les données correspondent à une moyenne de deux expériences réalisées à 298K sauf pour le PNP- $\beta$ -Gal. 1 et 2 correspondent aux données d'après (Kadam et al., 2011) et (Cecioni et al., 2012), respectivement.

## 9.4. Conclusion

Parmi les composés testés deux possèdent une forte affinité pour PA-IL ( $K_d \sim 5-6 \mu M$ ): le DEG144 et le JR127. La substitution du groupement naphthalène par un groupement 4-méthylcoumarine ne permet pas d'augmenter l'affinité envers PA-IL mais peut accroître le nombre de liaisons hydrogènes avec des résidus de la protéine et favoriser une meilleure stabilisation du sucre.

Des essais de cristallogénèse pour le complexe PA-IL/JR127 ont été menés en utilisant les conditions utilisées précédemment pour cristalliser PA-IL mais aucun cristal n'a été obtenu. Il serait intéressant de cribler d'autres conditions de co-cristallisation.

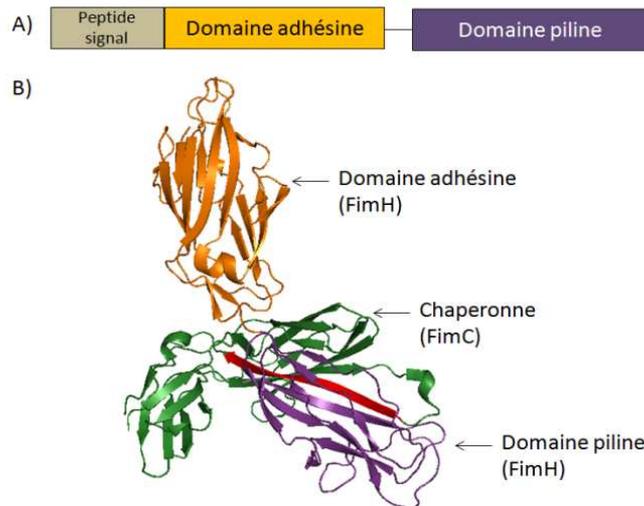
---

## 10. Les adhésines CupE6 et CupB6

---

Des travaux réalisés par l'équipe de Sophie de Bentzmann ont permis d'identifier plusieurs opérons de type Cup chez *Pseudomonas aeruginosa* et de montrer l'implication des adhésines Cup dans la formation de microcolonies (Vallet et al., 2004). Dans le cadre d'une approche de thérapie antiadhésive, ces adhésines sont prises pour cibles et sont donc soumises à une étude fonctionnelle et structurale. Mon travail de thèse a consisté à caractériser les adhésines CupE6 et CupB6 sur le plan structural et biochimique. Ces adhésines comme toutes les adhésines de type CU sont composées de deux domaines ; un domaine adhésine et un domaine piline situés aux extrémités N et C-terminales respectivement (Figure 71). Le domaine adhésine permet la fixation du pathogène aux cellules hôtes par la reconnaissance spécifique de récepteurs. Chez les adhésines homologues PapG et FimH chez *E. coli*, ce domaine est également appelé domaine lectine en raison de la présence d'un site de reconnaissance aux carbohydrates. Le domaine piline permet l'accroche de l'adhésine à l'extrémité du pilus. Ce dernier est assemblé par des liaisons non covalentes qui lient l'extension N-terminale de la sous-unité piline (n) au sillon hydrophobe de la sous unité-piline (n+1). La présence d'un sillon hydrophobe dans le domaine piline ne permet pas d'obtenir l'adhésine recombinante sous forme soluble. Pour résoudre ce problème, plusieurs stratégies peuvent être mises en place. Notre intérêt étant porté sur le domaine adhésine seul, la première stratégie a consisté à produire uniquement ce domaine. Dans une deuxième stratégie, nous avons choisi d'exprimer l'adhésine avec sa chaperonne. En effet la chaperonne va stabiliser par l'apport de son brin  $\beta$  G le sillon hydrophobe de la piline. Cette stratégie a permis d'obtenir par exemple le complexe FimH/FimC d'*E. coli* sous forme soluble et de résoudre sa structure 3-D (Choudhury et al., 1999) (Figure 71). Des essais d'expression/purification des domaines adhésine CupB6 et CupE6 ainsi que du complexe CupE6/E4 ont été effectués.

L'étude sur l'adhésine CupB6 fait suite aux travaux de thèse du Dr Bertrand Blanchard qui avaient permis l'obtention du complexe adhésine CupB6/chaperonne CupB4. Durant mes travaux de thèse, j'ai entrepris plusieurs expériences afin de mieux caractériser ce complexe pour obtenir des cristaux ainsi qu'identifier un ligand naturel saccharidique de CupB6.



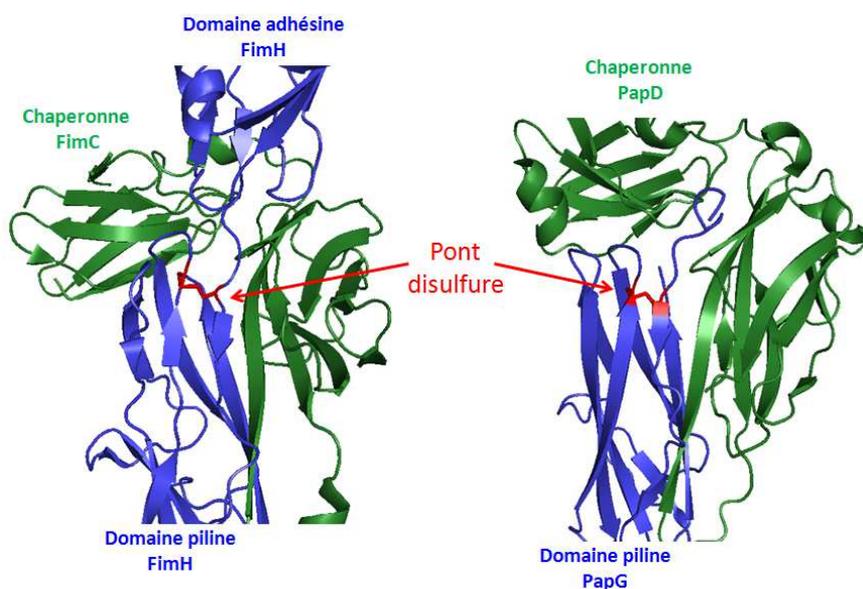
**Figure 71:** A) Illustration de la composition de l'adhésine CU. B) Structure atomique du complexe FimH/FimC d'*E. coli*. Le brin  $\beta$  G qui stabilise le sillon hydrophobe du domaine piline est représenté en rouge. (D'après Choudhury et al., 1999).

## 10.1. L'adhésine CupE6

### 10.1.1. Expression du domaine adhésine de CupE6

#### 10.1.1.1. Délimitation du domaine adhésine par analyse de séquence

Les résultats du PSI-Blast réalisés par Phyre2 (Kelley and Sternberg, 2009) mettent en évidence la très forte conservation de deux résidus cystéine qui forment un pont disulfure à l'apex des domaines pilines CU (Figure 72). La présence du premier résidu cystéine marque la partie N-terminale du domaine piline.



**Figure 72:** Représentation du pont disulfure très conservé au sein des pilines CU.



séquences prises en compte pour exprimer le domaine adhésine de CupE6 sous forme recombinante. Le domaine COG5430 est représenté par la ligne rouge, et les sous-domaines SCPU sont représentés par une ligne verte. B) Etude des régions désordonnées de CupE6. L'analyse est réalisée par un script conçu par le Dr Ambroise Desfosses, (EMBL, Grenoble) qui permet de rassembler les résultats de différents programmes de prédiction de désordre. Les acides aminés encadrés en rouges sont prédits comme désordonnés. C) Représentation des constructions choisies pour exprimer le domaine adhésine de CupE6.

L'alignement entre les pilines CupE1, CupE2, CupE3 et l'adhésine CupE6 met en évidence un domaine piline qui pourrait être délimité facilement grâce à la conservation de cette cystéine (C176 pour CupE6). Le domaine piline des différentes sous-unités CupE possède de nombreux résidus conservés (Figure 73). Bien que les extensions N-ter des sous-unités pilines CupE1-3 ne possèdent pas le motif conservé impliqué dans le mécanisme de l'échange de brin donneur, la conservation de la cystéine permet de démarquer la fin de cette région (Figure 73) (Giraud et al., 2011). Partant de l'hypothèse que le domaine piline de CupE6 débute au niveau de la cystéine 176, la séquence GIGPV pourrait représenter le linker qui sépare les deux domaines. Afin d'exprimer le domaine adhésine de CupE6, plusieurs constructions dépourvues du peptide signal de 24 acides aminés ont été réalisées. En premier lieu, la construction CupE6 25-170 a été choisie puis deux autres constructions, CupE6 25-146 et CupE6 25-151, qui prennent en compte le début du domaine COG5430. Les 12 premiers acides aminés de CupE6 identifiés comme désordonnés ont aussi été enlevés par la suite dans la construction CupE6 37-170 (Figure 73). L'analyse a été réalisée par un script (conçu par le Dr Ambroise Desfosses) qui permet de faire une synthèse sous la forme d'un graphique de tous les résultats obtenus par différents programmes de prédiction de désordre. D'après le graphique généré pour CupE6, toutes les régions en dessous de la valeur seuil de 0,5 sont considérées comme désordonnées.

#### 10.1.1.2. Tests d'expression du domaine adhésine CupE6

Les séquences nucléotidiques de toutes les constructions plasmidiques réalisées possèdent des codons rares. Pour une expression optimale, les tests d'expression ont été réalisés avec des souches capables d'exprimer les codons rares. Tous les résultats de tests d'expression sont résumés dans le tableau 7. L'utilisation du vecteur pProEx-HtB ne permettant pas d'exprimer le domaine adhésine CupE6 sous forme soluble, le vecteur pET-M30 a été utilisé pour la construction CupE6 37-170. Ce type de vecteur possède une étiquette GST volumineuse qui favorise l'expression de la protéine sous forme soluble (Smyth et al., 2003). Malgré la présence d'une très forte expression, la protéine CupE6 37-170 reste insoluble. La

température d'expression pouvant également avoir un impact sur le repliement de la protéine, deux températures ont été testées, 30°C et 16°C (Vera et al., 2007).

Construction	Souche	Milieu	Conditions d'expression	Résultats
pProEx-HtB-CupE6 25-146	RosettaGami2 (DE3) pLysS	LB	DO <sub>600</sub> 0.8- (0,2mM) IPTG 16°C-16H	Corps d'inclusion
			DO <sub>600</sub> 0.8- (0,2 mM) IPTG 30°C-3H	Aucune bande visible par coloration au Bleu de Coomassie après purification sur colonne nickel
pProEx-HtB-CupE6 25-151	RosettaGami2 (DE3) pLysS	LB	DO <sub>600</sub> 0.8- (0,2 et 0,8 mM) IPTG- 16°C-16H	Corps d'inclusion- Vérification par immunomarquage
		LB- TB	DO <sub>600</sub> 0.8- (0,2 mM) IPTG 30°C-3H	Corps d'inclusion
	BL21 (DE3) RIL	DO <sub>600</sub> 0.8- (0,2mM) IPTG 16°C-16H		
	ArcticExpress (DE3) RIL	DO <sub>600</sub> 0.6- (0,2mM) IPTG 16°C-16H		
pProEx-HtB-CupE6 25-170	RosettaGami2 (DE3) pLysS	LB	DO <sub>600</sub> 0.6- (0,4mM) IPTG 30°C-3H	Aucune bande visible par coloration au Bleu de Coomassie après purification sur colonne nickel
	BL21 (DE3) RIL	LB	DO <sub>600</sub> 0.7- (0,4mM) IPTG 30°C-3H	
	ArcticExpress (DE3) RIL	LB	DO <sub>600</sub> 0.9- (0,4mM) IPTG 30°C-3H	
pProEx-HtB-CupE6 37-170	RosettaGami2 (DE3) pLysS	LB-TB	DO <sub>600</sub> 1- (0,2mM) IPTG 16°C-16H	Corps d'inclusion Vérification par immunomarquage
		Milieu autoinductible	37°C pendant 7 H et 16°C pendant 16H	
	BL21 (DE3) pLysS	LB-TB	DO <sub>600</sub> 1- (0,2mM) IPTG 16°C-16H	
		Milieu autoinductible	37°C pendant 7 H et 16°C pendant 16H	
	ArcticExpress (DE3) pLysS	LB-TB	DO <sub>600</sub> 1- (0,2mM) IPTG 16°C-16H	
		Milieu autoinductible	37°C pendant 7 H et 16°C pendant 16H	
pET-M30-CupE6 37-170	BL21 (DE3) pLysS	LB	DO <sub>600</sub> 0.8- (0,2 mM) IPTG 30°C-3H	
	ArcticExpress (DE3) pLysS		DO <sub>600</sub> 0.8- (0,2 mM) IPTG 30°C-3H	

**Tableau 7: Conditions d'expression testées pour l'expression du domaine adhésine CupE6.**

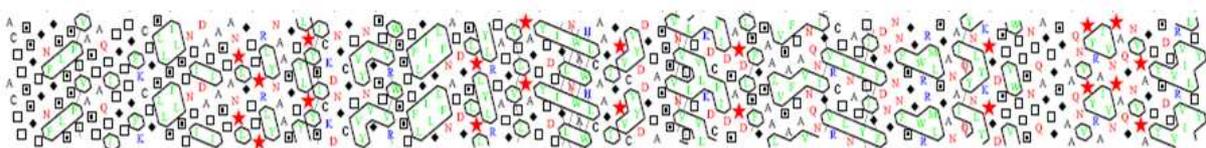
Un alignement de séquences réalisé par PSI-Blast dans Phyre2 met en évidence la conservation de 6 cystéines (positions 26, 58, 102, 107, 178 et 188) entre le domaine adhésine CupE6 et plusieurs domaines de protéines SCU (Spore Coat U). Des analyses bioinformatiques notamment avec le serveur DiANNA (Ferrè and Clote, 2006) indiquent la

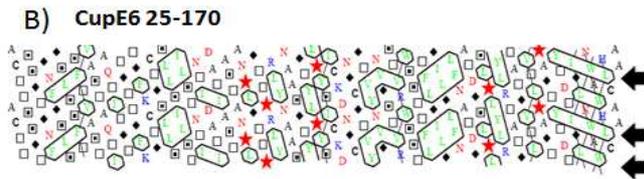
présence potentielle de plusieurs ponts disulfures. Ce programme utilise un réseau neural réalisé à partir de la prédiction de la structure secondaire, d'un alignement de séquence et d'une prédiction des cystéines oxydées pour ensuite prédire l'existence de potentiels ponts disulfures. La mauvaise formation des ponts disulfures dans le cytoplasme peut être la raison pour laquelle le domaine CupE6 est insoluble. Bien que la souche RosettaGami2 est conçue pour favoriser la formation des ponts disulfures, son utilisation dans cette étude n'a pas permis d'obtenir le domaine adhésine CupE6 sous forme soluble. La souche ArcticExpress qui par l'expression de deux chaperonnes, Cpn10 et Cpn60, peut aider à solubiliser certaines protéines à basse température a également été utilisée mais sans succès.

Parmi les différentes stratégies que l'on peut trouver dans la littérature, la solution la plus appropriée pour cette protéine serait de l'exprimer dans le périplasma. En effet toutes les sous-unités piline Cup sont repliées dans le périplasma avant d'être exportées à la surface de la bactérie. Contrairement au cytoplasme, le périplasma abrite un réseau d'enzymes redox qui catalyse la formation et l'isomérisation des ponts disulfures. Le clonage du gène *cupE6* avec son peptide signal permettrait l'export de la protéine dans le périplasma. Cette méthode oblige de placer l'étiquette qui aide à la purification en C-terminale de la protéine ce qui n'est pas optimal pour l'expression. L'utilisation d'un plasmide comme pMAL-p2x (BioLabs) permet à la fois d'exporter la protéine d'intérêt dans le périplasma (grâce à la présence du peptide signal MalE) mais également de la purifier (grâce à la fusion avec MalE).

Une analyse par HCA (Hydrophobic Cluster Analysis) (Gaboriaud et al., 1987) met en évidence un cluster hydrophobe en C-terminal des constructions CupE6 25-170 et CupE6 37-170, ce qui peut également être la cause de l'insolubilité de la protéine (Figure 74). L'ajout d'une partie du linker où d'un résidu chargé serait un moyen de masquer cette région hydrophobe. Les constructions qui ont permis de résoudre la structure atomique des domaines adhésines FimH et PapG contiennent une partie du linker. En ce qui concerne la construction de FimH, l'ajout de la thréonine vient sûrement stabiliser le brin  $\beta$  formé par trois valines (Bouckaert et al., 2005). Pour la construction du domaine adhésine PapG, les résidus NIGG du linker ont été ajoutés (Dodson et al., 2001). Il serait intéressant de cloner les constructions CupE6 25-174 et CupE6 37-174 dans un vecteur qui permet une expression périplasmique.

#### A) CupE6





**Figure 74: Représentation graphique HCA de A) CupE6 et B) cupE6 25-170.** Les flèches noires indiquent des patches hydrophobes qui peuvent être responsables de l'insolubilité des protéines CupE6 25-170 et CupE6 37-170.

## 10.2. Expression du complexe CupE6/E4

Face aux difficultés rencontrées pour exprimer le domaine adhésine CupE6, des expériences visant à co-exprimer le complexe CupE6/E4 ont été réalisées. Les gènes (dépourvus de leur peptide signal), qui codent pour l'adhésine CupE6 et la chaperonne CupE4, ont été clonés dans le vecteur pETDUET. Les 12 premiers acides aminés de CupE6 étant définis comme désordonnés, deux constructions ont été réalisées : pETDUET-CupE6 25/CupE4 et pETDUET-CupE6 37/CupE4.

Comme décrit ci-dessus, la séquence nucléique de CupE6 présente des codons rares et la protéine comporterait aussi au moins un pont disulfure. Les tests d'expression ont donc été réalisés avec des souches capables de coder pour des codons rares et/ou d'aider à la formation des ponts disulfures. Le tableau 8 résume les différentes conditions testées. Aucune expression du complexe CupE6/CupE4 n'a été observée dans ces conditions.

Construction	Souche	Milieu	IPTG	DO <sub>600</sub>	temps et température de culture	Résultats du test d'expression	Résultats de la purification sur résine de Nickel
pETDUET CupE6 25/CupE4	Rosetta gami2 (DE3) pLysS	LB	0,2 mM	0,6	16°C -16H	Aucune expression Vérification par immunomarquage	Aucune bande visible après coloration au bleu de Coomassie
			0,5 mM				
			0,2 mM	1			
	TB	0,2 mM	1				
	BL21 (DE3) RIL	LB	0,2 mM	0,8			
Artic Express (DE3) RIL	LB	0,2 mM	0,6				
pETDUET CupE6 37/CupE4	Rosetta gami2 (DE3) pLysS	TB	0,4 mM	0,7	30°C-3H Ou 16°C-16H		
	BL21 (DE3) RIL						
	Artic Express (DE3) RIL						

**Tableau 8: Conditions d'expression testées pour le complexe CupE6/E4.**

Plusieurs facteurs peuvent expliquer la non-expression du complexe comme une mutation dans le vecteur pETDUET. Un des moyens de vérifier cette hypothèse serait de réaliser un clonage dont l'expression de la protéine est connue. L'autre solution plus coûteuse serait de séquencer entièrement le vecteur d'expression. Il est également possible que la protéine CupE6 soit toxique pour les souches utilisées. Dans ce cas, la souche d'*E.coli* C43 conçue pour des protéines toxiques serait plus adéquate.

### 10.3. Conclusion

Les tests d'expression réalisés dans le cytoplasme avec les différentes constructions n'ont pas permis de produire le domaine adhésine CupE6 sous forme soluble. La présence de ponts disulfures est certainement responsable de l'expression du domaine CupE6 dans les corps d'inclusion chez *E. coli*. Dans la littérature, les domaines adhésines de PapG et FimH ainsi que toutes les sous-unités pilines Pap et Fim sont exprimées dans le périplasme. Seule la sous-unité FimA est exprimée dans le cytoplasme mais la purification de cette protéine nécessite une renaturation à partir des corps d'inclusion. De plus, une étape permettant de catalyser la formation du pont disulfure est également réalisée (Puorger et al., 2011). L'expression du domaine adhésine CupE6 dans le périplasme pourrait permettre l'obtention de la protéine sous forme soluble et correctement repliée. Parmi les constructions réalisées, les constructions CupE6 25/170 et CupE6 37/170 semblent les plus représentatives du domaine adhésine. Cependant il serait très intéressant de réaliser également les constructions CupE6 25/174 et CupE6 37/174 où l'ajout des 4 acides aminés GIGP peut également jouer sur la solubilité du domaine (Dodson et al., 2001).

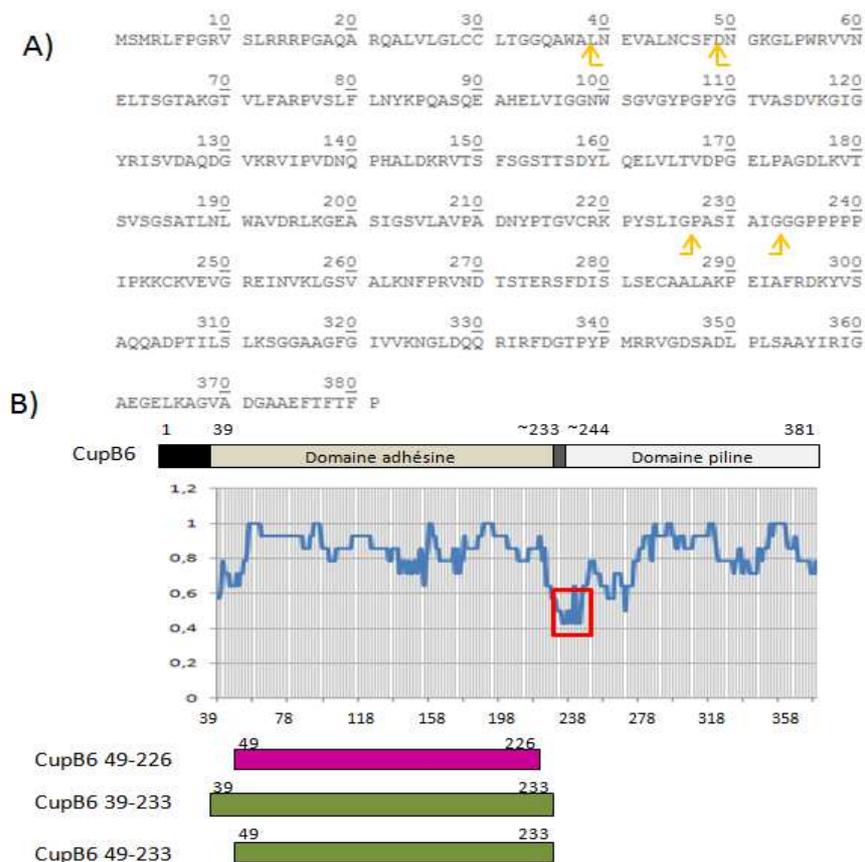
Face aux difficultés rencontrées à exprimer le domaine adhésine, la deuxième stratégie qui consiste à co-exprimer l'adhésine entière CupE6 avec sa chaperonne a été entreprise. Les résultats des tests d'expression ne montrent toutefois aucune expression visible. Il est possible que le complexe soit toxique pour les différentes souches utilisées. Dans ce cas, la souche d'*E.coli* C43 conçue pour des protéines toxiques serait plus adéquate. Là encore l'expression dans le périplasme serait peut-être plus appropriée au vu du nombre de ponts disulfures prédits dans le domaine adhésine de CupE6.



noir. La structure secondaire de FimH et la prédiction de la structure secondaire de CupB6 sont indiquées. B) L'extension N-ter de la piline CupB1 est encadrée et les résidus prédits qui correspondent aux résidus P1 à P4 sont indiqués par des points rouges. Les étoiles correspondent aux résidus conservés entre les deux séquences CupB6 et CupB1.

Cet alignement a permis de délimiter l'extrémité N-terminale du domaine piline de CupB6 (résidus 244-381) et de proposer une extrémité C-terminale pour le domaine adhésine de CupB6 (résidus 39-233). La présence d'un domaine Pfam00419 (245-381) qui caractérise le domaine piline et qui est présent chez l'adhésine PapG ainsi que les sous-unités pilines majeures PapA, CupA1, CupB1, CupC1 et CupD1 conforte le postulat que le domaine piline de CupB6 comprend les résidus 244-381. Un alignement entre la séquence qui correspondrait au domaine piline CupB6 et celle de la sous unité piline CupB1 indique la présence de nombreux résidus conservés. L'extension N-ter de CupB1 a été définie suite à un alignement entre PapA et CupB1 qui montre la conservation des résidus P1 à P4 (Figure 75 B).

Partant donc de l'hypothèse que les 38 premiers acides aminés correspondent au peptide signal de CupB6 et que le domaine adhésine de CupB6 se situe entre les résidus 39-233, plusieurs constructions ont été réalisées (Figure 76).



**Figure 76: Constructions choisies pour exprimer le domaine adhésine de CupB6.** A) Séquence protéique de CupB6. Les flèches orange délimitent les séquences utilisées dans les constructions. B)

Représentation schématique de CupB6 et des trois constructions choisies pour exprimer le domaine adhésine de CupB6. Les résultats du script prédisent une région désordonnée (encadrée en rouge) qui pourrait correspondre au linker. Les 10 premiers acides aminés du domaine adhésine ont été retirés dans deux constructions car ils s'avèrent constituer une région moins ordonnée.

Dans toutes les constructions, le peptide signal est retiré. Une des constructions correspond au domaine adhésine entier (39-233) tel que nous l'avons défini. Dans les deux autres constructions, les 10 premiers acides aminés pouvant être moins ordonnés, sont retirés ainsi que sept résidus en C-terminal.

#### 10.4.1.2. Test d'expression et de purification

Les tests d'expression réalisés avec le vecteur pProEx-HtB permettant la fusion d'une étiquette hexahistidine clivable par la protéase TEV n'ont pas permis l'expression du domaine adhésine de CupB6. Pour aider l'expression, la fusion d'une étiquette plus volumineuse à l'extrémité N-terminale du domaine de CupB6 a été testée. Les différentes constructions réalisées sont résumées dans le tableau 9.

Construction	Souche	Milieu	Induction	DO <sub>600</sub>	temps et température de culture	Résultats du test d'expression	
pProEx-HtB CupB6 49-226	RosettaGami 2 (DE3)pLysS	LB-TB	0,2 mM IPTG	0,8	16°C -16H	Aucune bande après purification sur résine Nickel	
	BL21 (DE3) RIL						
	ArcticExpress (DE3) RIL						
pET-M30 CupB6 49-226	RosettaGami 2 (DE3)pLysS						
	BL21 (DE3) RIL						
	ArcticExpress (DE3) RIL						
pET-M30 CupB6 49-233	ArcticExpress (DE3)pLysS	LB		20% arabinose	0,6	30°C-3H ou 16°C -16H	Expression et protéine soluble
	pET-32a CupB6 49-226						
pET-32a CupB6 49-233	ArcticExpress (DE3) RIL						
	BL21 (DE3) RIL						
pBAD-M41 CupB6 49-226	ArcticExpress (DE3) RIL						
	BL21 (DE3) RIL						
pBAD-M41 CupB6 49-233	ArcticExpress (DE3) RIL						
	BL21 (DE3) RIL						
pBAD-M41 CupB6 39-233	ArcticExpress (DE3) RIL						
	BL21 (DE3) RIL						

**Tableau 9: Constructions et conditions des tests d'expression réalisés pour produire le domaine adhésine de CupB6.**

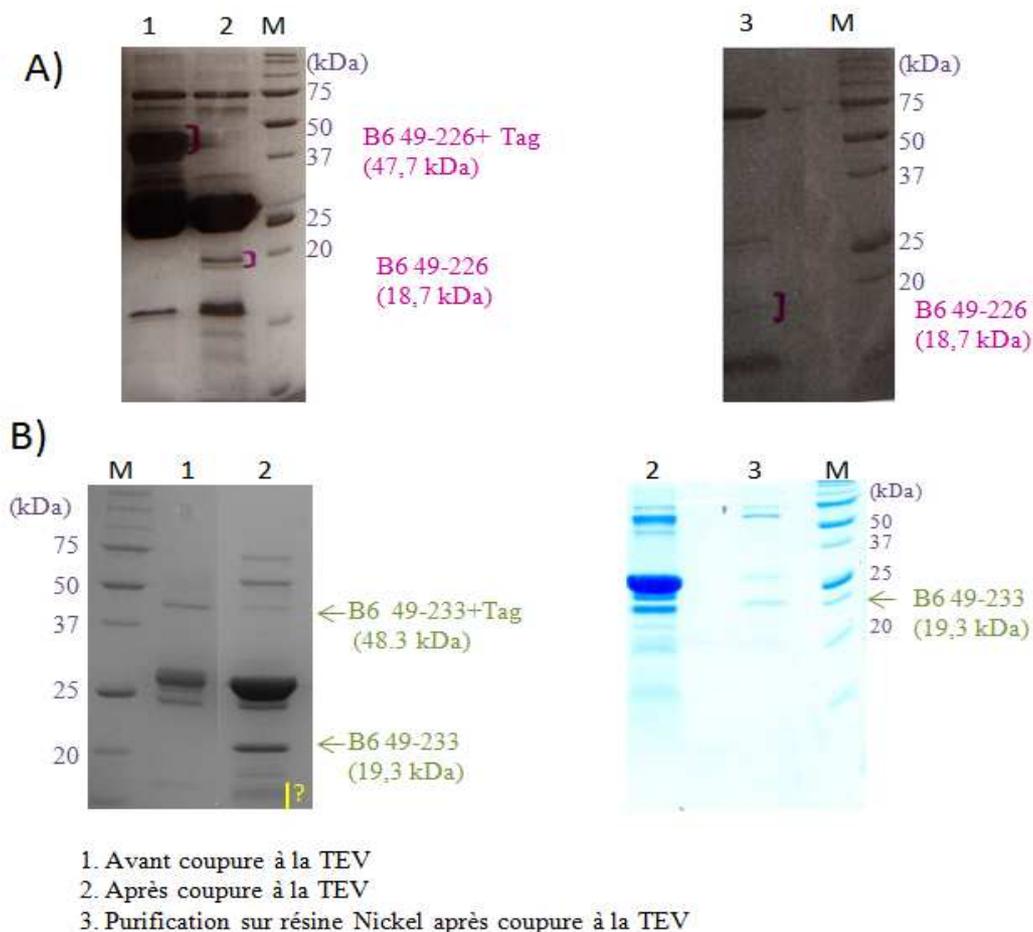
Des premiers essais de mise au point d'un protocole de purification ont été faits avec les constructions pET-M30-CupB6 49-226 et pET-M30-CupB6 49-233. Le vecteur pET-M30 permet la fusion d'une étiquette hexahistidine couplée à la glutathion-S-transférase en N-terminal du domaine adhésine de CupB6. La présence d'un site de clivage par la protéase TEV permet d'éliminer les étiquettes GST et hexahistidine. Les différents protocoles de purification testés sont résumés dans le tableau 10.

Construction	Conditions de culture	Etapes de purification	Résultats
pET-M30 CupB6 49-226	DO <sub>600</sub> 0,7 0,2mM IPTG 30°C-3H	- Résine de Nickel (batch) - Résine de Glutathione Sepharose 4B (Batch) - Coupure à la TEV - Résine de Nickel (batch)	Contaminants (Bandes 70, 57, 30, 15 kDa)
		- Ajout 1% triton X100-30 min - Résine de Nickel (batch) - Coupure à la TEV sur la résine de Glutathione Sepharose 4B (Batch) - Résine de Nickel (batch)	Problème au niveau de la coupure à la TEV
	DO <sub>600</sub> 0,7 0,2mM IPTG 16°C-16H	- Ajout 1% triton X100-30 min - Résine de Nickel (+100mM Lysine et Arginine+10% glycérol ) - Coupure à la TEV - Résine de Nickel	Contaminants (Bandes 70, 57, 30, 15 kDa)
	- Ajout 1% triton X100-30 min - Résine de Nickel (+1mM urée et 10% glycérol) - Coupure à la TEV - Résine de Nickel		
	- Ajout 1% triton X100-30 min - Résine de Nickel (+100mM Lysine et Arginine+10% glycérol) - Filtration sur gel S75 - Echangeuse d'anions		
pET-M30 CupB6 49-233	DO <sub>600</sub> 0,8 0,2mM IPTG 16°C-16H	- Résine de Nickel - Coupure à la TEV - Résine de Nickel	
		- Résine de Nickel (+100mM Lysine et Arginine) - Filtration sur gel S75	
		- Résine de Nickel (+1mM urée et 10% glycérol ) - Filtration sur gel S75	

**Tableau 10: Résumé des différents protocoles de purification testés avec les constructions pET-M30-CupB6 49-226 et pET-M30-CupB6 49-233.**

La protéine CupB6 49-226 est peu soluble et la présence de nombreux contaminants rend l'étape de purification très difficile. Afin d'augmenter la solubilité de la protéine, 1 % de triton X100 (selon les recommandations du fournisseur) ainsi que 100 mM de Lysine et arginine ont été testés (Golovanov et al., 2004). L'ajout de 1 M d'urée et de 10 % de glycérol

permet de limiter l'agrégation de la protéine observée. Deux bandes sont obtenues après l'expression et également après coupure à la TEV (Figure 77). Il se peut que la protéine CupB6 49-226 ne soit pas stable et que l'observation des deux bandes soit la résultante d'un clivage en C-terminal. La protéine CupB6 49-226 est beaucoup plus soluble et son élution à 500 mM d'imidazole facilite l'étape de purification. Après clivage de l'étiquette GST, la présence de plusieurs bandes laisse penser que la protéine CupB6 49-226 est dégradée (Figure 77). Cependant dans les deux cas, aucun protocole de purification n'a permis d'obtenir le domaine adhésine pur. En effet, plusieurs bandes contaminantes sont obtenues même après la dernière étape de purification où la protéine sans l'étiquette n'est pas retenue par la colonne nickel (Figure 77). Parmi ces contaminants, une bande de 30 kDa pourrait correspondre à l'étiquette GST seule. Malgré une purification sur résine nickel et/ou glutathion ou sur filtration sur gel, cette bande est toujours présente.



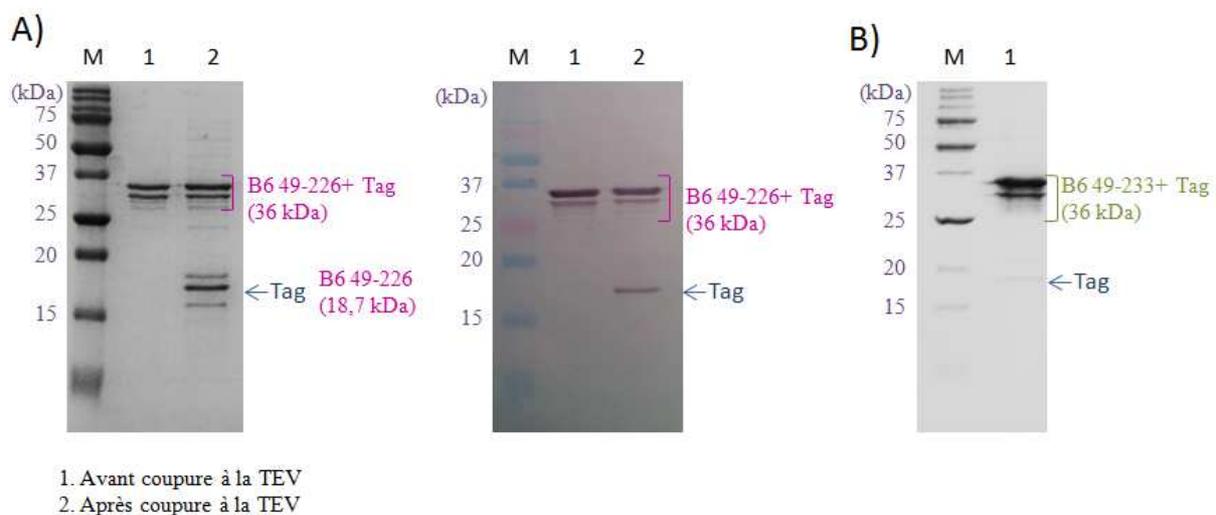
**Figure 77: Résultats sur gel SDS-PAGE 12% de l'obtention du domaine adhésine A) CupB6 49-226 et B) CupB6 49-233 après plusieurs étapes de purification et de clivage de l'étiquette His<sub>6</sub>-GST. La ligne jaune montre une potentielle dégradation de CupB6 49-233.**

D'autres essais de purification ont été réalisés utilisant les constructions pET32-TEV-CupB6 49-226 et pET32-TEV-CupB6 49-233. Le vecteur pET32-TEV comporte plusieurs étiquettes dont la protéine thioredoxine, une séquence S et une étiquette hexahistidine clivable par la protéase TEV. Les différents protocoles de purification testés sont résumés dans le tableau 11.

Construction	Conditions de culture	Etapes de purification	Résultats
pET32-TEV CupB6 49-226	DO <sub>600</sub> 0,6- 0,2mM IPTG 30°C-3H	- Résine de Zinc	Dégradation de la protéine Après clivage obtention de trois bandes
pET32-TEV CupB6 49-233	DO <sub>600</sub> 0,6- 0,2mM IPTG 16°C-16H	- Résine sur Nickel - filtration sur gel S200 - Coupure à la TEV - Résine sur Nickel - filtration sur gel S200	

**Tableau 11: Résumé des différents protocoles de purification testés avec les constructions pET32-TEV -CupB6 49-226 et pET32-TEV-CupB6 49-233.**

Après la première étape de purification par chromatographie d'affinité grâce à l'étiquette hexahistidine sur résine IMAC (Immobilized Metal Affinity Chromatography), deux bandes principales ainsi que plusieurs autres sont observées pour les deux constructions pET32-TEV-CupB6 49-226 et pET32-TEV-CupB6 49-233. Un immunomarquage utilisant l'étiquette S, permet d'identifier ces bandes et de remarquer la présence d'une dégradation du domaine adhésine CupB6 en C-terminal. La purification est réalisée avec des inhibiteurs de protéases. Lors du clivage de l'étiquette, deux bandes qui correspondent à CupB6 sont observées (Figure 78).



**Figure 78: Dégradation du domaine adhésine de CupB6 A) CupB6 49-226 et B) CupB6 49-233 après la première étape de purification par chromatographie IMAC (Immobilized Metal Affinity Chromatography).**

Cette dégradation est également observée avec les constructions pBAD-M41-CupB6 49-226 et pBAD-M41-CupB6 49-233. Dans l'hypothèse où le retrait des 10 premiers acides aminés déstabilisent la protéine, la construction pBAD-M41-CupB6 39-233 a été réalisée mais une dégradation est toujours observée.

L'observation de deux bandes qui correspondent au domaine adhésine CupB6 montre que la construction n'est pas stable. Une analyse par spectrométrie de masse permettrait de redéfinir une construction plus courte et plus stable. Les constructions qui ont permis d'obtenir les structures cristallographiques du domaine adhésine FimH et PapG d'*E. coli* comportent une partie du linker qui sépare les deux domaines. Il serait également intéressant de tester des constructions comprenant une partie du linker ce qui pourrait potentiellement rendre le domaine adhésine plus stable.

Différentes bandes contaminantes à 70 kDa, 57 kDa, 40 kDa peuvent correspondre aux chaperonines Dnak, GroEL, et GrpE respectivement. L'ajout de 2 mM ATP et de 10 mM MgSO<sub>4</sub> dans le lysat cellulaire avant l'étape de purification serait un moyen de les dissocier de notre protéine d'intérêt et d'améliorer la purification (Thain et al., 1996).

## **10.5. Le complexe CupB6/B4**

Durant les travaux de thèse du Dr. Bertrand Blanchard, les gènes qui codent pour l'adhésine CupB6 et pour la chaperonne CupB4 avaient été clonés dans le plasmide de co-expression pETDUET. Le protocole de purification mis en place était composé de trois étapes : (1) une chromatographie d'affinité sur résine nickel, (2) une filtration sur gel (Superdex 75) et (3) une échangeuse d'anions.

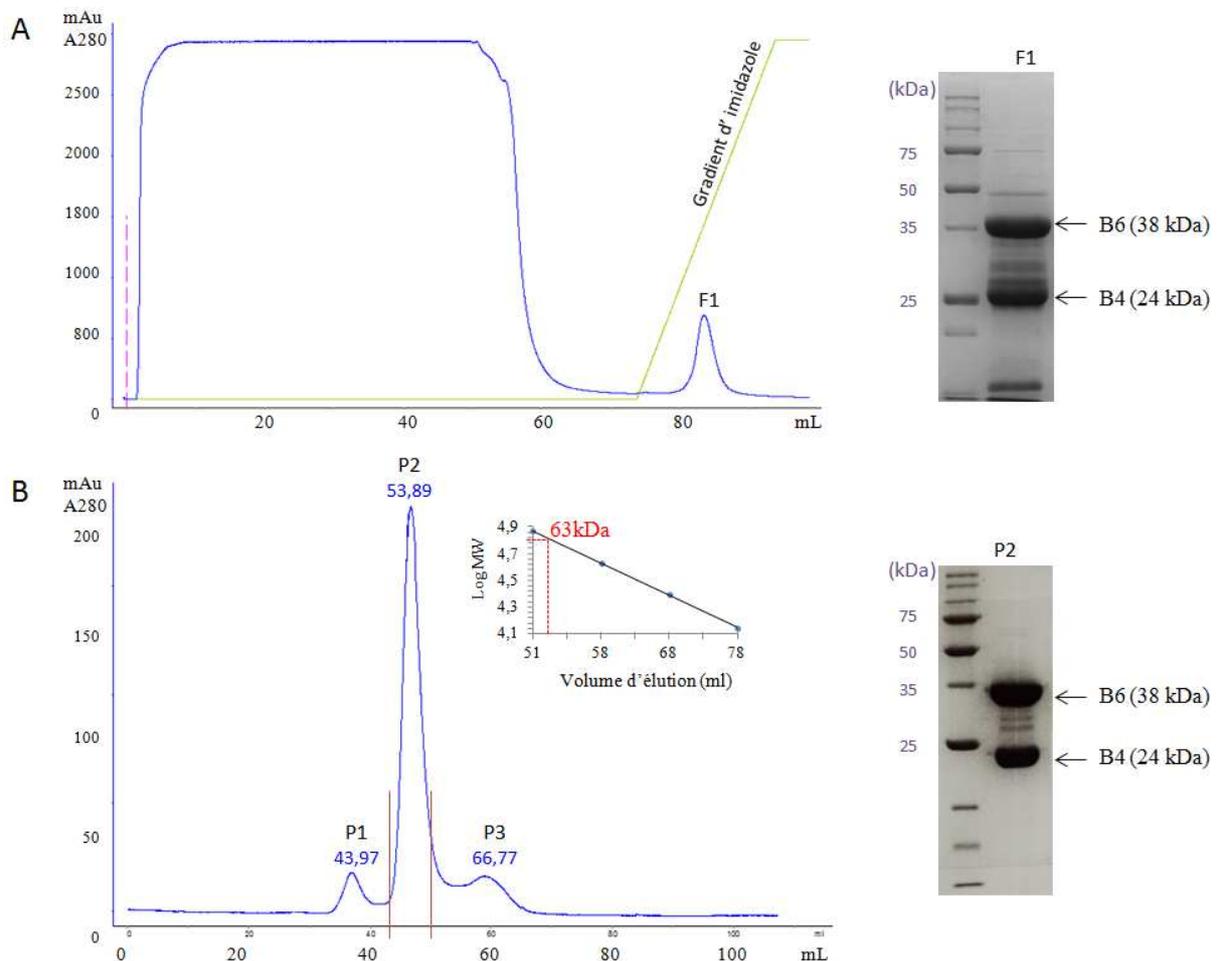
Afin d'identifier la spécificité de CupB6, le complexe marqué avait été envoyé au Consortium for Functional Genomic. Les résultats obtenus indiquaient un faible signal pour le Lewis b.

### **10.5.1. Optimisation de la purification de CupB6/B4**

Le complexe CupB6/B4 recombinant est co-exprimé dans la souche d'*E. coli* RosettaGami2 pLysS permettant l'expression de CupB6 qui comporte plusieurs codons rares et un pont disulfure. L'étiquette hexahistidine présente à l'extrémité N-terminale de l'adhésine permet la première étape de purification de CupB6/B4 par chromatographie d'affinité sur résine nickel. L'élution du complexe se fait à 300 mM d'imidazole. L'ajout de 50 mM de lysine et d'arginine dans le tampon d'équilibration améliore la purification sur la colonne nickel. En effet, un seul pic est obtenu après l'élution par gradient de 20 mM à 500 mM d'imidazole (Figure 79). L'analyse SDS-PAGE 12% qui correspond au pic d'élution concentré montre une

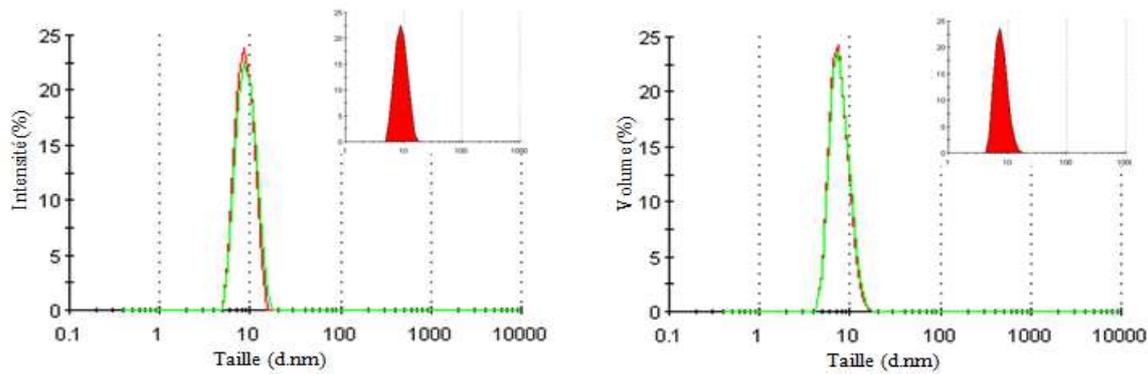
faible présence de contaminants. Sans l'ajout de la lysine et l'arginine, deux pics confondus sont obtenus et la plus grande présence de contaminants demande une étape de déssalage suivie d'une purification par échangeuse d'anions.

La seconde étape de purification du complexe se fait par une filtration sur gel en utilisant une Superdex 75 Hiload 16/60 (GE healthcare). Le complexe CupB6/B4 est élué à 53,89 ml ce qui correspond selon la courbe de calibration de la colonne à une protéine de 63kDa soit à un monomère de CupB6/B4 (62 kDa) (Figure 79). L'éluion du pic P1 dans le volume mort de la colonne de filtration sur gel correspond à des agrégats du complexe et le pic P3 à des contaminants. L'analyse sur gel SDS-PAGE du pic P2 concentré à 3,5 mg.ml<sup>-1</sup> montre que le complexe est pur (Figure 79).



**Figure 79: Profils de purification et analyse sur gel SDS-PAGE qui correspond à l'éluion du complexe CupB6/B4 A) par chromatographie sur colonne nickel et B) par filtration sur gel (Superdex 75).**

Une analyse par diffusion de lumière (DLS) a été réalisée afin d'estimer la polydispersité du complexe (Figure 80).



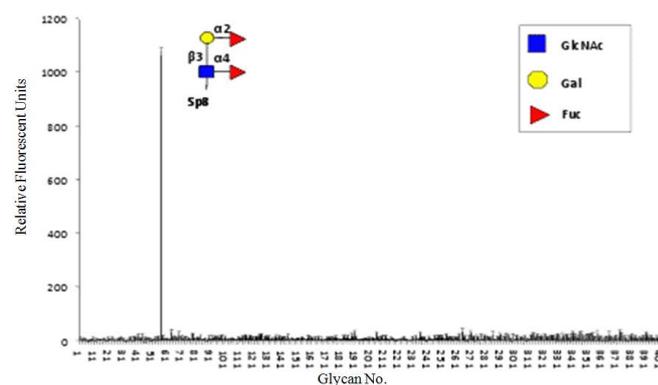
**Figure 80: Analyse du complexe CupB6/B4 par diffusion dynamique de la lumière.** Distribution du rayon hydrodynamique en intensité (à gauche) et en volume (à droite).

La présence d'un unique pic montre que l'échantillon est très homogène et l'indice de polydispersité de 0,1 (inférieur à 0,25) indique que le complexe est monodisperse (Figure 80). Ces deux paramètres sont essentiels pour effectuer les tests de cristallogénèse d'une protéine. Pour 2 L de culture en milieu LB, environ 100µl de complexe pur CupB6/B4 à 4 mg.ml<sup>-1</sup> sont obtenus en moyenne.

### 10.5.2. Étude de la spécificité de CupB6 pour les oligosaccharides

#### 10.5.2.1. Identification de ligands saccharidiques par puces à sucres

D'après les résultats des puces à sucres (Consortium for Functional Glycomics), le Lewis b serait le ligand naturel de CupB6. Bien que les intensités pour les 406 autres sucres testés soient nulles, le signal mesuré pour le Lewis b reste faible (Figure 81).



**Figure 81: Résultats du consortium for Functional Glycomics pour le complexe CupB6/B4.** Résultats de la thèse du Dr. Blanchard Bertrand.

Ce résultat a été vérifié par une mesure d'ITC. Le thermogramme obtenu ne présente aucune titration entre le complexe CupB6/B4 et le Lewis b. Afin de vérifier que cela ne soit pas dû à

une faible affinité entre le Lewis b et CupB6, des expériences d'ITC devraient être réalisées en conditions d'excès du ligand (Lewis b).

#### 10.5.2.2. Etude de la spécificité de CupB6 par SPR

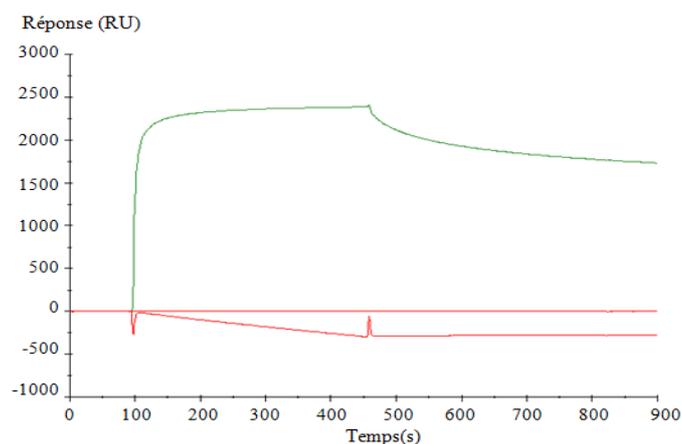
Des expériences de résonance plasmonique de surface (SPR) ont été entreprises pour vérifier si le Lewis b est le ligand naturel de CupB6/B4. L'avantage de cette technique est qu'elle consomme peu de protéine par rapport à l'ITC.

Dans un premier temps, une puce à sucre streptavidine/ Lewis b est réalisée. Sur le canal FC2 de la puce CM5 est fixé du PAA-Lewis b alors que le canal FC1, qui sert de référence, reste recouvert de streptavidine.

L'optimisation de la capture de streptavidine a permis de fixer une très grande quantité Lewis b (768 RU (response units) de PAA-Lewis b ont été fixés au lieu de 50 RU). La protéine BambL de *Burkholderia ambifaria* qui présente une forte affinité pour le fucose est utilisée comme témoin positif. Le sensorgramme obtenu présente un signal positif composé d'une phase d'association qui se termine par un plateau de saturation et d'une phase de dissociation pour le témoin positif (Figure 82). Le signal négatif obtenu pour CupB6/B4 indique que soit la chaperonne ou soit l'adhésine possède une affinité pour la streptavidine (288 RU) ou pour le dextran. L'interaction observée n'est pas spécifique puisque aucune condition de régénération comme 2M de NaCl ne permettent de décrocher le complexe.

La potentielle présence d'une hélice  $\alpha$  polyproline due à la répétition de 5 prolines au niveau du putatif linker de CupB6 laisse penser que le complexe CupB6/B4 se fixe à la streptavidine de la puce. En effet ce type d'hélice joue un rôle très important dans de nombreuses interactions protéine-protéine (Mansiaux et al., 2011). De plus lors de la purification par filtration sur gel Superdex 75 HiLoad 16/60 (GE healthcare), le complexe n'est pas retenu par la résine composée principalement de dextran.

Dans l'hypothèse que CupB6/B4 interagisse avec la streptavidine, il serait intéressant de réaliser des essais de saturation à la BSA afin de ne détecter uniquement les interactions avec le ligand.

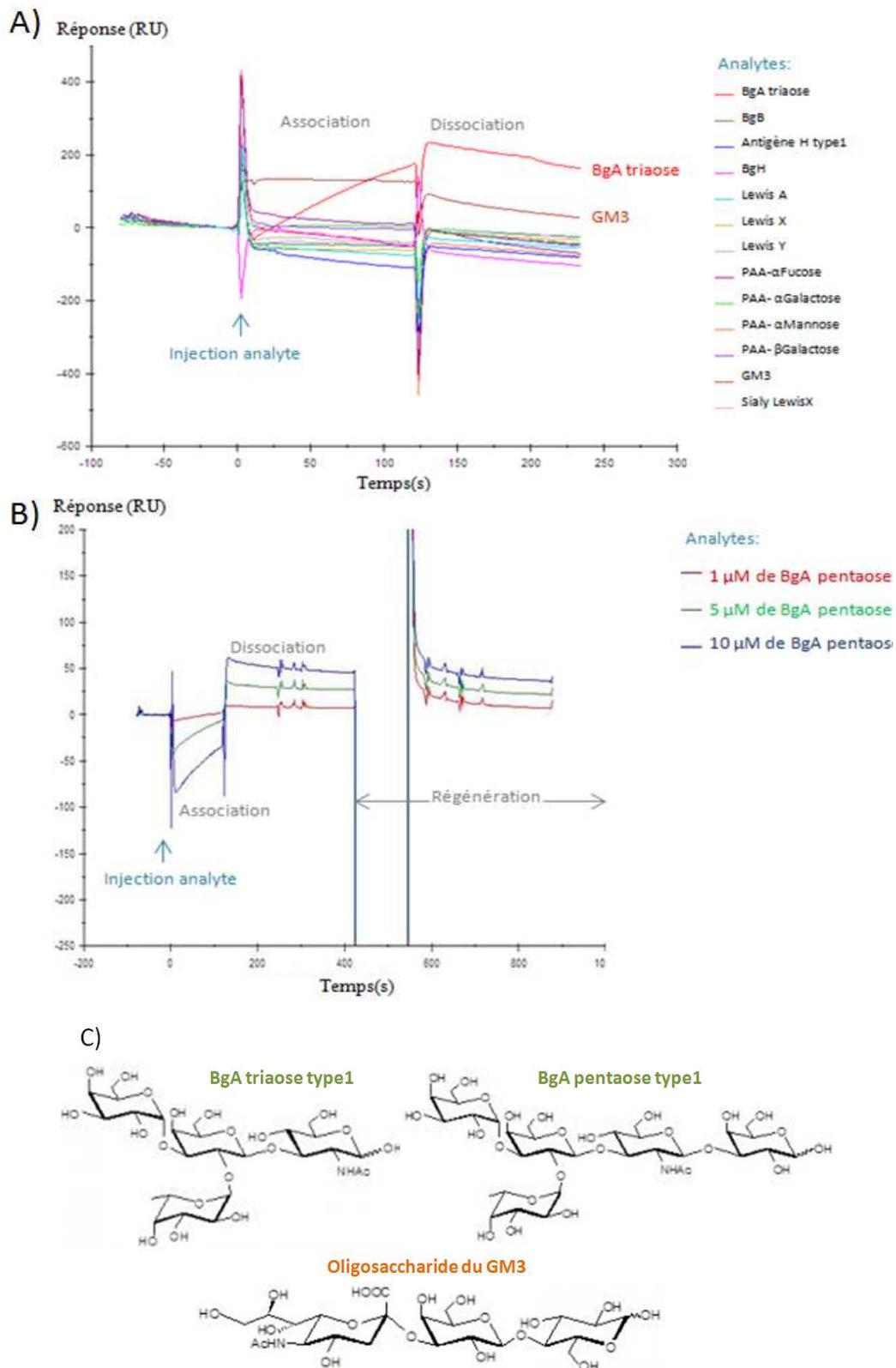


**Figure 82: Mesure de l'interaction entre CupB6/B4 et le Lewis b par SPR.** Le sensorgramme obtenu est issu de la soustraction de la cellule FC2 qui contient le Lewis b et FC1 qui contient la streptavidine. Le signal en vert correspond à la protéine témoin BambL (25 $\mu$ M) et le signal rouge à CupB6/B4 (27 $\mu$ M).

Suite aux interactions non spécifiques avec la streptavidine, il est clair que fixer le sucre à la puce n'est pas adéquat pour l'étude de la spécificité de CupB6/B4. L'autre stratégie consiste à utiliser la protéine comme ligand et le sucre comme analyte. Afin de percevoir un signal d'interaction ligand/analyte, plus de 10000 RU de CupB6/B4 doivent être capturés sur la puce. La présence de lysines dans le complexe (19 lysines pour l'adhésine CupB6 et 1 pour la chaperonne CupB4) rend possible la fixation de CupB6/B4 sur la puce CM5 par couplage amine. Par cette méthode également appelé couplage direct, plus de 10000 RU de CupB6/B4 ont été capturés sur la piste FC2. Parmi les différents analytes testés, une fixation de 200 RU de BgA triose (Blood group A antigen triose) et de 50 RU pour la partie oligosaccharidique du lipooligosaccharide GM3 (Ganglioside sugar 3'Sialyllactose) est observée sur le sensorgramme (Figure 83 A). Une autre expérience utilisant différentes concentrations du sucre BgA pentaose a permis de montrer que la fixation du sucre est dépendante de la concentration (Figure 83 B). Une fixation de 50 RU est observée à 10  $\mu$ M de BgA pentaose. La même expérience a été réalisée avec le Lewis b mais aucune fixation significative n'a été observée. La fixation du BgA pentaose est plus faible que celle du BgA triose, mais il est fort probable que le BgA triose se fixe également de manière concentration-dépendante au complexe. Ces résultats indiquent que CupB6/B4 pourrait présenter une affinité pour le BgA triose et une plus faible affinité pour le GM3 et le BgA pentaose.

Les problèmes majeurs liés à la méthode de couplage direct pour CupB6/B4 sont l'instabilité du complexe que l'on peut remarquer par la ligne de base qui ne se stabilise pas et l'étape de régénération. En effet bien que plusieurs protocoles de régénération ont été testés,

aucune régénération efficace n'a pu être obtenue durant l'expérience. (L'expérience de couplage direct de CupB6/B4 a été réalisée par le Dr Audfray Aymeric).

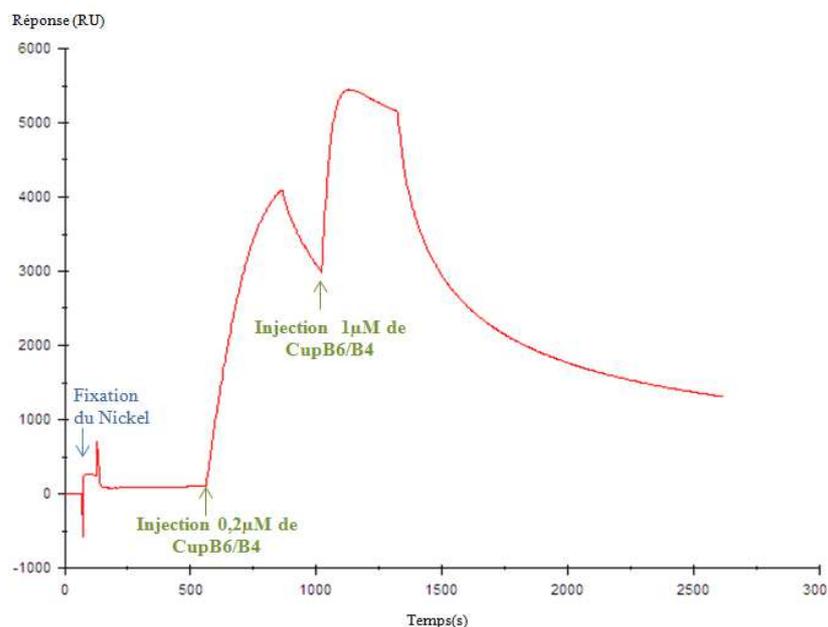


**Figure 83: Etude de l'affinité de CupB6/B4 par la méthode de couplage direct.** A) et B) Les sensorgrammes obtenus sont issus de la soustraction de la cellule FC2 où CupB6/B4 est fixé et de

FC1 qui contient de la streptavidine. Les courbes ont été soustraites par rapport à celle du tampon pour rétablir la ligne de base qui n'est pas stable. C) Structure des sucres qui interagissent avec CupB6/B4.

Le fait de ne pas pouvoir régénérer la puce lorsque le complexe est fixé par couplage direct sur CM5 nous a orienté vers une puce de type NTA. En effet le principal avantage de cette puce basée sur l'affinité ion nickel/étiquette hexahistidine permet une rapide régénération. La présence d'une étiquette hexahistidine à l'extrémité N-terminale de CupB6 permet l'utilisation d'une puce NTA. Un protocole a d'abord été mis en place avec la protéine RSL de *Ralstonia Solanacearum*.

Les différents essais de capture de CupB6/B4 sur la puce NTA n'ont pas été concluants. Différents paramètres tels que la variation du pH (de 7 à 8), la concentration du complexe, la concentration en sel, le temps d'injection de la protéine n'ont pas permis de fixer plus de protéine. En effet la grande instabilité du complexe limite la fixation à 1000 RU de CupB6/B4 (Figure 84), ce qui ne permet pas de visualiser une potentielle interaction avec l'analyte d'environ 670 Da. La constante décroissance de la ligne de base observée avec la puce NTA mais également par couplage direct sur CM5 peut être due au fait que le complexe agrège ce qui ne permet pas une fixation optimale.



**Figure 84: Essais de fixation du complexe CupB6/B4 sur la puce NTA.**

D'après les résultats obtenus par couplage direct sur CM5, l'hypothèse que CupB6 possède une affinité avec le BgA triose, GM3 et le BgA pentaose doit être confirmée par l'obtention d'un  $K_d$ . Une étude cinétique a été réalisée avec l'analyte BgA triose utilisant la méthode par couplage direct sur CM5. Durant l'expérience, cinq concentrations croissantes de BgA triose

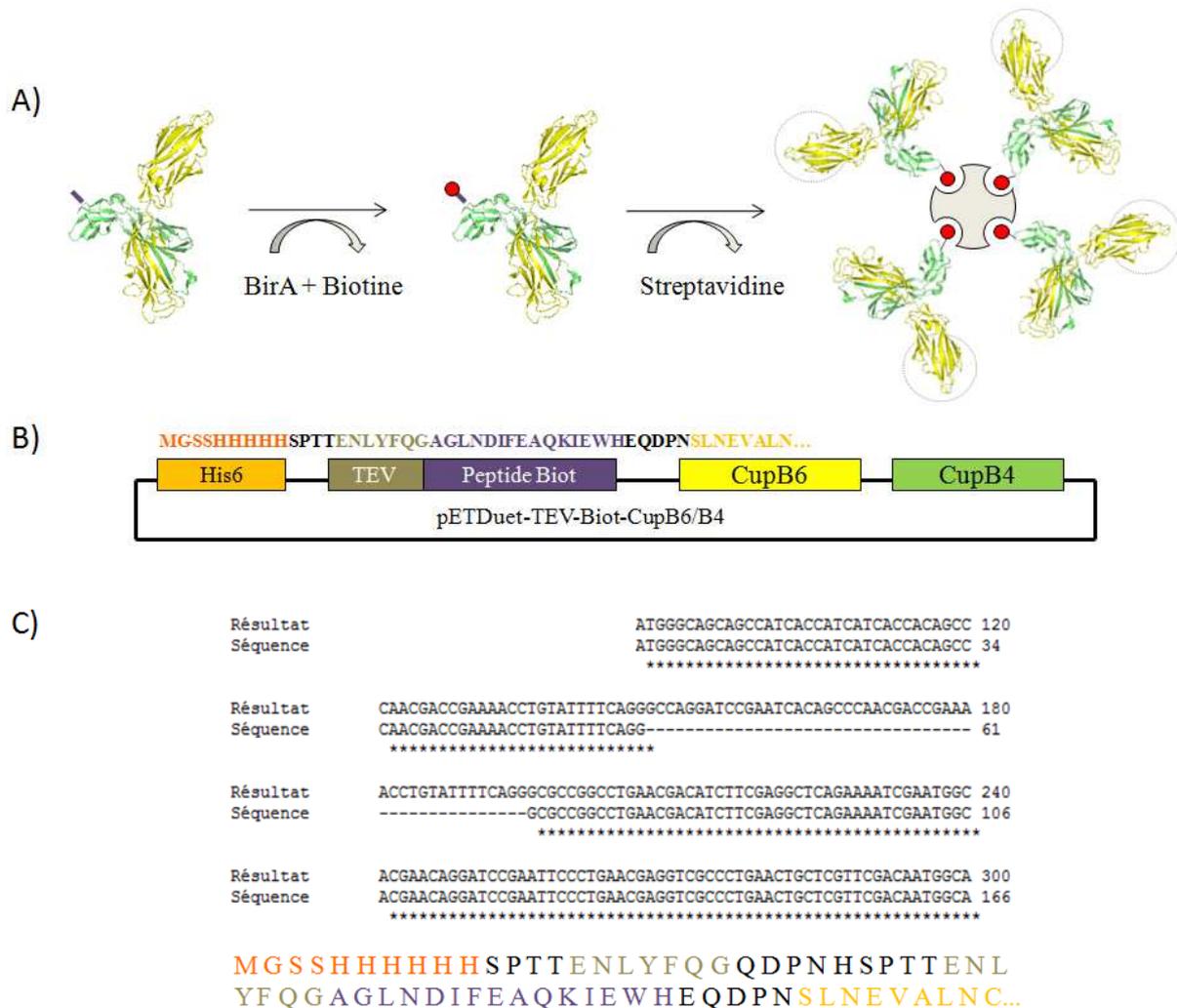
sont injectées à la suite après la phase de dissociation. Les résultats ne sont cependant pas interprétables en raison de l'instabilité de la ligne de base. Une mesure par ITC en excès de ligand permettrait d'infirmer si CupB6/B4 présente une affinité pour BgA triose. Cette expérience n'a pas été faite par manque de ligand BgA triose. Elle a cependant été réalisée avec le GM3. Les résultats obtenus par ITC ne montrent aucune titration entre CupB6/B4 et le GM3. Il est important de noter qu'après l'expérience, le complexe CupB6/B4 était entièrement précipité. La mesure réalisée par ITC indique qu'il n'y a pas d'interaction entre CupB6/B4 et la partie saccharidique du GM3, sous réserve que le complexe n'ait pas précipité avant la première injection.

#### 10.5.2.3. Création d'un tétramère artificiel de CupB6/B4

La difficulté d'identifier la spécificité de CupB6 pour les oligosaccharides peut être liée à sa monovalence. En effet la faible affinité que les lectines présentent pour les sucres est en général compensée par leur multivalence c'est-à-dire la présence de plusieurs sites de reconnaissance. La création d'un tétramère artificiel du complexe CupB6/B4 serait un moyen d'augmenter l'avidité et donc d'augmenter l'intensité du signal d'interaction entre un ligand potentiel et l'adhésine CupB6. Cette stratégie avait été utilisée dans l'étude de plusieurs lectines SIGNs (specific intercellular adhesion molecule 3 (ICAM-3)-grabbing nonintegrin) de souris pour obtenir de meilleurs résultats avec les puces à sucres (Powlesland et al., 2006). Le couple biotine/(strept)avidine constitue la liaison non covalente la plus forte connue ( $K_d = 10^{-15}$  M) (Miller and Tausig, 1964). Un tétramère de streptavidine fixe quatre molécules de biotine. Ces particularités peuvent être utilisées pour créer un tétramère artificiel de CupB6/B4 suite à sa biotinylation spécifique (Figure 85 A). Pour cela, le peptide de reconnaissance de la biotine ligase BirA (AGLNDIFEAQKIEWH) a été inséré par PCR en N-terminal de CupB6 après l'étiquette hexahistidine. Il permettra l'attachement covalent de la biotine sur le résidu lysine de ce peptide par BirA. Afin de pouvoir enlever l'étiquette hexahistidine le peptide ENLYFQG qui est reconnu par la protéase TEV (tobacco etch virus) a également été inséré par PCR. Le plasmide pETDuet-TEV-Biot-CupB6/B4 a ainsi été obtenu (Figure 85 B).

Les résultats du séquençage indiquent que la séquence qui correspond au peptide reconnu par la TEV est en double (Figure 85 C). Une erreur dans le design d'amorce en est la cause (cf 5.1.2). La présence de deux peptides reconnus par la TEV ne pose en soi pas de problème puisque après coupure qui s'effectue entre le résidu Q et G, ces deux peptides sont enlevés. De plus, la séquence du peptide reconnu par BirA a été correctement insérée. Aucune

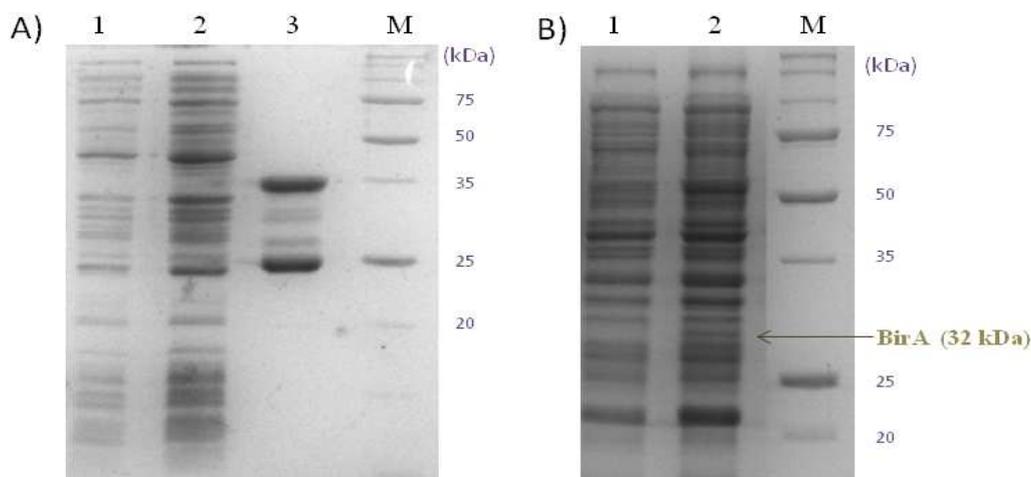
mutation n'a été observée et le cadre de lecture de l'adhésine CupB6 est bien respecté (Figure 85 C).



**Figure 85: Création d'un tétramère artificiel du complexe CupB6/B4 utilisant l'interaction biotine-streptavidine.** A) Représentation des différentes étapes pour l'obtention du tétramère de CupB6/B4. Le complexe FimH/FimC d'*E. coli* a été utilisé pour une représentation qui se rapproche le plus de la réalité. La chaperonne FimC est en vert et l'adhésine FimH en jaune. Le site de reconnaissance aux sucres de l'adhésine est représenté par un cercle. B) Représentation du plasmide pETDuet-TEV-Biot-CupB6/B4 à obtenir. C) Alignement des séquences nucléiques entre le résultat du séquençage du plasmide pETDuet-TEV-Biot-CupB6/B4 et la séquence attendue ainsi que la séquence protéique qui correspond au résultat du séquençage.

Des premiers essais d'expression ont été réalisés avec la souche d'*E. coli* K12 AVB100 (Avidity) qui contient le gène *birA* dans son chromosome et permet de surexprimer la ligase en même temps que le complexe et de biotinyler *in vivo*. Aucune production du complexe n'a été observée. La présence de plusieurs codons rares dans gène *cupB6* pouvait expliquer l'absence d'expression de CupB6/B4 car la souche AVB100 ne permet pas de les produire.

Afin d'apporter ces codons rares au cours de l'expression, le plasmide pLysSRARE2 a donc été co-transformé dans la souche AVB100. Là encore aucune expression du complexe n'a été observée (Figure 86). La deuxième stratégie consistait à co-exprimer la protéine BirA et le complexe dans la souche Rosetta2. Le gène *birA* a été cloné dans le vecteur pCDFDuet qui possède une origine de répllication différente du plasmide pETDuet-TEV-Biot-CupB6/B4. Les tests d'expression ont permis d'identifier la production de BirA mais pas du complexe CupB6/B4 (Figure 86). Il est possible que l'ajout de 38 acides aminés en N-terminal de l'adhésine CupB6 gêne son repliement et donc son expression. Dans les études qui utilisent cette stratégie, le peptide reconnu par BirA est inséré en C-terminal de la protéine. Avec CupB6, l'insertion en C-terminal est impossible car elle générerait la formation du complexe adhésine/chaperonne. Mais l'insertion du peptide reconnu par BirA en C-terminal de la chaperonne pourrait être un moyen de palier à ce problème. La formation d'un tétramère où la biotine est présente en C-terminal de la chaperonne permettrait également une meilleure orientation du site de reconnaissance aux sucres de l'adhésine (Figure 86).



**Figure 86: Tests d'expression de CupB6/B4 avec le plasmide pETDuet-TEV-Biot-CupB6/B4.** A) Analyse par SDS-PAGE du test d'expression de CupB6/B4 utilisant la souche AVB100 pLysSRARE2 et B) la souche Rosetta2 co-transformée avec le plasmide pCDFDuet-BirA et pETDuet-TEV-Biot-CupB6/B4. Les pistes 1 et 2 correspondent aux extraits totaux de la culture, respectivement avant et après induction. Le complexe CupB6/B4 sans le peptide est utilisé comme marqueur en piste 3. La taille attendue de l'adhésine CupB6 qui comporte l'étiquette His6 et les deux peptides est de 41kDa.

### 10.5.3. Cristallogénèse du complexe CupB6/B4

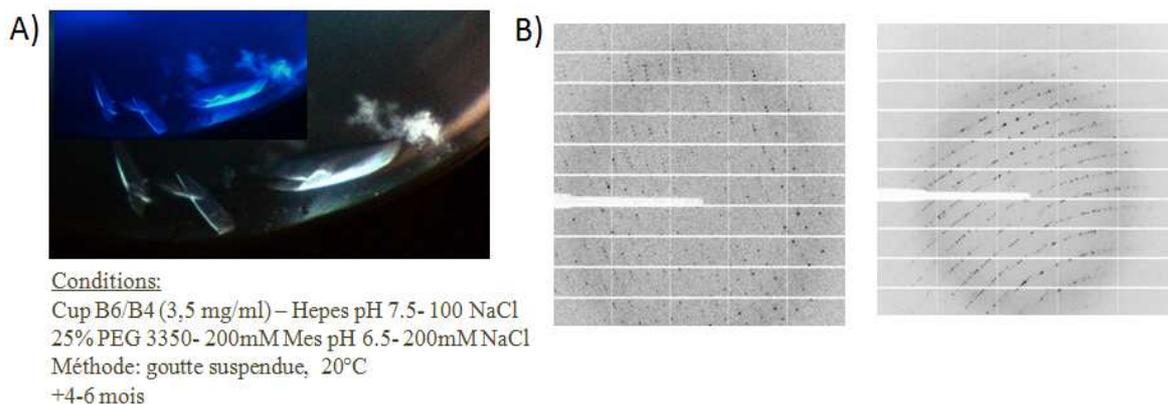
Le complexe CupB6/B4 a été envoyé à la Plateforme de cristallisation robotisée des protéines de l'EMBL, Grenoble. Les kits suivants ont été utilisés : PACT et JCSG (Qiagen/ Nextal) et les plate-5, 4, 1 et 2 (Hampton Research). Dans le tampon 20 mM Tris-HCl pH 7,5- 100 mM

NaCl et avec le complexe à une concentration de 5 mg.ml<sup>-1</sup>, aucun cristal n'a été obtenu. Toutes ces conditions ont également été testées avec le complexe CupB6/B4 méthylé par du diméthylamine-borane. L'addition de groupements méthyl sur l'amine de la chaîne latérale des lysines peut favoriser la cristallisation par « immobilisation » de ces chaînes latérales (Walter et al., 2006). Cette stratégie n'a pas permis d'obtenir des cristaux.

Les kits JCSG+suite (Qiagen), Crystal Strategy Screen 1 et 2, Morpheus et MIDAS (Molecular Dimensions Ltd) ont été testés au laboratoire avec différentes concentrations en complexe CupB6/B4.

Des cristaux ont été obtenus entre 4 à 6 mois dans 25 % PEG3350- 200 mM MES pH 6,5- 200 mM NaCl pour le complexe à une concentration de 3,5 mg.ml<sup>-1</sup> dans de l'HEPES pH 7,5- 100 mM de NaCl (Figure 87). Les cristaux ont été congelés avec 20 % de glycérol puis testés à l'ESRF sur la ligne BM30A. Des données ont pu être collectées pour un cristal à une résolution de 2,3 Å et appartiennent au groupe d'espace orthorhombique P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub> avec une maille de a=31Å, b=51Å, c= 128Å,  $\alpha=\beta=\gamma=90^\circ$ . Après analyse, il s'avère que l'unité asymétrique ne peut contenir une protéine de poids moléculaire supérieur à 25 kDa. Deux hypothèses peuvent être envisagées, soit le complexe se dissocie et la chaperonne cristallise ou soit le complexe est lysé et un fragment de dégradation cristallise.

Suite à cette observation, une étude de protéolyse ménagée (cf 10.6.3.3) a été entreprise. D'après les résultats, les cristaux obtenus pourraient correspondre à un produit de lyse de l'adhésine CupB6 ou bien à celui de la chaperonne. Des essais de remplacement moléculaire utilisant la structure de CupB2 qui possède 32 % d'identité avec CupB4 n'ont pas aboutis. Dans l'hypothèse que les cristaux pouvaient contenir le domaine piline de CupB6, des essais de remplacement moléculaire avec les domaines pilines des adhésines FimH et PagG ont aussi été tentés sans succès. Il n'existe aucune modèle pour le domaine adhésine à ce jour. Ces cristaux étant obtenus entre 4 à 6 mois suite probablement à une protéolyse, il a été impossible de les reproduire et de déterminer quelle partie avait cristallisé. Des expériences de protéolyse limitée ont été menées sur le complexe et seront décrites ultérieurement.

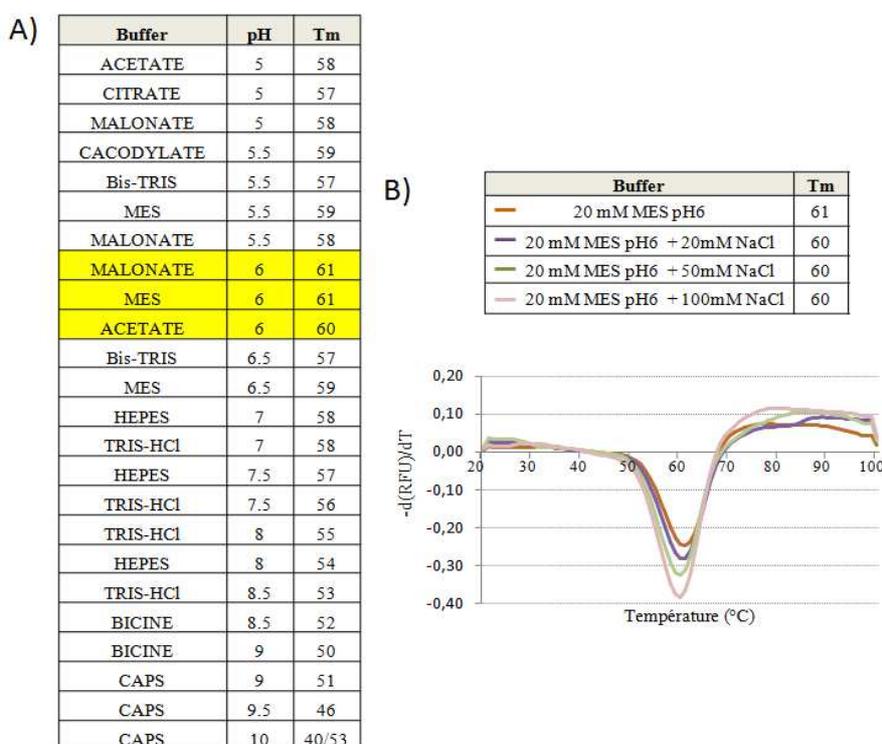


**Figure 87: A) Cristaux et B) clichés de diffractions à 2,3Å de CupB4.**

Les difficultés rencontrées lors des tests de cristallo-genèse peuvent être dues au fait que le complexe ne soit pas dans des conditions optimales.

### 10.5.3.1. Optimisation des conditions de cristallisation par TSA

Des expériences de déstructuration thermique des protéines (TSA) ont été entreprises pour définir un tampon plus optimal pour le complexe CupB6/B4. Plusieurs tampons à une concentration de 100 mM dont le pH varie de 5 à 10 ont été testés. Il apparaît que le complexe CupB6/B4 présente une meilleure stabilité ( $T_m$  de 61) à pH 6 (Figure 88 A). La présence du NaCl à différentes concentrations n'a pas un effet sur la stabilité de CupB6/B4 (Figure 88 B).



**Figure 88: Optimisation du tampon de CupB6/B4 par TSA.** A) Tableau regroupant les  $T_m$  de CupB6/B4 déterminés par TSA pour différents tampons. Les conditions où le complexe CupB6/B4 est

le plus stable sont surlignées en jaune. B) Etude de l'influence du sel sur la stabilité du complexe CupB6/B4.

Des expériences de TSA sur plusieurs additifs pouvant aider à la cristallisation du complexe CupB6/B4 ont également été réalisées (Figure 89).

Pour certains additifs tels que le TCEP (tris(2-carboxyl)phosphine), le polyvinylpyrrolidone k15, ou le Yttrium (III) chloride hexahydrate, deux T<sub>m</sub> sont obtenus. Ce qui peut s'expliquer par le fait que ces additifs déstabilisent ou dissocient le complexe CupB6/B4. Ces composés ne doivent pas être utilisés dans les essais de cristallogénèse.

Mesurer la stabilité thermique d'une protéine peut également permettre l'identification d'un ligand. En effet la formation d'un complexe protéine/ligand peut en augmentant sa stabilité augmenter le T<sub>m</sub> de la protéine (Epps et al., 2001). Parmi tous les sucres testés aucune variation du T<sub>m</sub> de CupB6/B4 n'a été observée excepté pour l'acide neuraminique qui en acidifiant le pH diminue le T<sub>m</sub> du complexe. Ces résultats indiquent qu'il n'a pas été possible d'identifier un ligand pour CupB6 par TSA. Il est possible que la formation d'un potentiel complexe CupB6/sucre n'augmente pas la stabilité de CupB6/B4. En effet la valeur du T<sub>m</sub> de CupB6/B4 est élevée et la monovalence du CupB6 peut conduire à une faible affinité pour les sucres. De plus, seuls des monosaccharides ont été testés et il est possible que le ligand de CupB6 soit un oligosaccharide.

Cette étude a aussi permis de cribler différents additifs (T<sub>m</sub> varie de 60 à 61) qui ne déstabilisent pas le complexe et pourraient être ajoutés dans des essais de cristallogénèse de CupB6/B4.

Additifs (Divers)	Concentration	T <sub>m</sub>
Phenol	10 mM	60
Dimethyl sulfoxide	0,03	60
Sodium bromide	10 mM	60
6- Aminohexanoic acid	0,03	60
1,5-Diaminopentane dihydrochloride	0,03	55
1,6-Diaminohexane	0,03	42
1,8-Diaminooctane	0,03	61
Glycyl-glycyl-glycine	30 mM	61
Spermidine	10 mM	50
Spermidine tetrahydrochloride	10 mM	58
Hexamine cobalt (III) chloride	10 mM	58
Sarcosine	10 mM	
β-Nicotinamide adenine dinucleotide hydrate	10 mM	55
Adenosine-5'-triphosphate disodium salt hydrate	10 mM	46/61

Additifs (Divers)	Concentration	T <sub>m</sub>
TCEP hydrochloride	10 mM	21/25
GSH (L-Glutathione reduced)	1 mM	60
Ethylenediaminetetraacetic acid disodium salt dihydrate	10 mM	60
EDTA	10 mM	61
Ethylene glycol	0,03	60
Glycerol	0,03	61
NDSB-195	300mM	59
NDSB-201	200mM	58
NDSB-211	200mM	61
NDSB-221	200mM	59
NDSB-256	200mM	59
CYMAL-7	0.015mM	61
Benzamidine hydrochloride	0,02	52

Additifs (acides aminés)	Concentration	Tm
L-Proline	10 mM	61
L-Alanine	10 mM	61
L-Glutamine	10 mM	61
L-Arginine	10 mM	53
L-Lysine	10 mM	61
L-Glycine	10 mM	61
L-Serine	10 mM	61

Additifs (Multivalents)	Concentration	Tm
Barium chloride dihydrate	10 mM	60
Cadmium chloride hydrate	10 mM	53
Calcium chloride dihydrate	10 mM	59
Cobalt (II)chloride hexahydrate	10 mM	56
Copper (II) choride dihydrate	10 mM	47
Magnesium choride hexahydrate	10 mM	59
Manganese (II) choride tetrahydrate	10 mM	58
Strontium chloride hexahydrate	10 mM	59
Yttrium (III) chloride hexahydrate	10 mM	43/53
Zinc chloride	10 mM	52
Nickel (II) chloride hexahydrate	10 mM	56
Chromium (III) chloride hexahydrate	10 mM	41
Praseodymium (III) acetate hydrate	10 mM	47

Additifs (polymères)	Concentration	Tm
Polyvinylpyrrolidone K15	0.5%	33/60
Detran sulfate sodium salt (Mr 5,000)	0,03	49
Pentaerythritol ethoxylate	0,04	60
Polyethylene glycol 3,350	0,01	61

Additifs (carbohydrates)	Concentration	Tm
D-(+)-Glucose monohydrate	0,03	61
Sucrose	0,03	61
Xylitol	0,03	61
D-Sorbitol	0,03	61
myo-Inositol	1.2%	61
D-(+)-Trehalose dihydrate	0,03	61
D-(+)-Galactose	0,03	61
N-acetyl-D-glucosamine	10 mM	61
L-Fucose	10 mM	61
D-Glucose	10 mM	61
N-acetyl-D-galactosamine	10 mM	61
L-Arabinose	10 mM	61
Xylose	10 mM	61
D-Mannose	10 mM	61
Neuraminic acid	10 mM	42
L-Rhamnose	10 mM	61

Additifs (sels)	Concentration	Tm
Ammonium sulfate	100 mM	58
Potassium chloride	100 mM	59
Lithium chloride	100 mM	59
Sodium chloride	200 mM	60
Sodium iodide	100 mM	53
Sodium thiocyanate	200 mM	59
Potassium sodium tartrate tetrahydrate	100 mM	56
Sodium citrate tribasic dihydrate	100 mM	61
Cesium chloride	100 mM	59
Sodium malonate pH 7	100 mM	58

Figure 89: Détermination du Tm pour le complexe CupB6/B4 en présence de certains additifs.

### 10.5.3.2. Optimisation des conditions de cristallisations

Suite à ces expériences de thermostabilité, de nouveaux essais de cristallisation ont été menés pour le complexe CupB6/B4 dans un tampon plus optimal sur la Plateforme de cristallisation robotisée des protéines de l'EMBL. Les kits suivant ont été utilisés : PACT et JCSG (Qiagen/Nextal), Wizard I+II (Emerald BioSystems) et les plate-4, 3, et 2 (Hampton Research). Aucun cristal n'a été obtenu à une concentration de 4 mg.ml<sup>-1</sup> dans 20 mM MES pH 6, 100 mM NaCl et à une concentration de 3,5 mg.ml<sup>-1</sup> dans 20 mM MES pH 6. Pour toutes les plaques (sauf plate-3) environ 50 % de gouttes claires et des gouttes contenant des précipitations sont obtenues. La présence de 70 % de gouttes claires pour la plaque plate-3 (Hampton Research), indique que la concentration de CupB6/B4 n'est pas assez élevée et il serait intéressant de tester avec une concentration de 6-5 mg.ml<sup>-1</sup>.

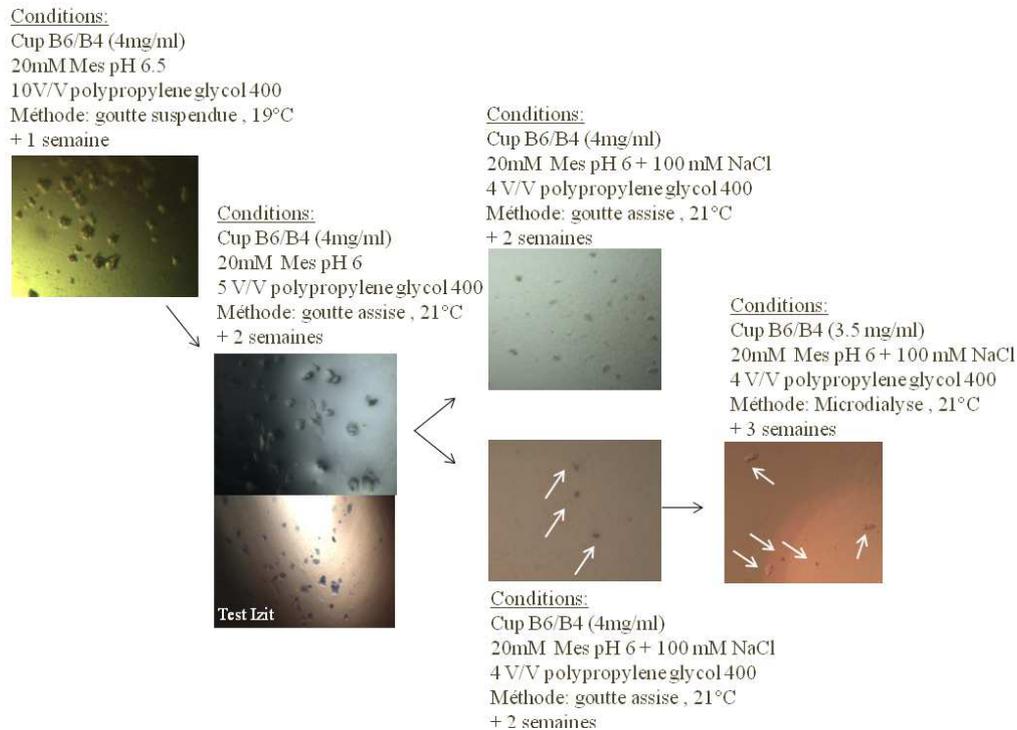
Dans un premiers temps, les essais de cristallisation étaient réalisés avec une concentration de 5 mg.ml<sup>-1</sup> de CupB6/B4. Mais dans plus de 60 % des conditions, une rapide et forte précipitation du complexe était obtenue. Suite à ces observations les essais de cristallisation ont été réalisés avec des concentrations plus basses comme 4,5- 4- 3,5- 3 mg.ml<sup>-1</sup>.

Au laboratoire, différents kits commerciaux ont également été testés par la méthode de goutte suspendue: CSS-I et CSSII (Clear Strategy Screen), Morpheus et Midas (Molecular Dimension Ltd) ; Kit Crystal Screen 1 et 2 (Hampton Research). Avec le kit Midas idéal pour les complexes de protéines, des microcristaux ont été obtenus dans 10 % v/v de polypropylène glycol 400 (Figure 90). A partir de ces microcristaux, la méthode de l'ensemencement a été réalisée mais aucun cristal n'a été obtenu. Cette technique permet de favoriser la croissance de quelques grands cristaux au lieu d'une multitude de petits.

Une phase d'optimisation des conditions de cristallisation a été entreprise afin d'obtenir des cristaux plus gros. Pour ce faire, différentes concentrations de l'agent précipitant, le polypropylène glycol 400, et du complexe CupB6/B4 ont été testées avec ou sans la présence de différents sels par la méthode de la goutte suspendue. Afin de ralentir la diffusion de la vapeur, une goutte d'huile a été placée entre le réservoir et la goutte pendue. Seuls des microcristaux ont été obtenus avec 5 % v/v de polypropylène glycol 400 à 3,5 mg.ml<sup>-1</sup> de CupB6/B4 et 100 mM de chlorure de lithium ou 100 mM d'ammonium sulfate.

Différentes températures ont également été testées sur la machine TG40 (Molecular Dimensions Ltd) qui utilise la méthode de goutte assise. Malgré des problèmes de condensation, des cristaux sont apparus uniquement à 21°C dans 5 % v/v de polypropylène glycol 400 et à 4 mg.ml<sup>-1</sup> de CupB6/B4 après une semaine. La coloration en bleu des cristaux après le test Izit (Hampton Research) a permis de confirmer qu'il s'agissait bien de cristaux de protéines et non de sel (Figure 90). Un test de diffraction sur ces cristaux n'a montré aucune diffraction qui peut s'expliquer par la très petite taille du cristal pêché. Les cristaux obtenus étant superposés les uns aux autres, d'autres essais ont été réalisés utilisant la méthode de la goutte assise à 21°C et avec 4 mg.ml<sup>-1</sup> de CupB6/B4 dans 20 mM MES pH 6 -100 mM NaCl. Des microcristaux ont été obtenus avec 5- 4 -3 % v/v de polypropylène glycol 400 (Figure 90). Tous les additifs criblés par TSA qui n'influent pas le Tm de CupB6/B4 (61-60) ont alors été ajoutés aux 5 % v/v de polypropylène glycol 400. L'effet du pH allant de 2,2 à 11 a également été testé. Parmi toutes les conditions testées, des microcristaux ont été obtenus avec l'additif glycyl-glycyl-glycine (Figure 90). Un premier essai utilisant la microdialyse avec 4 % v/v de polypropylène glycol 400 et 3,5 mg.ml<sup>-1</sup> de CupB6/B4 dans 20 mM MES pH 6 – 100 mM NaCl a permis d'obtenir des cristaux plus gros. D'autres tests utilisant la microdialyse doivent être effectués avec l'ajout de la glycyl-glycyl-glycine et d'un sel. Il serait également intéressant de tester d'autres tampons de CupB6/B4 à pH 6 comme le malonate, l'imidazole, l'acide succinique, l'ADA (N-(2-Acetamide)-iminodiacetic acid) et à pH 6,5 le MOPS (3-(N-morpholino)propanesulfonic

acid) qui n'ont pas été testés. Une autre étape d'optimisation serait de réaliser les essais de cristallogenèse en capillaire dans un gel d'agarose. En effet cette technique de cristallisation par contre-diffusion en microgravité permet de réduire le défaut d'empilement grâce à une croissance plus lente des cristaux (Gavira and García-Ruiz, 2002).



**Figure 90: Optimisation des conditions de cristallisation de CupB6/B4.**

Des séparations de phases ont aussi été obtenues dans certaines conditions (Figure 91). Une variation de certains paramètres de ces dernières comme la concentration de l'agent précipitant, du pH, la concentration de sel pourrait peut-être permettre d'obtenir des cristaux.

Observation de la goutte	Conditions de cristallisation	Observation de la goutte	Conditions de cristallisation
	- 35% Polypropylene glycol 400 - 0.1M K/Na phosphate - [CupB6/B4]= 4mg/ml - goutte pendue		- 20% Polyacrylate 2100 - 0,2M sodium chloride - 0,1M Bicine pH9
	- 35% Sokalan CP42 - 0.1M Bis-tris NaOH - [CupB6/B4]= 4mg/ml - goutte pendue		- 20% Glascol W13 - 0,2M sodium sulfate - 0,1M HEPES-NaOH pH7,5
	- 5% Polyvinyl alcohol typell - 0,2M potassium acetate - 0,1M HEPES-NaOH pH7	- [CupB6/B4]= 4mg/ml - goutte pendue	- 25% Jeffamine D2000 - 0,2M Imidazole pH7
- [CupB6/B4]= 4mg/ml - goutte pendue	- 8% Polyvinyl alcohol typell - 0,2M sodium chloride - 0,1M Bicine pH9		

**Figure 91: Différentes conditions où une séparation de phases est observée pour CupB6/B4.**

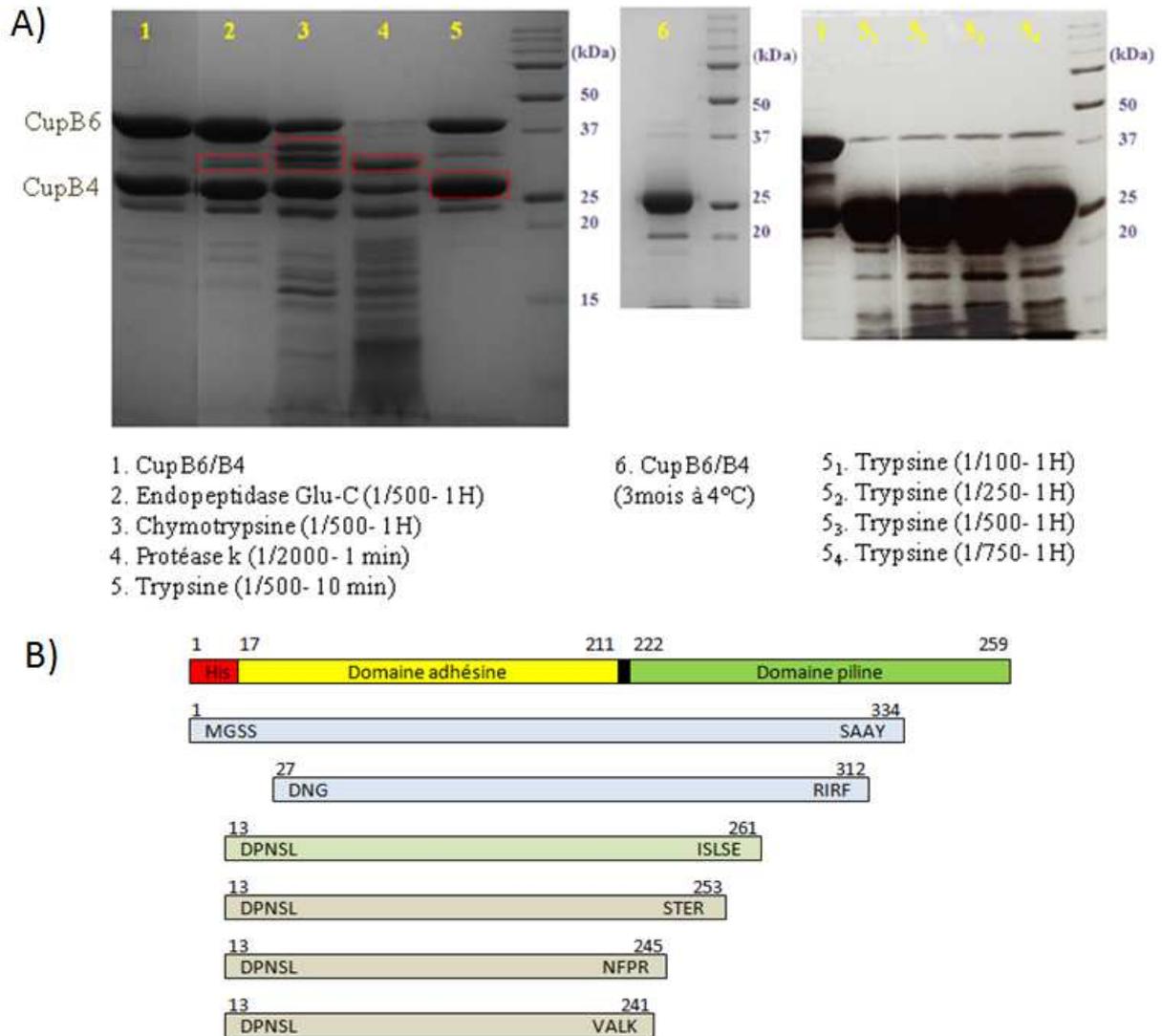
### 10.5.3.3. La protéolyse ménagée de CupB6/B4

La présence de régions désordonnées et flexibles au sein du complexe CupB6/B4 peut être la cause des difficultés rencontrées à obtenir des cristaux. Des expériences de protéolyse ménagée ont été entreprises afin de déterminer si un complexe plus stable pouvait être obtenu. Plusieurs protéases ont été testées et certaines bandes de protéolyse ont été analysées par la plateforme de spectrométrie de masse à l'IBS, Grenoble.

La digestion avec de la chymotrypsine au 1/500<sup>ème</sup> pendant 60 min permet d'obtenir deux bandes proches de la taille de CupB6. Après analyse, la bande de 35,3 kDa obtenue sur gel après digestion correspondrait à CupB6 1-334 et la bande de 30,3 kDa à CupB6 27-312 (Figure 92). Il ne paraît cependant pas judicieux d'enlever les 25 ou 47 acides aminés en C-terminal de CupB6 dans de nouvelles constructions plasmidiques au vue des différentes structures des pilines Pap et Fim d'*E. coli*. En effet ils sont impliqués dans la formation du sillon hydrophobe (le brin  $\beta$  F). Sans la stabilisation de la chaperonne, il est fort possible que CupB6 1-334 ou CupB6 27-312 soient insolubles ou agrègent.

Les résultats de protéolyse ménagée obtenus avec l'endopeptidase Glu-spécifique et la trypsine sont les plus intéressants. En effet, malgré différentes concentrations testées pendant 1h, il semble qu'une seule bande correspondant à l'adhésine CupB6 soit obtenue (Figure 92). L'analyse par spectrométrie de masse des bandes générées par l'endopeptidase Glu-C révèle que la bande de 26,3 kDa également retrouvée avec la chymotrypsine et la protéase K pourrait correspondre à CupB6 13-261.

Les résultats obtenus avec la trypsine n'ont pas été analysés mais le profil de digestion ressemble à celui obtenu pour le complexe CupB6/B4 après stockage au frigidaire. En effet, malgré l'ajout d'inhibiteurs de protéases durant la purification de CupB6/B4, une lyse du complexe a été observée après 3 mois à 4°C (Figure 92). L'analyse par séquençage N-terminal a permis d'identifier différents fragments correspondant à la lyse de la chaperonne et de l'adhésine. En ce qui concerne l'adhésine CupB6, seule l'étiquette hexahistidine est lysée en N-terminal. D'après la taille du peptide estimée entre 24-25 kDa et les sites de coupure spécifiques de la trypsine, le fragment de lyse pourrait correspondre à CupB6 13-241 (24 kDa), à CupB6 13-245 (24,5 kDa) ou à CupB6 13-253 (25,4 kDa). Des expériences de spectrométrie de masse devraient être menées afin de déterminer où a eu lieu la lyse. Faute de matériel, elles n'ont pas pu être réalisées en même temps que le séquençage N-terminal.



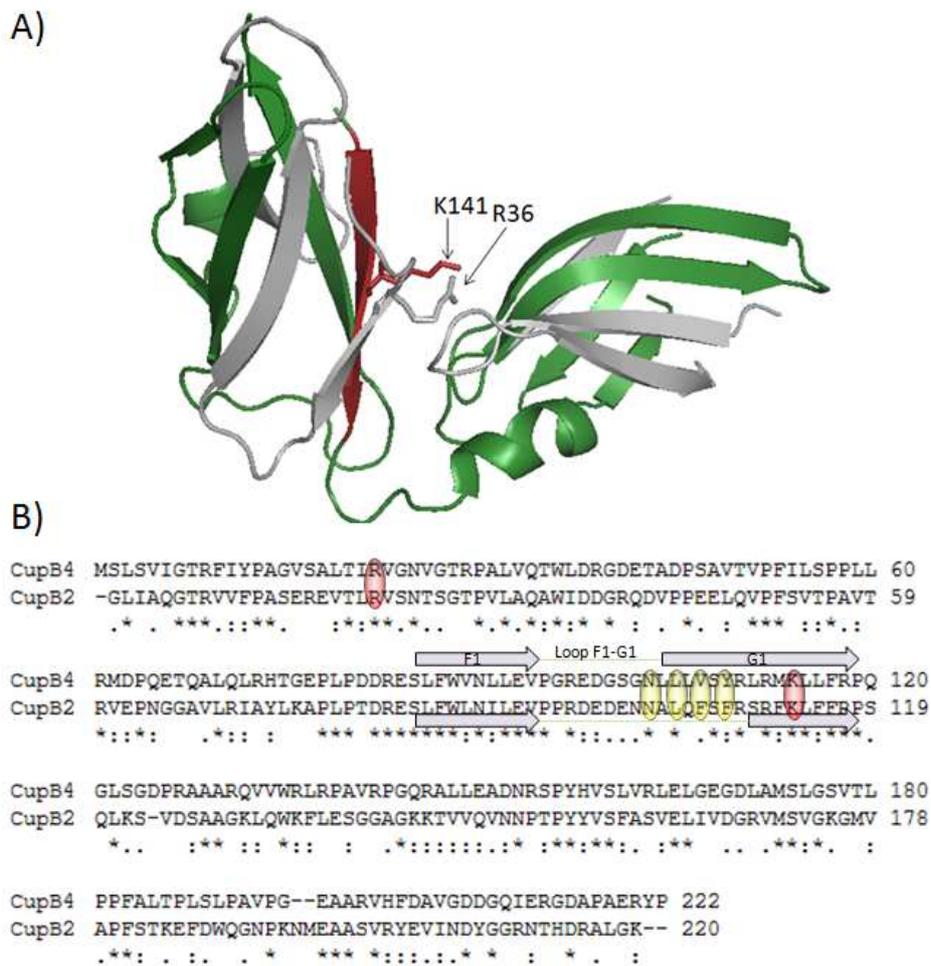
**Figure 92: Protéolyse ménagée de CupB6/B4 par différentes protéases.** A) Analyse sur gel SDS-PAGE des profils de digestion de CupB6/B4 par différentes protéases. B) Représentation des fragments de lyse de CupB6 avec la protéase chymotrypsine (bleu), la protéase Glu-C (vert) et après 3 mois à 4°C (gris).

Les trois autres bandes observées de 20- 18- 15,5 kDa sur gel SDS correspondraient respectivement à CupB4 45-120, CupB4 25-188, CupB4 45-185. La résolution de la chaperonne CupB2 permet de visualiser les fragments qui sont lysés (Figure 93).

Les résultats de protéolyse ménagée mettent en évidence une bande de 26,5 kDa ou de 24- 25 kDa qui correspondrait au domaine adhésine et une partie du domaine piline de CupB6. La taille de ces fragments est compatible avec la taille du produit de lyse estimée qui a cristallisé (cf 10.6.3). La trypsine et l'endopeptidase Glu-C étant les deux protéases qui permettent d'obtenir ces fragments, il serait intéressant de réaliser de la protéolyse *in situ* afin de reproduire les cristaux obtenus (Bai et al., 2007). Il serait également intéressant de réaliser les

constructions suivantes : CupB6 13-241 ; CupB6 13-245 et CupB6 13-253. D'après les prédictions des domaines, les coupures ont lieu en N-terminal du domaine piline. Il est possible que l'obtention de protéines solubles et monodisperses soit difficile.

Afin de limiter cette lyse, les conditions de cristallisation devraient être réalisées à 4°C. De plus l'ajout d'EDTA (qui ne modifie pas le Tm de CupB6/B4) dans le tampon de filtration sur gel permettrait d'inhiber les métalloprotéases. L'ajout de 50 mM d'acides aminés chargés dans le tampon pourrait également protéger le complexe contre la protéolyse (Golovanov et al., 2004).



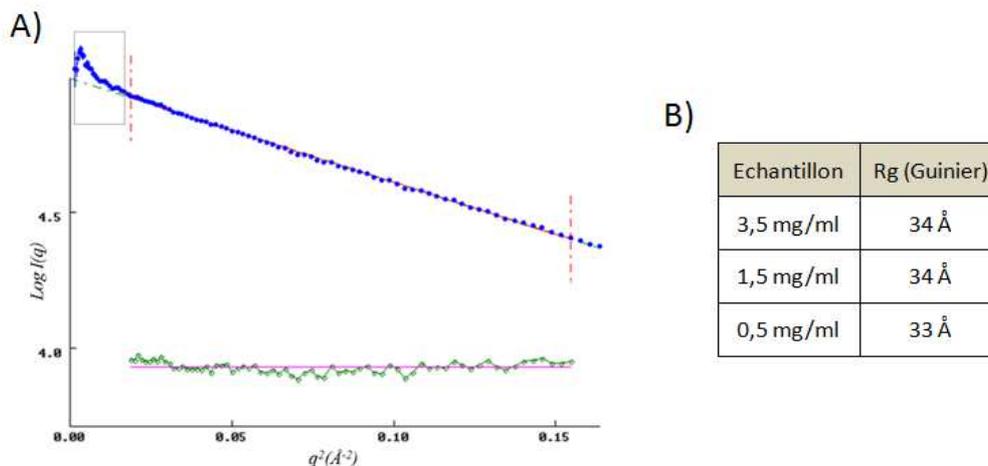
**Figure 93:** A) Représentation sur la structure CupB2 des fragments dégradés lors de la lyse de CupB4. En gris sont représentés les parties de CupB4 qui sont lysées. Le brin  $\beta$  G qui permet de stabiliser le domaine piline figure en rouge. Les deux acides aminés qui jouent un rôle important dans l'interaction avec la chaperonne et le domaine piline sont représentés en bâton. B) Alignement de la séquence de CupB4 et de CupB2. Le brin  $\beta$  G1 et F1 sont représentés par une flèche grise. Les résidus qui interagissent avec le sillon hydrophobe de la piline sont surlignés en jaune. Les résidus arginine et lysine très conservés chez les chaperonnes CU sont surlignés en rouge.

## 10.5.4. Caractérisation du complexe CupB4/CupB6 par SAXS

### 10.5.4.1. Evaluation et traitement des données collectées

Des études de diffusion des rayons X aux petits angles ont été entreprises afin de mieux caractériser le complexe CupB6/B4.

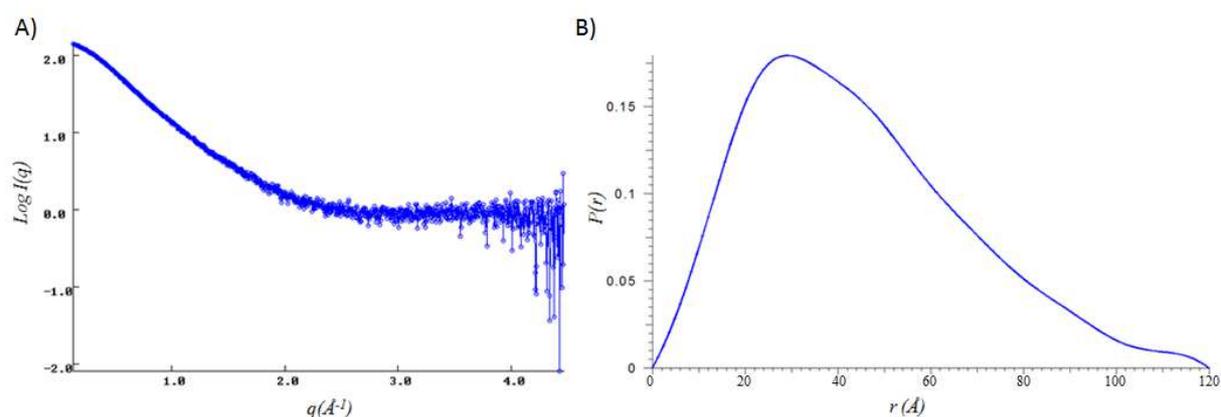
D'après la loi de Guinier le rayon de giration  $R_g$  ainsi que l'intensité  $I(0)$  peuvent être déterminés. Pour toutes les concentrations testées, un  $R_g$  d'environ 34 Å est observé ce qui correspond à une particule proche de 62 kDa, soit la taille du complexe CupB6/CupB4 (Figure 94). Aucune augmentation du  $R_g$  n'est observée lorsque la concentration de l'échantillon augmente (Figure 94). Donc à une concentration de 3,5 mg.ml<sup>-1</sup>, le complexe ne forme pas d'agrégats dus à la concentration et reste monodisperse, ce qui est important pour une potentielle cristallisation du complexe. Il aurait été intéressant de comparer également les valeurs de  $I(0)/c$  à différentes concentrations. Mais la présence d'agrégats observés sur les graphiques de Guinier à toutes les concentrations testées (3,5- 1,5- 0,5 mg.ml<sup>-1</sup>) ne permet pas de définir de  $I(0)$  (Figure 94). Ce type d'agrégat peut être dû à des dommages causés par les radiations pendant la collecte et peut être amélioré par une optimisation du tampon. En théorie l'ajout de DTT permet d'éliminer ce type d'agrégat, mais la présence d'un pont disulfure dans le domaine piline nous empêche d'utiliser cet agent réducteur. D'autres mesures de SAXS doivent être réalisées afin d'obtenir des courbes de diffusion sans la présence de ces agrégats. L'ajout de 10% de glycérol ou de l'additif glycyglycyl-glycine dans le tampon ainsi que la diminution de la température, du débit et du pourcentage de transmission pendant la collecte des données doivent être testés afin d'essayer d'éliminer ces agrégats.



**Figure 94: Calcul du rayon de giration  $R_g$  à partir de l'approximation de Guinier.** A) Graphique qui permet après traitement de définir le  $R_g$  pour l'échantillon CupB6/B4 à 3,5 mg.ml<sup>-1</sup>. La courbe de diffusion (en bleu) est étirée jusqu'à ce qu'elle soit en accord avec la courbe verte en pointillé qui

représente la loi de Guinier. L'étirement se limite à  $qR_g = 1,3$ . L'agrégation du complexe indiqué par un cadre gris n'est pas prise en compte dans l'analyse des données. Les valeurs qui sont utilisées pour déterminer le  $R_g$  sont comprises entre les deux traits verticaux rouges. Il est important que la ligne en vert foncé qui représente le nombre de points sélectionnés pour l'analyse du  $R_g$  soit en accord avec la ligne rose. B) Tableau regroupant les valeurs du  $R_g$  pour chaque concentration du complexe testée.

La détermination du  $R_g$  par l'approximation de Guinier représente également un indice de qualité des données. Pour une concentration de CupB6/B4 (64 kDa) à  $3,5 \text{ mg.ml}^{-1}$ , le  $R_g$  est déterminé avec 72 points de la courbe de diffusion aux petits angles (il est nécessaire d'avoir au moins 30 points pour une protéine de 100 kDa). La qualité des données est estimée à 90 %. Pour étudier l'enveloppe à basse résolution du complexe CupB6/B4, la courbe de diffusion qui correspond à une concentration de  $3,5 \text{ mg.ml}^{-1}$  a été sélectionnée (Figure 95). En effet les données aux petits angles et aux grands angles pour cette courbe de diffusion sont de meilleure qualité que celles obtenues avec les concentrations plus basses. Les 25 premiers points associés à l'agrégation du complexe ont été retirés causant une perte d'information aux petits angles.



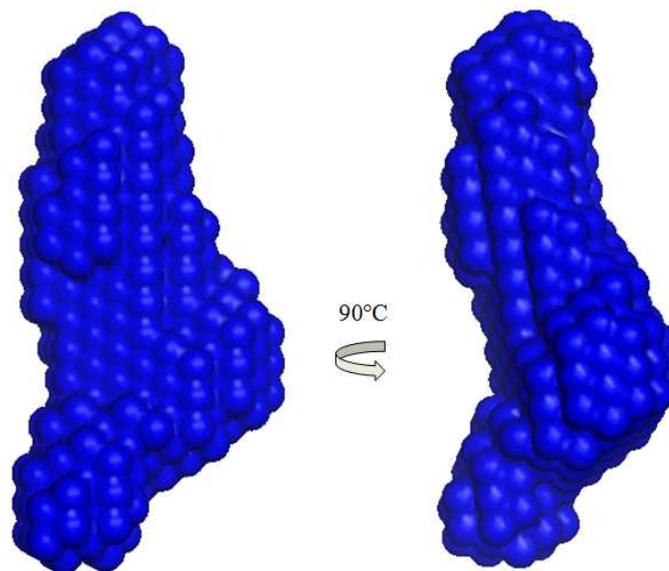
**Figure 95: A) Courbe de diffusion et B) la fonction de distribution  $P(r)$  obtenue avec une concentration de CupB6/B4 égale à  $3,5 \text{ mg/ml}$ .**

Le programme GNOM permet de calculer la fonction de distribution  $P(r)$  à partir de la courbe de diffusion par une transformée de Fourier indirecte. La forme de la fonction  $P(r)$  indique qu'il s'agit d'une protéine multi-domaines qui est très allongée (au maximum de  $P(r)$ ,  $r = 29 \text{ \AA}$ , ce qui est inférieur de  $D_{\max}/2$ ). La longueur maximale  $D_{\max}$  du complexe déterminée lorsque  $P(r) = 0$  pour une large valeur de  $r$  est de  $120 \text{ \AA}$ . D'après la fonction de distribution  $P(r)$  calculé, le programme GNOM donne un  $R_g$  qui est de  $35,2 \text{ \AA}$ . Le  $R_g$  calculé par GNOM correspond au  $R_g$  défini par l'approximation de Guinier ( $34 \text{ \AA}$ ), ce qui indique que les données qui vont servir au calcul de l'enveloppe du complexe sont correctes.

A partir de la courbe de distribution  $P(r)$ , 20 modèles ont été calculés par le programme DAMMIF afin de tenir compte de toutes les conformations que le complexe peut adopter en solution. La comparaison de ces différents modèles par le programme DAMMAVER détermine un indice NSD (Normalized Spatial Discrepancy) qui mesure le degré de similarité entre les modèles. L'indice NSD de tous les modèles calculés est inférieur à 1. Il varie de 0,8 à 0,87 ce qui indique une bonne corrélation. La superposition de tous les modèles sélectionnés permet de générer un modèle initial pour le calcul de l'enveloppe finale à basse résolution par la méthode *ab initio*.

#### 10.5.4.2. La modélisation *ab initio*

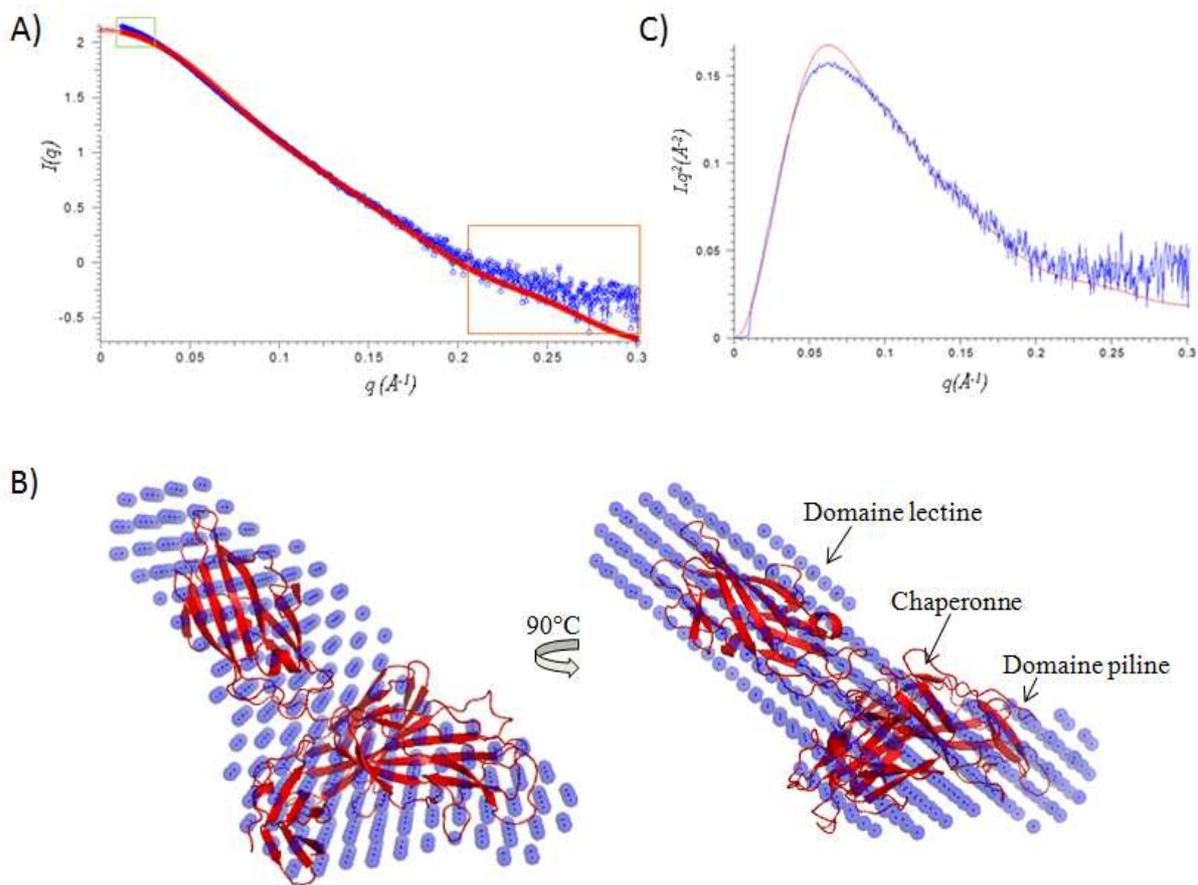
La courbe théorique calculée à partir de l'enveloppe reconstituée à basse résolution est mise en accord avec la courbe de diffusion expérimentale. Un  $\chi^2$  égal à 1 montre que l'enveloppe calculée est en bonne adéquation avec les données expérimentales. Un  $\chi^2 > 1$  indique que l'enveloppe calculée n'est pas de bonne qualité et doit être considérée avec réserve. Le  $D_{\max}$  mesuré à partir de l'enveloppe dans Pymol (121,45 Å) correspond au  $D_{\max}$  donné par la fonction de distribution (120 Å) (Figure 96).



**Figure 96: Enveloppe à basse résolution modélisée par *ab initio* du complexe CupB6/B4.**

La disponibilité de la structure atomique du complexe FimH/FimC d'*E. coli*, a permis grâce au programme CRY SOL d'évaluer l'enveloppe à basse résolution modélisée *ab initio*. La courbe de diffusion théorique, calculée à partir de la structure cristallographique du complexe FimH/FimC, a été mise en accord avec la courbe de diffusion expérimentale (Figure 97 A).

D'après les résultats d'accord entre les deux courbes de diffusion, une différence de taille ainsi qu'une différence de conformation entre les deux complexes sont observées. Le complexe FimH/FimC est de 56 kDa alors que le complexe CupB6/B4 est de 62 kDa. La superposition de la structure atomique et de l'enveloppe à basse résolution par le programme SUPERCOMB met aussi en évidence une différence de conformation (Figure 97 B). Ce résultat n'est pas surprenant car si les chaperonnes CU présentent de forte similarité structurales, les adhésines sont très différentes. Cette étude permet d'appuyer l'exactitude du modèle *ab initio*.



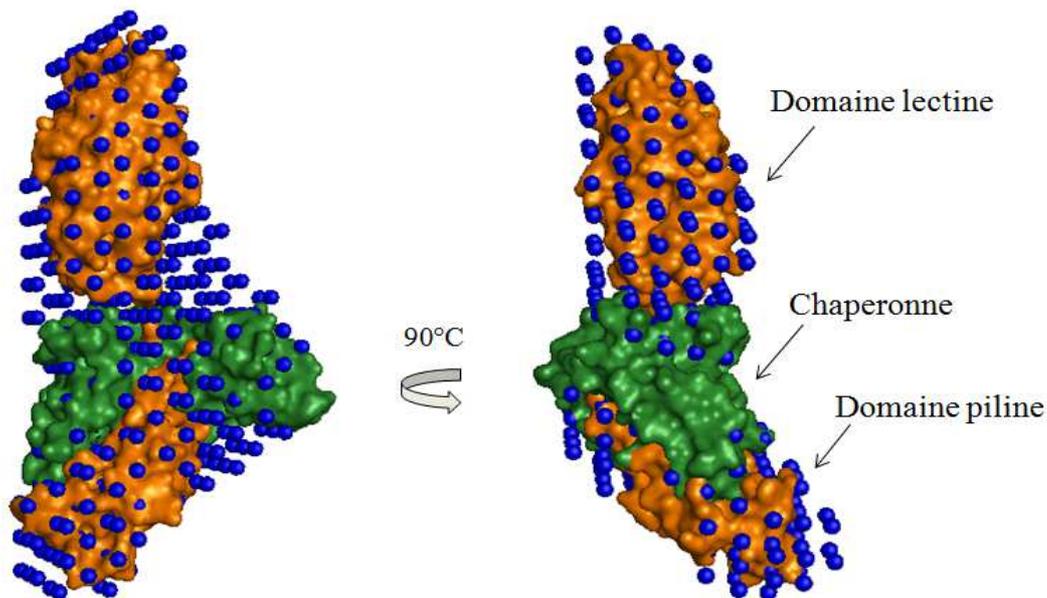
**Figure 97: Evaluation de l'exactitude de l'enveloppe à basse résolution A) et B) ainsi que le bon repliement C) du complexe CupB6/B4.** A) La courbe de diffusion expérimentale (en bleu) est mise en accord avec la courbe théorique (en rouge) calculée à partir de la structure atomique FimH/FimC d'*E. coli* par le programme CRY SOL. L'encadré vert montre une différence de taille et l'encadré orange une différence de conformation entre les deux complexes. B) La superposition par le programme SUPERCOMB de l'enveloppe CupB6/B4 (en bleu) et de la structure cristallographique de FimH/FimC (en rouge). La visualisation des structures se fait par Pymol. C) Etude du repliement de CupB6/B4 (en bleu) par la représentation de Kratky. En rouge figure la représentation de Kratky calculée à partir de la courbe de diffusion théorique de FimH/FimC obtenue par CRY SOL. Cette courbe est utilisée comme témoin du bon repliement du complexe CupB6/B4.

La représentation de Kratky montrant une courbe en forme de cloche similaire à celle calculée pour FimH/FimC indique que le complexe est correctement replié. Afin d'étudier la présence d'une flexibilité entre les domaines adhésine et piline de CupB6, la méthode EOM (ensemble organization method) a été utilisée.

#### 10.5.4.3. Etude de la flexibilité par EOM

A partir de la structure atomique du domaine adhésine de FimH, de la structure atomique de la chaperonne FimC complexée au domaine piline de FimH, puis de la séquence qui correspond au complexe CupB6/B4, un modèle de base a été calculé par le programme Pre-BRUNCH. Avec ce modèle de base où le domaine lectine CupB6 est séparé du complexe chaperonne/domaine piline par le linker de CupB6/B4, 1000 conformations ont été générées aléatoirement afin de recouvrir l'ensemble de l'enveloppe à basse résolution de CupB6/B4. Afin d'étudier la présence d'une flexibilité entre les deux domaines de CupB6, N=50 conformations sont sélectionnées par l'algorithme génétique en fonction de la courbe de diffusion expérimentale parmi les 1000 conformations générées. D'après la valeur de l'indice NSD qui est de 1.44 (inférieur à 1,5) et celle de la distribution du  $R_g$  qui est de 13,3 Å (inférieure à 20 Å), les résultats indiquent qu'il n'y a pas de grande flexibilité entre le domaine lectine et piline de CupB6 (Bernadó et al., 2007). Dans cette étude, l'utilisation de cette méthode a simplement mis en évidence l'absence d'une grande flexibilité entre les deux domaines de CupB6. Dans le cas d'une forte flexibilité entre les deux domaines, la méthode EOM aurait permis de représenter les différentes conformations que le complexe CupB6/B4 adopte en solution. L'étude EOM a été réalisée avec l'aide du Dr Julien Pérard (UVHCI-Grenoble).

L'absence d'une forte flexibilité du domaine adhésine de CupB6 a été vérifiée lorsque les structures atomiques du domaine lectine FimH, du domaine piline FimH et de la chaperonne CupB2 qui présente 32 % d'identité avec la chaperonne CupB4, ont été positionnées à l'intérieur de la reconstruction à basse résolution (Figure 98). La présence de flexibilité entre les deux domaines de CupB6 aurait généré une enveloppe à basse résolution plus globulaire. Bien que les conformations du domaine adhésine CupB6 et FimH soient différentes, il reste facile de définir sur l'enveloppe à basse résolution la position du domaine adhésine de CupB6.



**Figure 98: Modèle illustrant la position du domaine lectine et piline de l'adhésine ainsi que la chaperonne dans l'enveloppe à basse résolution de CupB6/B4.** Les surfaces du domaine lectine et piline de l'adhésine FimH sont représentées en orange. La surface de la chaperonne CupB2 figure en vert.

## 10.6. Conclusion

Les constructions dans les vecteurs possédant une étiquette volumineuse ont permis l'expression du domaine adhésine CupB6 sous forme soluble. Cependant lors des essais de purification, la présence d'une dégradation du domaine a été constatée. Les constructions choisies n'étant pas stables, d'autres doivent être réalisées en incluant une partie du linker.

Des études menées sur le complexe CupB6/B4 ont été entreprises afin de mieux le caractériser et d'obtenir des cristaux. Dans un premier temps, l'optimisation du pH de CupB6/B4 défini par TSA a permis d'obtenir des microcristaux avec 10 % v/v de polypropylène glycol 400. L'étude d'optimisation des conditions de cristallisation n'a pas encore permis d'obtenir des cristaux exploitables. Cependant certains paramètres ont été définis et d'autres essais d'optimisation doivent être réalisés.

Des cristaux correspondant à un fragment de lyse du complexe d'environ 25-20 kDa ont également été obtenus. Des essais sans résultat de remplacement moléculaire à partir de la structure de la chaperonne CupB2 présentant 32 % d'identité avec CupB4 laissent penser que les cristaux pourraient correspondre à un produit de protéolyse de CupB6. D'après les études de protéolyse ménagée, de spectrométrie de masse, et de séquençage N-ter, plusieurs fragments de protéolyse d'une taille de 25,5 et 24-25 kDa correspondant à CupB6 ont été identifiés. Des essais visant à reproduire ces cristaux doivent être entrepris.

La dégradation du complexe peut être à l'origine des difficultés rencontrées à obtenir des cristaux de CupB6/B4. Des précautions doivent être prises pour limiter cette lyse comme réaliser les essais de cristallisation à 4°C. La présence d'une grande flexibilité entre les domaines piline et lectine serait un autre paramètre qui pourrait gêner à la cristallisation du complexe. Cette hypothèse n'est pas valide puisque les expériences SAXS ont montré l'absence d'une grande flexibilité entre les domaines.

Les différentes études entreprises pour identifier un ligand naturel de cupB6 avec le complexe n'ont pas abouti. La monovalence et l'instabilité (agrégation) du complexe rend l'étude très difficile. L'obtention du domaine adhésine CupB6 serait un moyen de contourner les problèmes rencontrés.

# **Conclusions et Perspectives**



---

## 11. Conclusions et perspectives

---

Mes travaux de thèse s'inscrivent dans le contexte général de l'analyse structure-fonction des lectines. Ils ont été divisés en trois parties et m'ont permis d'aborder différents aspects de la biologie structurale tels que la biologie moléculaire, la biochimie, différentes techniques qui permettent de caractériser les interactions protéine-sucres, la diffusion des rayons X aux petits angles et la cristallographie aux rayons X.

La première partie de mes travaux consistait à caractériser par ITC et cristallographie aux rayons X les interactions entre la lectine de légumineuse, PELa, et certains de ses ligands naturels. Ce travail était en continuité avec celui du Dr Raquel Benavides qui avait montré sur puce à sucres auprès du « Consortium for Functional Glycomics » que PELa possédait une forte spécificité pour les *N*-glycanes de type complexe et asymétriques avec une courte branche 1-6 qui comporte un résidu  $\alpha$ -mannose en sucre non-réducteur et une branche 1-3 qui possède un LacNAc (N-acetyl-lactosamine) ou un LacNAc sialylé sur le résidu  $\alpha$ -mannose. Les travaux que j'ai menés ont permis de déterminer les paramètres thermodynamiques de deux ligands naturels potentiels de PELa présentant des branches 1-3 et 1-6 soit symétriques soit asymétriques. Ces deux composés présentent la même affinité en ITC alors que sur la puce à sucres testée, la présence d'un GlcNAc sur la branche 1-6 diminue de manière significative l'affinité de la lectine PELa. La différence d'entropie observée indique cependant que la présence d'un GlcNAc comme sucre réducteur sur la branche 1-6 est défavorable à l'interaction. Ensuite, j'ai entrepris une étude structurale par cristallographie aux rayons X des interactions mises en place entre le ligand d'intérêt et le site de reconnaissance de PELa. La structure du complexe PELa/*N*-glycane complexe et symétrique a montré que le GlcNAc sur la branche 1-6 possède une conformation coûteuse en énergie dans le site -1 de PELa. Ceci permet d'expliquer la spécificité inhabituelle de PELa pour les *N*-glycanes de type complexe asymétriques. En effet les lectines à mannose/glucose comme ConA, DGL, les lectines de *P. satium* et *Lens culinaris* possèdent généralement une spécificité pour les *N*-glycanes de type oligomannosique.

Les lectines de légumineuses sont utilisées en biotechnologie pour une variété d'applications comme antifongiques, agents antitumoraux, antibactériens, ou encore comme inhibiteurs de la reverse transcriptase du VIH1 (Lam and Ng, 2011). Elles sont aussi utilisées pour la purification de protéines glycosylées par chromatographie d'affinité comme la ConA. La

spécificité particulière de PelA pourrait en faire un outil biotechnologique et il serait intéressant d'étudier ses applications potentielles.

Les deux autres parties de mes travaux de thèse ont portés sur la caractérisation de lectines chez *Pseudomonas aeruginosa*. L'une d'elles était la lectine soluble spécifique du galactose appelée PA-IL. Les deux autres étaient des lectines putatives retrouvées aux extrémités des pili assemblés par le système chaperonne-usher, appelées CupB6 et CupE6. Le but ultime de tous ces travaux est la conception de glycomimétiques (composés à base de sucre) en vue d'une thérapie antiadhésive pour lutter contre les infections à *P. aeruginosa* et contre le biofilm notamment chez les patients atteints de la mucoviscidose.

La lectine PA-IL est étudiée au laboratoire depuis une dizaine d'année et est très bien connue d'un point de vue structural et fonctionnel. En collaboration avec des chimistes des sucres, plusieurs séries de glycomimétiques ont été conçues et testés pour évaluer leur potentiel en tant que anti-adhésif (Kadam et al., 2011; Cecioni et al., 2012; Reynolds et al., 2012). Dans cette thématique, mes travaux de thèse consistaient à évaluer le potentiel d'inhibition de plusieurs molécules monovalentes synthétisées par l'équipe du Pr R. Roy de l'université du Québec à Montréal. Des travaux préliminaires menés par le Dr Bertrand Blanchard avaient mis en évidence un très bon inhibiteur : le DEG144. Cet inhibiteur est composé d'un groupement naphtalène relié par une liaison thiol à un  $\beta$ Gal. Plusieurs stratégies ont été mises en place pour augmenter l'affinité de ce composé et donc son potentiel d'inhibition. L'une des stratégies consistait substituer le groupement naphtalène par un groupement 4-méthylcoumarine. Une étude par ITC a montré que la présence des deux atomes d'oxygène du composé JR127 pouvait augmenter la stabilité de l'interaction avec PA-IL. Cependant aucune différence significative au niveau de l'affinité n'a été observée entre les deux composés, DEG144 et JR127. Il est à noter que les deux molécules monovalentes possèdent une affinité ( $K_d \sim 5-6 \mu\text{M}$ ) qui est très forte pour PA-IL.

Dans une approche antiadhésive, il serait intéressant dans un premier temps d'étudier l'impact respectif des deux composés DEG144 et JR127 sur l'infection et le biofilm. Pour ce dernier, il existe différentes techniques de microscopie mais la plus simple reste la coloration au cristal violet. Pour l'infection, des tests *in vivo* sur un modèle murin permettraient d'étudier si les composés permettent une amélioration de la survie et une diminution de la perméabilité des capillaires alvéolaires induite par *P. aeruginosa*. Ce type d'étude a permis de montrer l'effet protecteur de plusieurs monosaccharides tel l'alpha-méthyle-D-galactose et l'alpha-méthyle-L-fucose (Chemani et al., 2009). Il a été aussi montré chez des patients atteints de

mucoviscidose que l'inhalation de galactose et fucose diminuait la quantité des bactéries présentes dans leurs sécrétions bronchiques (Hauber et al., 2008). En diminuant la population bactérienne et la formation du biofilm, l'inhalation de ce type de composé (DEG144) combiné à une antibiothérapie serait un moyen d'augmenter l'efficacité des antibiotiques et de lutter contre l'infection.

Afin de maximiser l'impact de la thérapie antiadhésive, les différentes lectines présentes chez *P. aeruginosa* doivent être ciblées de manière très spécifique. L'équipe de Sophie de Bentzmann a identifié différentes adhésines associées au pili CU qui pourraient être des lectines impliquées dans la formation du biofilm et dans l'attachement de la bactérie aux cellules hôtes. La majorité de mon travail de thèse portait sur les adhésines CupE6 et CupB6.

Dans un premier temps, des premiers travaux ont été réalisés sur l'adhésine CupE6. Différentes constructions ont été réalisées afin d'obtenir le domaine adhésine de CupE6 isolé. Les tests d'expression réalisés dans le cytoplasme n'ont conduit qu'à l'expression de ce domaine sous forme insoluble. L'éventuelle présence de ponts disulfures pourrait être responsable de l'expression sous forme de corps d'inclusion chez *E. coli* suite à leur mauvaise formation. L'expression du domaine adhésine CupE6 dans le périplasma pourrait permettre l'obtention de la protéine sous forme soluble et correctement repliée. De nouvelles constructions tels que CupE6 25/170 ; CupE6 37/170 ; CupE6 25/174 et CupE6 37/174 devraient être réalisées utilisant le vecteur pMAL-p2x (BioLab) qui permet une expression périplasmique.

Face aux difficultés rencontrées pour exprimer le domaine adhésine seul, une deuxième stratégie qui consiste à co-exprimer l'adhésine entière CupE6 avec sa chaperonne CupE4 a été entreprise. La chaperonne permet de stabiliser le repliement incomplet du domaine piline. Les tests d'expression entrepris sur différentes constructions n'ont donné aucune expression visible. Il est possible que le complexe soit toxique dans les différentes souches utilisées. Dans ce cas, la souche d'*E. coli* C43 conçue pour des protéines toxiques serait sans doute plus adéquate. Là encore une expression dans le périplasma serait plus appropriée au vu du nombre de ponts disulfures prédits pour l'adhésine CupE6.

En ce qui concerne l'adhésine CupB6, des tests d'expression ont également été réalisés avec plusieurs constructions qui correspondent au domaine adhésine. L'utilisation d'étiquettes volumineuses telles que la MBP, la thioredoxine ou la GST, a permis de produire ce domaine sous forme soluble. L'utilisation de ce type d'étiquette est également intéressante car elles ont permis de résoudre la structure de plusieurs protéines (Smyth et al., 2003). Les essais de

purification ont cependant mis en évidence que les constructions testées ne sont pas stables et lors des essais de purification des dégradations du domaine sont observées. D'autres constructions doivent être réalisées notamment en jouant sur l'extrémité C-terminale de ce domaine. L'ajout de quelques résidus imputés au linker pourrait peut-être stabiliser ce domaine au vue des constructions choisies pour l'expression du domaine adhésine FimH et PapG d'*E. coli*, (Dodson et al., 2001; Bouckaert et al., 2005).

L'expression du complexe CupB6/B4 avait été obtenue durant la thèse du Dr Blanchard et un premier protocole de purification avait été mis au point. Les premiers essais d'identification de la spécificité de CupB6 sur puces à sucre avaient été menés mais controversés. J'ai poursuivi l'étude visant à identifier des ligands de CupB6 grâce à des expériences de résonance plasmonique de surface et d'ITC mais les résultats ne sont pas concluants. La monovalence de CupB6, la faible affinité des lectines pour les sucres ainsi que l'instabilité du complexe en sont probablement les principales causes. Afin d'augmenter l'avidité de CupB6, nous avons voulu créer un tétramère artificiel suite à l'ajout d'une séquence de biotinylation à l'extrémité N-terminale de CupB6 pour une complexation avec la streptavidine. Malheureusement, cette stratégie n'a pas été poursuivie en raison des problèmes d'expression rencontrés. L'obtention d'un tétramère artificiel avec le domaine adhésine de CupB6 serait peut-être plus aisé. En effet en se basant sur les structures de FimH et PapG, l'insertion du peptide reconnu par la biotine ligase BirA en C-terminal ne devrait pas être un obstacle pour l'expression du domaine adhésine. Ce tétramère permettrait en augmentant l'avidité d'obtenir de meilleurs résultats avec les puces à sucres du « Consortium for Functional Glycomics ». Il serait très intéressant d'entreprendre une étude pour identifier la spécificité de CupB6 avec un biacore Flexchip. Ce type de biacore présente deux avantages : une puce qui permet de tester 400 sucres différents et la possibilité d'injecter la bactérie entière. Dans l'équipe de Sophie de Bentzmann, plusieurs mutants exprimant uniquement un seul type de pili Cup ont été obtenus. Cette étude serait un moyen d'étudier la spécificité des différentes adhésines Cup.

Des cristaux qui correspondent à un produit de protéolyse du complexe ont été obtenus par le Dr. Bertrand Blanchard et par moi-même. Les données collectées montrent que les cristaux sont composés d'une protéine de taille comprise entre 26,5 kDa et de 20 kDa. Les essais de remplacement moléculaire avec les homologues de la chaperonne CupB4 ou du domaine piline de FimH ou PapG n'ont pas aboutis. Des études de protéolyse ménagée, de spectrométrie de masse et de séquençage N-terminal sur le complexe ont permis d'identifier des fragments de 26,3 kDa et 25-24 kDa qui pourraient correspondre principalement au domaine adhésine ainsi qu'à une petite partie du domaine piline, et un fragment de 20 kDa qui

correspond à la chaperonne CupB4. Réaliser de la protéolyse *in situ* en utilisant les protéases trypsine et endopeptidase Glu-C serait peut-être un moyen de reproduire ces cristaux rapidement (Bai et al., 2007). L'autre stratégie consiste à réaliser les constructions plasmidiques correspondant à ces nouveaux fragments stables de CupB6 identifiés (CupB6 13-261 ; CupB6 13-241 ; CupB6 13-245 et CupB6 13-253) et de réaliser des tests d'expression. La lyse du complexe observée environ 3 semaines après la purification peut expliquer la difficulté de cristallisation. De nouvelles précautions peuvent être prises pour essayer d'endiguer le phénomène comme réaliser la purification et les essais de cristallisation à 4°C, utiliser de l'EDTA pourrait ralentir la lyse du complexe.

Les difficultés rencontrées pour cristalliser le complexe CupB6/B4 nous ont aussi conduits à sa caractérisation plus poussée et à une amélioration du protocole de purification. Des expériences par TSA ont permis dans un premier temps de définir le pH où le complexe est le plus stable (6.0) et d'obtenir des microcristaux. Des essais d'optimisation n'ont pas permis d'obtenir des cristaux exploitables mais ont néanmoins permis de cribler certains paramètres de cristallisation. L'apparition de cristaux sous forme d'amas pourrait être optimisée par l'utilisation de capillaire en gel d'agarose. Il serait également intéressant de réaliser des essais de cristallisation sans l'étiquette hexahistidine. L'obtention de fragments d'anticorps Fab spécifiques du complexe CupB6/B4 permettrait aussi d'envisager une co-cristallisation. Cette technique peut dans certains cas stabiliser certaines régions de la protéine et permettre l'obtention de cristaux. L'existence d'un linker proline et glycine riche entre les domaines adhésine et piline de CupB6 qui est relativement plus long que dans les autres adhésines homologues pourrait être à l'origine d'une certaine flexibilité qui gênerait sa cristallisation. Bien que l'étude de SAXS ne conforte pas cette hypothèse, la mutation de certains résidus glycines et prolines ou le remplacement du linker par celui de l'adhésine FimH serait une stratégie intéressante à utiliser pour obtenir des cristaux de CupB6/B4. La répétition de 5 prolines au niveau du linker pourrait être responsable de la formation d'une hélice  $\alpha$  polyproline. La présence de cette hélice pourrait expliquer l'interaction du complexe avec la streptavidine que nous avons observée par résonance plasmonique de surface. En effet ce type d'hélice joue un rôle très important dans de nombreuses interactions protéine-protéine (Mansiaux et al., 2011). La particularité du pilus CupB est d'une part la présence des protéines CupB5 qui présentent 44% de similarité avec les TPSA et d'autre part la présence d'un domaine POTRA au niveau de la protéine « usher », CupB3. Des études menées dans l'équipe de Sophie de Bentzmann ont montré que la présence de la chaperonne CupB4 dans le périplasme était nécessaire à l'export et la présence des cupB5 sur le pilus. Il est possible

d'imaginer que l'hélice  $\alpha$  polyproline permette l'interaction entre le complexe CupB6/B4 et CupB5. L'autre hypothèse pourrait laisser penser que l'adhésine CupB6 présente un site de reconnaissance pour les sucres mais également pour des protéines.

Une troisième stratégie qui n'a jamais été abordée dans ces travaux serait d'exprimer l'adhésine CupB6 avec l'extension Nter de la sous-unité CupB1. Contrairement au brin  $\beta$  G de la chaperonne, l'extension Nter de la sous-unité piline entraîne une conformation dite fermée du domaine piline (Sauer et al., 2002a). Cette conformation fermée augmente la stabilité de l'adhésine et donc les chances que celle-ci cristallise.

Des travaux portés sur les pili de type P d'*E. coli* ont abouti à la synthèse de molécules capables de bloquer la formation du pilus. Ces molécules également appelées pilicides se fixent sur une zone de la chaperonne PapD connue pour interagir avec la protéine « usher ». En inhibant l'interaction entre la protéine « usher » et le complexe chaperonne/piline, les molécules bloquent la formation de pili et permettent une forte réduction de l'adhésion aux cellules et donc une moindre formation du biofilm (Pinkner et al., 2006; Aberg et al., 2007; Chorell et al., 2010). Les chaperonnes étant très conservées, il serait intéressant de tester l'effet de ces pilicides sur les pili Cup de *P. aeruginosa*. Une étude visant à inhiber la synthèse des pili Cup chez *P. aeruginosa* pourrait être envisagée dans l'optique de combiner l'action des pilicides et des glycomimétiques au sein d'une thérapie destinée aux patients atteints de la mucoviscidose.

# **Références bibliographiques**



## A

- Aberg, V., Fällman, E., Axner, O., Uhlin, B.E., Hultgren, S.J., and Almqvist, F. (2007). Pilicides regulate pili expression in *E. coli* without affecting the functional properties of the pilus rod. *Mol Biosyst* 3, 214–218.
- Alhede, M., Bjarnsholt, T., Jensen, P.Ø., Phipps, R.K., Moser, C., Christophersen, L., Christensen, L.D., Van Gennip, M., Parsek, M., Høiby, N., et al. (2009). *Pseudomonas aeruginosa* recognizes and responds aggressively to the presence of polymorphonuclear leukocytes. *Microbiology (Reading, Engl.)* 155, 3500–3508.
- Alkawash, M.A., Soothill, J.S., and Schiller, N.L. (2006). Alginate lyase enhances antibiotic killing of mucoid *Pseudomonas aeruginosa* in biofilms. *APMIS* 114, 131–138.
- Allen, W.J., Phan, G., and Waksman, G. (2012). Pilus biogenesis at the outer membrane of Gram-negative bacterial pathogens. *Curr. Opin. Struct. Biol.* 22, 500–506.
- Anderson, G.G., Moreau-Marquis, S., Stanton, B.A., and O'Toole, G.A. (2008). In vitro analysis of tobramycin-treated *Pseudomonas aeruginosa* biofilms on cystic fibrosis-derived airway epithelial cells. *Infect. Immun.* 76, 1423–1433.
- Arora, S.K., Dasgupta, N., Lory, S., and Ramphal, R. (2000). Identification of two distinct types of flagellar cap proteins, FliD, in *Pseudomonas aeruginosa*. *Infect. Immun.* 68, 1474–1479.
- Arora, S.K., Ritchings, B.W., Almira, E.C., Lory, S., and Ramphal, R. (1998). The *Pseudomonas aeruginosa* flagellar cap protein, FliD, is responsible for mucin adhesion. *Infect. Immun.* 66, 1000–1007.
- Avichezer, D., and Gilboa-Garber, N. (1987). PA-II, the L-fucose and D-mannose binding lectin of *Pseudomonas aeruginosa* stimulates human peripheral lymphocytes and murine splenocytes. *FEBS Lett.* 216, 62–66.

## B

- Bai, Y., Auperin, T.C., and Tong, L. (2007). The use of *in situ* proteolysis in the crystallization of murine CstF-77. *Acta Crystallographica Section F Structural Biology and Crystallization Communications* 63, 135–138.
- Bajolet-Laudinat, O., Girod-de Bentzmann, S., Tournier, J.M., Madoulet, C., Plotkowski, M.C., Chippaux, C., and Puchelle, E. (1994). Cytotoxicity of *Pseudomonas aeruginosa* internal lectin PA-I to respiratory epithelial cells in primary culture. *Infect. Immun.* 62, 4481–4487.
- Baldus, S.E., Engelmann, K., and Hanisch, F.-G. (2004). MUC1 and the MUCs: a family of human mucins with impact in cancer biology. *Crit Rev Clin Lab Sci* 41, 189–231.
- Barasch, J., and al-Awqati, Q. (1993). Defective acidification of the biosynthetic pathway in cystic fibrosis. *J. Cell Sci. Suppl.* 17, 229–233.
- Benz, R., and Hancock, R.E. (1981). Properties of the large ion-permeable pores formed from protein F of *Pseudomonas aeruginosa* in lipid bilayer membranes. *Biochim. Biophys. Acta* 646, 298–308.
- Bernadó, P., Mylonas, E., Petoukhov, M.V., Blackledge, M., and Svergun, D.I. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* 129, 5656–5664.

Bernadó, P., and Svergun, D.I. (2012). Analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Methods Mol. Biol.* 896, 107–122.

Bernardi, A., and Cheshev, P. (2008). Interfering with the sugar code: design and synthesis of oligosaccharide mimics. *Chemistry* 14, 7434–7441.

Blanchard, B., Nurisso, A., Hollville, E., Tétaud, C., Wiels, J., Pokorná, M., Wimmerová, M., Varrot, A., and Imberty, A. (2008). Structural basis of the preferential binding for globo-series glycosphingolipids displayed by *Pseudomonas aeruginosa* lectin I. *J. Mol. Biol.* 383, 837–853.

Bordi, C., and De Bentzmann, S. (2011). Hacking into bacterial biofilms: a new therapeutic challenge. *Annals of Intensive Care* 1, 19.

Boucher, R.C. (2003). Regulation of airway surface liquid volume by human airway epithelia. *Pflugers Arch.* 445, 495–498.

Boucher, R.C. (2004). New concepts of the pathogenesis of cystic fibrosis lung disease. *Eur Respir J* 23, 146–158.

Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhrer, M., Hung, C.-S., Pinkner, J., Slättegård, R., Zavialov, A., et al. (2005). Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* 55, 441–455.

Brown, M.T., Delalez, N.J., and Armitage, J.P. (2011). Protein dynamics and mechanisms controlling the rotational behaviour of the bacterial flagellar motor. *Current Opinion in Microbiology* 14, 734–740.

Bucior, I., Pielage, J.F., and Engel, J.N. (2012). *Pseudomonas aeruginosa* Pili and Flagella Mediate Distinct Binding and Signaling Events at the Apical and Basolateral Surface of Airway Epithelium. *PLoS Pathog* 8, e1002616.

Buts, L., Garcia-Pino, A., Imberty, A., Amiot, N., Boons, G.-J., Beeckmans, S., Versées, W., Wyns, L., and Loris, R. (2006). Structural basis for the recognition of complex-type biantennary oligosaccharides by *Pterocarpus angolensis* lectin. *FEBS Journal* 273, 2407–2420.

Byrd, M.S., Sadovskaya, I., Vinogradov, E., Lu, H., Sprinkle, A.B., Richardson, S.H., Ma, L., Ralston, B., Parsek, M.R., Anderson, E.M., et al. (2009). Genetic and biochemical analyses of the *Pseudomonas aeruginosa* Psl exopolysaccharide reveal overlapping roles for polysaccharide synthesis enzymes in Psl and LPS production. *Mol. Microbiol.* 73, 622–638.

## C

Cai, X., Wang, R., Filloux, A., Waksman, G., and Meng, G. (2011). Structural and functional characterization of *Pseudomonas aeruginosa* CupB chaperones. *PLoS ONE* 6, e16583.

Cantón, R., Cobos, N., De Gracia, J., Baquero, F., Honorato, J., Gartner, S., Alvarez, A., Salcedo, A., Oliver, A., and García-Quetglas, E. (2005). Antimicrobial therapy for pulmonary pathogenic colonisation and infection by *Pseudomonas aeruginosa* in cystic fibrosis patients. *Clin. Microbiol. Infect.* 11, 690–703.

Cecioni, S., Faure, S., Darbost, U., Bonnamour, I., Parrot-Lopez, H., Roy, O., Taillefumier, C., Wimmerová, M., Praly, J.-P., Imberty, A., et al. (2011). Selectivity among two lectins: probing the effect of topology, multivalency and flexibility of “clicked” multivalent glycoclusters. *Chemistry* 17, 2146–2159.

Cecioni, S., Praly, J.-P., Matthews, S.E., Wimmerová, M., Imberty, A., and Vidal, S. (2012). Rational design and synthesis of optimized glycoclusters for multivalent lectin-carbohydrate interactions: influence of the linker arm. *Chemistry* 18, 6250–6263.

Chabre, Y.M., Giguère, D., Blanchard, B., Rodrigue, J., Rocheleau, S., Neault, M., Rauthu, S., Papadopoulos, A., Arnold, A.A., Imberty, A., et al. (2011). Combining glycomimetic and multivalent strategies toward designing potent bacterial lectin inhibitors. *Chemistry* 17, 6545–6562.

Chemani, C., Imberty, A., De Bentzmann, S., Pierre, M., Wimmerová, M., Guery, B.P., and Faure, K. (2009). Role of LecA and LecB lectins in *Pseudomonas aeruginosa*-induced lung injury and effect of carbohydrate ligands. *Infect. Immun.* 77, 2065–2075.

Chevance, F.F.V., and Hughes, K.T. (2008). Coordinating assembly of a bacterial macromolecular machine. *Nature Reviews Microbiology* 6, 455–465.

Chmiel, J.F., and Davis, P.B. (2003). State of the Art: Why do the lungs of patients with cystic fibrosis become infected and why can't they clear the infection? *Respiratory Research* 4, 8.

Chorell, E., Pinkner, J.S., Phan, G., Edvinsson, S., Buelens, F., Remaut, H., Waksman, G., Hultgren, S.J., and Almqvist, F. (2010). Design and synthesis of C-2 substituted thiazolo and dihydrothiazolo ring-fused 2-pyridones: pilicides with increased antivirulence activity. *J. Med. Chem.* 53, 5690–5695.

Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S.J., and Knight, S.D. (1999). X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* 285, 1061–1066.

Cioci, G., Mitchell, E.P., Gautier, C., Wimmerová, M., Sudakevitz, D., Pérez, S., Gilboa-Garber, N., and Imberty, A. (2003). Structural basis of calcium and galactose recognition by the lectin PA-IL of *Pseudomonas aeruginosa*. *FEBS Letters* 555, 297–301.

Colomb, F., Krzewinski-Recchi, M.-A., El Machhour, F., Mensier, E., Jaillard, S., Steenackers, A., Harduin-Lepers, A., Lafitte, J.-J., Delannoy, P., and Groux-Degroote, S. (2012). TNF regulates sialyl-Lewis(x) and 6-sulfo-sialyl-Lewis(x) expression in human lung through up-regulation of ST3GAL4 transcript isoform BX. *Biochimie* 94, 2045–2053.

Craig, L., Volkmann, N., Arvai, A.S., Pique, M.E., Yeager, M., Egelman, E.H., and Tainer, J.A. (2006). Type IV pilus structure by cryo-electron microscopy and crystallography: implications for pilus assembly and functions. *Mol. Cell* 23, 651–662.

Croft, L., Beatson, S.A., Whitchurch, C.B., Huang, B., Blakeley, R.L., and Mattick, J.S. (2000). An interactive web-based *Pseudomonas aeruginosa* genome database: discovery of new genes, pathways and structures. *Microbiology (Reading, Engl.)* 146 ( Pt 10), 2351–2364.

## D

D'Argenio, D.A., Calfee, M.W., Rainey, P.B., and Pesci, E.C. (2002). Autolysis and autoaggregation in *Pseudomonas aeruginosa* colony morphology mutants. *J. Bacteriol.* 184, 6481–6489.

Davril, M., Degroote, S., Humbert, P., Galabert, C., Dumur, V., Lafitte, J.J., Lamblin, G., and Roussel, P. (1999). The sialylation of bronchial mucins secreted by patients suffering from cystic fibrosis or from chronic bronchitis is related to the severity of airway infection. *Glycobiology* 9, 311–321.

Diggle, S.P., Stacey, R.E., Dodd, C., Cámara, M., Williams, P., and Winzer, K. (2006). The galactophilic lectin, LecA, contributes to biofilm development in *Pseudomonas aeruginosa*. *Environ. Microbiol.* 8, 1095–1104.

Dodson, K.W., Pinkner, J.S., Rose, T., Magnusson, G., Hultgren, S.J., and Waksman, G. (2001). Structural basis of the interaction of the pyelonephritic *E. coli* adhesin to its human kidney receptor. *Cell* 105, 733–743.

Döring, G., Conway, S.P., Heijerman, H.G., Hodson, M.E., Høiby, N., Smyth, A., and Touw, D.J. (2000). Antibiotic therapy against *Pseudomonas aeruginosa* in cystic fibrosis: a European consensus. *Eur. Respir. J.* 16, 749–767.

Dupeux, F., Röwer, M., Seroul, G., Blot, D., and Márquez, J.A. (2011). A thermal stability assay can help to estimate the crystallization likelihood of biological samples. *Acta Crystallogr. D Biol. Crystallogr.* 67, 915–919.

## E

Eidam, O., Dworkowski, F.S.N., Glockshuber, R., Grütter, M.G., and Capitani, G. (2008). Crystal structure of the ternary FimC-FimF(t)-FimD(N) complex indicates conserved pilus chaperone-subunit complex recognition by the usher FimD. *FEBS Lett.* 582, 651–655.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2126–2132.

Epp, S.F., Köhler, T., Plésiat, P., Michéa-Hamzehpour, M., Frey, J., and Pechère, J.C. (2001). C-terminal region of *Pseudomonas aeruginosa* outer membrane porin OprD modulates susceptibility to meropenem. *Antimicrob. Agents Chemother.* 45, 1780–1787.

Epps, D.E., Sarver, R.W., Rogers, J.M., Herberg, J.T., and Tomich, P.K. (2001). The ligand affinity of proteins measured by isothermal denaturation kinetics. *Anal. Biochem.* 292, 40–50.

## F

Feldman, M., Bryan, R., Rajan, S., Scheffler, L., Brunnert, S., Tang, H., and Prince, A. (1998). Role of flagella in pathogenesis of *Pseudomonas aeruginosa* pulmonary infection. *Infect. Immun.* 66, 43–51.

Ferrè, F., and Clote, P. (2006). DiANNA 1.1: an extension of the DiANNA web server for ternary cysteine classification. *Nucleic Acids Res.* 34, W182–185.

Finck-Barbançon, V., Goranson, J., Zhu, L., Sawa, T., Wiener-Kronish, J.P., Fleiszig, S.M., Wu, C., Mende-Mueller, L., and Frank, D.W. (1997). ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity and epithelial injury. *Mol. Microbiol.* 25, 547–557.

Folkesson, A., Haagensen, J.A.J., Zampaloni, C., Sternberg, C., and Molin, S. (2008). Biofilm induced tolerance towards antimicrobial peptides. *PLoS ONE* 3, e1891.

Friedman, L., and Kolter, R. (2004). Two genetic loci produce distinct carbohydrate-rich structural components of the *Pseudomonas aeruginosa* biofilm matrix. *J. Bacteriol.* 186, 4457–4465.

## G

Gaboriaud, C., Bissery, V., Benchetrit, T., and Mornon, J.P. (1987). Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. *FEBS Lett.* 224, 149–155.

Garber, N., Guempel, U., Belz, A., Gilboa-Garber, N., and Doyle, R.J. (1992). On the specificity of the D-galactose-binding lectin (PA-I) of *Pseudomonas aeruginosa* and its strong binding to hydrophobic derivatives of D-galactose and thiogalactose. *Biochim. Biophys. Acta* 1116, 331–333.

Gavira, J.A., and García-Ruiz, J.M. (2002). Agarose as crystallisation media for proteins II: trapping of gel fibres into the crystals. *Acta Crystallogr. D Biol. Crystallogr.* *58*, 1653–1656.

Gilboa-Garber, N., Sudakevitz, D., Sheffi, M., Sela, R., and Levene, C. (1994). PA-I and PA-II lectin interactions with the ABO(H) and P blood group glycosphingolipid antigens may contribute to the broad spectrum adherence of *Pseudomonas aeruginosa* to human tissues in secondary infections. *Glycoconj. J.* *11*, 414–417.

Giraud, C., and De Bentzmann, S. (2012). Inside the complex regulation of *Pseudomonas aeruginosa* chaperone usher systems. *Environ. Microbiol.* *14*, 1805–1816.

Giraud, C., Bernard, C.S., Calderon, V., Yang, L., Filloux, A., Molin, S., Fichant, G., Bordi, C., and De Bentzmann, S. (2011). The PprA-PprB two-component system activates CupE, the first non-archetypal *Pseudomonas aeruginosa* chaperone-usher pathway system assembling fimbriae. *Environ. Microbiol.* *13*, 666–683.

Golovanov, A.P., Hautbergue, G.M., Wilson, S.A., and Lian, L.-Y. (2004). A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc.* *126*, 8933–8939.

Goochee, C.F., and Monica, T. (1990). Environmental effects on protein glycosylation. *Biotechnology (N.Y.)* *8*, 421–427.

Gossert, A.D., Bettendorff, P., Puorger, C., Vetsch, M., Herrmann, T., Glockshuber, R., and Wüthrich, K. (2008). NMR structure of the *Escherichia coli* type 1 pilus subunit FimF and its interactions with other pilus subunits. *J. Mol. Biol.* *375*, 752–763.

Von Götz, F., Häussler, S., Jordan, D., Saravanamuthu, S.S., Wehmhöner, D., Strüßmann, A., Lauber, J., Attree, I., Buer, J., Tümmler, B., et al. (2004). Expression analysis of a highly adherent and cytotoxic small colony variant of *Pseudomonas aeruginosa* isolated from a lung of a patient with cystic fibrosis. *J. Bacteriol.* *186*, 3837–3847.

Gupta, S.K., Berk, R.S., Masinick, S., and Hazlett, L.D. (1994). Pili and lipopolysaccharide of *Pseudomonas aeruginosa* bind to the glycolipid asialo GM1. *Infect. Immun.* *62*, 4572–4579.

Gustke, H., Kleene, R., Loers, G., Nehmann, N., Jaehne, M., Bartels, K.-M., Jaeger, K.-E., Schachner, M., and Schumacher, U. (2012). Inhibition of the bacterial lectins of *Pseudomonas aeruginosa* with monosaccharides and peptides. *Eur. J. Clin. Microbiol. Infect. Dis.* *31*, 207–215.

## H

Hahn, E., Wild, P., Hermanns, U., Sebbel, P., Glockshuber, R., Häner, M., Taschner, N., Burkhard, P., Aebi, U., and Müller, S.A. (2002). Exploring the 3D molecular architecture of *Escherichia coli* type 1 pili. *J. Mol. Biol.* *323*, 845–857.

Hartmann, M., Papavlassopoulos, H., Chandrasekaran, V., Grabosch, C., Beiroth, F., Lindhorst, T.K., and Röhl, C. (2012). Inhibition of bacterial adhesion to live human cells: Activity and cytotoxicity of synthetic mannosides. *FEBS Letters* *586*, 1459–1465.

Hatch, R.A., and Schiller, N.L. (1998). Alginate lyase promotes diffusion of aminoglycosides through the extracellular polysaccharide of mucoid *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* *42*, 974–977.

Hauber, H.-P., Schulz, M., Pforte, A., Mack, D., Zabel, P., and Schumacher, U. (2008). Inhalation with fucose and galactose for treatment of *Pseudomonas aeruginosa* in cystic fibrosis patients. *Int J Med Sci* *5*, 371–376.

Häussler, S., Ziegler, I., Löttel, A., Von Götz, F., Rohde, M., Wehmhöner, D., Saravanamuthu, S., Tümmler, B., and Steinmetz, I. (2003). Highly adherent small-colony variants of *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *J. Med. Microbiol.* *52*, 295–301.

He, J., Baldini, R.L., Déziel, E., Saucier, M., Zhang, Q., Liberati, N.T., Lee, D., Urbach, J., Goodman, H.M., and Rahme, L.G. (2004). The broad host range pathogen *Pseudomonas aeruginosa* strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. *Proc. Natl. Acad. Sci. U.S.A.* *101*, 2530–2535.

Hogardt, M., and Heesemann, J. (2010). Adaptation of *Pseudomonas aeruginosa* during persistence in the cystic fibrosis lung. *International Journal of Medical Microbiology* *300*, 557–562.

Holmgren, A., and Bränden, C.I. (1989). Crystal structure of chaperone protein PapD reveals an immunoglobulin fold. *Nature* *342*, 248–251.

Hoyle, B.D., and Costerton, J.W. (1991). Bacterial resistance to antibiotics: the role of biofilms. *Prog Drug Res* *37*, 91–105.

Hubbard, S.C., and Ivatt, R.J. (1981). Synthesis and processing of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* *50*, 555–583.

Hung, C.-S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C.G., Strouse, R., Langermann, S., Waksman, G., et al. (2002). Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.* *44*, 903–915.

Hwang, T.-C., and Sheppard, D.N. (2009). Gating of the CFTR Cl<sup>-</sup> channel by ATP-driven nucleotide-binding domain dimerisation. *J Physiol* *587*, 2151–2161.

## I

Ikeda, T., Oosawa, K., and Hotani, H. (1996). Self-assembly of the Filament Capping Protein, FliD, of Bacterial Flagella into an Annular Structure. *Journal of Molecular Biology* *259*, 679–686.

Imberty, A., wimmerová, M., Mitchell, E.P., and Gilboa-Garber, N. (2004). Structures of the lectins from *Pseudomonas aeruginosa*: insight into the molecular basis for host glycan recognition. *Microbes Infect.* *6*, 221–228.

## J

Jacob-Dubuisson, F., Kuehn, M., and Hultgren, S.J. (1993). A novel secretion apparatus for the assembly of adhesive bacterial pili. *Trends Microbiol.* *1*, 50–55.

Johansson, E.M.V., Crusz, S.A., Kolomiets, E., Buts, L., Kadam, R.U., Cacciarini, M., Bartels, K.-M., Diggle, S.P., Cámara, M., Williams, P., et al. (2008). Inhibition and Dispersion of *Pseudomonas aeruginosa* Biofilms by Glycopeptide Dendrimers Targeting the Fucose-Specific Lectin LecB. *Chemistry & Biology* *15*, 1249–1257.

## K

Kadam, R.U., Bergmann, M., Hurley, M., Garg, D., Cacciarini, M., Swiderska, M.A., Nativi, C., Sattler, M., Smyth, A.R., Williams, P., et al. (2011). A Glycopeptide Dendrimer Inhibitor of the Galactose-Specific Lectin LecA and of *Pseudomonas aeruginosa* Biofilms. *Angewandte Chemie International Edition* *50*, 10631–10635.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4, 363–371.

Klebensberger, J., Birkenmaier, A., Geffers, R., Kjelleberg, S., and Philipp, B. (2009). SiaA and SiaD are essential for inducing autoaggregation as a specific response to detergent stress in *Pseudomonas aeruginosa*. *Environ. Microbiol.* 11, 3073–3086.

Klein, T., Abgottspon, D., Wittwer, M., Rabbani, S., Herold, J., Jiang, X., Kleeb, S., Lüthi, C., Scharenberg, M., Bezençon, J., et al. (2010). FimH Antagonists for the Oral Treatment of Urinary Tract Infections: From Design and Synthesis to in Vitro and in Vivo Evaluation. *J. Med. Chem.* 53, 8627–8641.

Klinth, J.E., Castelain, M., Uhlin, B.E., and Axner, O. (2012). The influence of pH on the specific adhesion of P piliated *Escherichia coli*. *PLoS ONE* 7, e38548.

Köhler, T., Michea-Hamzehpour, M., Epp, S.F., and Pechere, J.C. (1999). Carbapenem activities against *Pseudomonas aeruginosa*: respective contributions of OprD and efflux systems. *Antimicrob. Agents Chemother.* 43, 424–427.

Kuchma, S.L., Griffin, E.F., and O'Toole, G.A. (2012). Minor Pilins of the Type IV Pilus System Participate in the Negative Regulation of Swarming Motility. *J. Bacteriol.*

Kulasekara, H.D., Ventre, I., Kulasekara, B.R., Lazdunski, A., Filloux, A., and Lory, S. (2005). A novel two-component system controls the expression of *Pseudomonas aeruginosa* fimbrial cup genes. *Molecular Microbiology* 55, 368–380.

## L

Lam, S.K., and Ng, T.B. (2011). Lectins: production and practical applications. *Appl. Microbiol. Biotechnol.* 89, 45–55.

Lamblin, G., Degroote, S., Perini, J.M., Delmotte, P., Scharfman, A., Davril, M., Lo-Guidice, J.M., Houdret, N., Dumur, V., Klein, A., et al. (2001). Human airway mucin glycosylation: a combinatorial of carbohydrate determinants which vary in cystic fibrosis. *Glycoconj. J.* 18, 661–684.

Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.

Lee, D.G., Urbach, J.M., Wu, G., Liberati, N.T., Feinbaum, R.L., Miyata, S., Diggins, L.T., He, J., Saucier, M., Déziel, E., et al. (2006). Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial. *Genome Biol.* 7, R90.

Lewis, H.A., Buchanan, S.G., Burley, S.K., Conners, K., Dickey, M., Dorwart, M., Fowler, R., Gao, X., Guggino, W.B., Hendrickson, W.A., et al. (2004). Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator. *EMBO J.* 23, 282–293.

Lillehoj, E.P., Kim, B.T., and Kim, K.C. (2002). Identification of *Pseudomonas aeruginosa* flagellin as an adhesin for Muc1 mucin. *Am J Physiol Lung Cell Mol Physiol* 282, L751–L756.

Loris, R., Hamelryck, T., Bouckaert, J., and Wyns, L. (1998). Legume lectin structure. *Biochimica Et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1383, 9–36.

Loris, R., Imberty, A., Beeckmans, S., Driessche, E.V., Read, J.S., Bouckaert, J., Greve, H.D., Buts, L., and Wyns, L. (2003). Crystal Structure of Pterocarpus angolensis Lectin in Complex with Glucose, Sucrose, and Turanose. *J. Biol. Chem.* 278, 16297–16303.

## M

Maccioni, H.J.F., Quiroga, R., and Spessott, W. (2011). Organization of the synthesis of glycolipid oligosaccharides in the Golgi complex. *FEBS Lett.* 585, 1691–1698.

Mansiaux, Y., Joseph, A.P., Gelly, J.-C., and De Brevern, A.G. (2011). Assignment of PolyProline II Conformation and Analysis of Sequence – Structure Relationship. *PLoS ONE* 6, e18401.

Marionneau, S., Cailleau-Thomas, A., Rocher, J., Le Moullac-Vaidye, B., Ruvoën, N., Clément, M., and Le Pendu, J. (2001). ABH and Lewis histo-blood group antigens, a model for the meaning of oligosaccharide diversity in the face of a changing world. *Biochimie* 83, 565–573.

Mathee, K., Ciofu, O., Sternberg, C., Lindum, P.W., Campbell, J.I., Jensen, P., Johnsen, A.H., Givskov, M., Ohman, D.E., Molin, S., et al. (1999). Mucoid conversion of *Pseudomonas aeruginosa* by hydrogen peroxide: a mechanism for virulence activation in the cystic fibrosis lung. *Microbiology (Reading, Engl.)* 145 ( Pt 6), 1349–1357.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* 40, 658–674.

McGaughey, G.B., Gagné, M., and Rappé, A.K. (1998).  $\pi$ -Stacking Interactions ALIVE AND WELL IN PROTEINS. *J. Biol. Chem.* 273, 15458–15463.

Meissner, A., Wild, V., Simm, R., Rohde, M., Erck, C., Bredenbruch, F., Morr, M., Römling, U., and Häussler, S. (2007). *Pseudomonas aeruginosa* cupA-encoded fimbriae expression is regulated by a GGDEF and EAL domain-dependent modulation of the intracellular level of cyclic diguanylate. *Environ. Microbiol.* 9, 2475–2485.

Mertens, H.D.T., and Svergun, D.I. (2010). Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *Journal of Structural Biology* 172, 128–141.

Mewe, M., Tielker, D., Schönberg, R., Schachner, M., Jaeger, K.-E., and Schumacher, U. (2005). *Pseudomonas aeruginosa* lectins I and II and their interaction with human airway cilia. *J Laryngol Otol* 119, 595–599.

Mikkelsen, H., Sivaneson, M., and Filloux, A. (2011). Key two-component regulatory systems that control biofilm formation in *Pseudomonas aeruginosa*. *Environ. Microbiol.* 13, 1666–1681.

Miller, A.K., and Tausig, F. (1964). Biotin-binding by parenterally-administered streptavidin or avidin. *Biochem. Biophys. Res. Commun.* 14, 210–214.

Mishra, M., Byrd, M.S., Sergeant, S., Azad, A.K., Parsek, M.R., McPhail, L., Schlesinger, L.S., and Wozniak, D.J. (2012). *Pseudomonas aeruginosa* Psl polysaccharide reduces neutrophil phagocytosis and the oxidative response by limiting complement-mediated opsonization. *Cell. Microbiol.* 14, 95–106.

Mitchell, E., Houles, C., Sudakevitz, D., Wimmerova, M., Gautier, C., Pérez, S., Wu, A.M., Gilboa-Garber, N., and Imberty, A. (2002). Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients. *Nature Structural & Molecular Biology* 9, 918–921.

Monroe, D. (2007). Looking for chinks in the armor of bacterial biofilms. *PLoS Biol.* 5, e307.

Moothoo, D.N., Canan, B., Field, R.A., and Naismith, J.H. (1999). Man alpha1-2 Man alpha-OMe-concanavalin A complex reveals a balance of forces involved in carbohydrate recognition. *Glycobiology* 9, 539–545.

Mouneimné, H., Robert, J., Jarlier, V., and Cambau, E. (1999). Type II topoisomerase mutations in ciprofloxacin-resistant strains of *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 43, 62–66.

Mu, X.-Q., and Bullitt, E. (2006). Structure and assembly of P-pili: a protruding hinge region used for assembly of a bacterial adhesion filament. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9861–9866.

## N

Newburg, D.S., Ruiz-Palacios, G.M., and Morrow, A.L. (2005). Human milk glycans protect infants against enteric pathogens. *Annu. Rev. Nutr.* 25, 37–58.

Nguyen, Y., Jackson, S.G., Aidoo, F., Junop, M., and Burrows, L.L. (2010). Structural characterization of novel *Pseudomonas aeruginosa* type IV pilins. *J. Mol. Biol.* 395, 491–503.

Nierengarten, J.-F., Iehl, J., Oerthel, V., Holler, M., Illescas, B.M., Muñoz, A., Martín, N., Rojo, J., Sánchez-Navarro, M., Cecioni, S., et al. (2010). Fullerene sugar balls. *Chem. Commun. (Camb.)* 46, 3860–3862.

Niv, Y., and Fass, R. (2011). The role of mucin in GERD and its complications. *Nature Reviews Gastroenterology & Hepatology* 9, 55–59.

Nuccio, S.-P., and Bäumler, A.J. (2007). Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol. Mol. Biol. Rev.* 71, 551–575.

## O

Ohlsson, J., Jass, J., Uhlin, B.E., Kihlberg, J., and Nilsson, U.J. (2002). Discovery of potent inhibitors of PapG adhesins from uropathogenic *Escherichia coli* through synthesis and evaluation of galabiose derivatives. *Chembiochem* 3, 772–779.

Ohlsson, J., Larsson, A., Haataja, S., Alajääski, J., Stenlund, P., Pinkner, J.S., Hultgren, S.J., Finne, J., Kihlberg, J., and Nilsson, U.J. (2005). Structure–activity relationships of galabioside derivatives as inhibitors of *E. coli* and *S. suis* adhesins: nanomolar inhibitors of *S. suis* adhesins. *Org. Biomol. Chem.* 3, 886–900.

Ohtsubo, K., and Marth, J.D. (2006). Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* 126, 855–867.

Ostedgaard, L.S., Baldursson, O., and Welsh, M.J. (2001). Regulation of the Cystic Fibrosis Transmembrane Conductance Regulator Cl<sup>-</sup> Channel by Its R Domain. *J. Biol. Chem.* 276, 7689–7692.

## P

Parge, H.E., Forest, K.T., Hickey, M.J., Christensen, D.A., Getzoff, E.D., and Tainer, J.A. (1995). Structure of the fibre-forming protein pilin at 2.6 Å resolution. , Published Online: 02 November 1995; | Doi:10.1038/378032a0 378, 32–38.

Parsek, M.R., and Fuqua, C. (2004). Biofilms 2003: Emerging Themes and Challenges in Studies of Surface-Associated Microbial Life. *J Bacteriol* 186, 4427–4440.

Phan, G., Remaut, H., Wang, T., Allen, W.J., Pirker, K.F., Lebedev, A., Henderson, N.S., Geibel, S., Volkan, E., Yan, J., et al. (2011). Crystal structure of the FimD usher bound to its cognate FimC-FimH substrate. *Nature* 474, 49–53.

Phillips, K., and De la Peña, A.H. (2011). The combined use of the ThermoFluor assay and ThermoQ analytical software for the determination of protein stability and buffer optimization as an aid in protein crystallization. *Curr Protoc Mol Biol Chapter 10*, Unit10.28.

Pinkner, J.S., Remaut, H., Buelens, F., Miller, E., Aberg, V., Pemberton, N., Hedenström, M., Larsson, A., Seed, P., Waksman, G., et al. (2006). Rationally designed small compounds inhibit pilus biogenesis in uropathogenic bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 103, 17897–17902.

Poole, K. (2005). Aminoglycoside resistance in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* 49, 479–487.

Porubsky, S., Speak, A.O., Salio, M., Jennemann, R., Bonrouhi, M., Zafarulla, R., Singh, Y., Dyson, J., Luckow, B., Lehuen, A., et al. (2012). Globosides but not Isoglobosides Can Impact the Development of Invariant NKT Cells and Their Interaction with Dendritic Cells. *J Immunol.*

Powlesland, A.S., Ward, E.M., Sadhu, S.K., Guo, Y., Taylor, M.E., and Drickamer, K. (2006). Widely divergent biochemical properties of the complete set of mouse DC-SIGN-related proteins. *J. Biol. Chem.* 281, 20440–20449.

Puorger, C., Vetsch, M., Wider, G., and Glockshuber, R. (2011). Structure, folding and stability of FimA, the main structural subunit of type 1 pili from uropathogenic *Escherichia coli* strains. *J. Mol. Biol.* 412, 520–535.

## R

Rebiere-Huët, J., Di Martino, P., and Hulen, C. (2004). Inhibition of *Pseudomonas aeruginosa* adhesion to fibronectin by PA-IL and monosaccharides: involvement of a lectin-like process. *Can. J. Microbiol.* 50, 303–312.

Remaut, H., Tang, C., Henderson, N.S., Pinkner, J.S., Wang, T., Hultgren, S.J., Thanassi, D.G., Waksman, G., and Li, H. (2008). Fiber formation across the bacterial outer membrane by the chaperone/usher pathway. *Cell* 133, 640–652.

Reynolds, M., Marradi, M., Imberty, A., Penadés, S., and Pérez, S. (2012). Multivalent gold glycoclusters: high affinity molecular recognition by bacterial lectin PA-IL. *Chemistry* 18, 4264–4273.

Rhim, A.D., Stoykova, L., Glick, M.C., and Scanlin, T.F. (2001). Terminal glycosylation in cystic fibrosis (CF): a review emphasizing the airway epithelial cell. *Glycoconj. J.* 18, 649–659.

Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.L., et al. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 245, 1066–1073.

Roberts, J.A., Marklund, B.I., Ilver, D., Haslam, D., Kaack, M.B., Baskin, G., Louis, M., Möllby, R., Winberg, J., and Normark, S. (1994). The Gal(alpha 1-4)Gal-specific tip adhesin of *Escherichia coli* P-fimbriae is needed for pyelonephritis to occur in the normal urinary tract. *Proc. Natl. Acad. Sci. U.S.A.* 91, 11889–11893.

Rodrigue, A., Quentin, Y., Lazdunski, A., Méjean, V., and Foglino, M. (2000). Two-component systems in *Pseudomonas aeruginosa*: why so many? *Trends Microbiol.* *8*, 498–504.

Römling, U., Schmidt, K.D., and Tümmler, B. (1997). Large chromosomal inversions occur in *Pseudomonas aeruginosa* clone C strains isolated from cystic fibrosis patients. *FEMS Microbiol. Lett.* *150*, 149–156.

Rosenfeld, M., Ramsey, B.W., and Gibson, R.L. (2003). *Pseudomonas* acquisition in young patients with cystic fibrosis: pathophysiology, diagnosis, and management. *Curr Opin Pulm Med* *9*, 492–497.

Ruer, S., Ball, G., Filloux, A., and De Bentzmann, S. (2008). The “P-usher”, a novel protein transporter involved in fimbrial assembly and TpsA secretion. *EMBO J.* *27*, 2669–2680.

## S

Sauer, F.G., Barnhart, M., Choudhury, D., Knight, S.D., Waksman, G., and Hultgren, S.J. (2000). Chaperone-assisted pilus assembly and bacterial attachment. *Curr. Opin. Struct. Biol.* *10*, 548–556.

Sauer, F.G., Fütterer, K., Pinkner, J.S., Dodson, K.W., Hultgren, S.J., and Waksman, G. (1999). Structural basis of chaperone function and pilus biogenesis. *Science* *285*, 1058–1061.

Sauer, F.G., Pinkner, J.S., Waksman, G., and Hultgren, S.J. (2002a). Chaperone priming of pilus subunits facilitates a topological transition that drives fiber formation. *Cell* *111*, 543–551.

Sauer, F.G., Remaut, H., Hultgren, S.J., and Waksman, G. (2004). Fiber assembly by the chaperone-usher pathway. *Biochim. Biophys. Acta* *1694*, 259–267.

Sauer, K., Camper, A.K., Ehrlich, G.D., Costerton, J.W., and Davies, D.G. (2002b). *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J. Bacteriol.* *184*, 1140–1154.

Sharma, V., and Surolia, A. (1997). Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J. Mol. Biol.* *267*, 433–445.

Sharon, N., and Lis, H. (2004). History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* *14*, 53R–62R.

Silby, M.W., Winstanley, C., Godfrey, S.A.C., Levy, S.B., and Jackson, R.W. (2011). *Pseudomonas* genomes: diverse and adaptable. *FEMS Microbiology Reviews* *35*, 652–680.

Sivaneson, M., Mikkelsen, H., Ventre, I., Bordi, C., and Filloux, A. (2011). Two-component regulatory systems in *Pseudomonas aeruginosa*: an intricate network mediating fimbrial and efflux pump gene expression. *Mol. Microbiol.* *79*, 1353–1366.

Smith, J.J., Travis, S.M., Greenberg, E.P., and Welsh, M.J. (1996). Cystic Fibrosis Airway Epithelia Fail to Kill Bacteria Because of Abnormal Airway Surface Fluid. *Cell* *85*, 229–236.

Smyth, D.R., Mrozkiewicz, M.K., McGrath, W.J., Listwan, P., and Kobe, B. (2003). Crystal structures of fusion proteins with large-affinity tags. *Protein Sci.* *12*, 1313–1322.

Spencer, D.H., Kas, A., Smith, E.E., Raymond, C.K., Sims, E.H., Hastings, M., Burns, J.L., Kaul, R., and Olson, M.V. (2003). Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*. *J. Bacteriol.* *185*, 1316–1325.

Spessott, W., Crespo, P.M., Daniotti, J.L., and Maccioni, H.J.F. (2012). Glycosyltransferase complexes improve glycolipid synthesis. *FEBS Lett.* 586, 2346–2350.

## T

Taccetti, G., Campana, S., Festini, F., Mascherini, M., and Döring, G. (2005). Early eradication therapy against *Pseudomonas aeruginosa* in cystic fibrosis patients. *Eur. Respir. J.* 26, 458–461.

Thain, A., Gaston, K., Jenkins, O., and Clarke, A.R. (1996). A method for the separation of GST fusion proteins from co-purifying GroEL. *Trends Genet.* 12, 209–210.

Thanassi, D.G., Saulino, E.T., and Hultgren, S.J. (1998). The chaperone/usher pathway: a major terminal branch of the general secretory pathway. *Curr. Opin. Microbiol.* 1, 223–231.

Tielker, D., Hacker, S., Loris, R., Strathmann, M., Wingender, J., Wilhelm, S., Rosenau, F., and Jaeger, K.-E. (2005). *Pseudomonas aeruginosa* lectin LecB is located in the outer membrane and is involved in biofilm formation. *Microbiology (Reading, Engl.)* 151, 1313–1323.

Le Trong, I., Aprikian, P., Kidd, B.A., Thomas, W.E., Sokurenko, E.V., and Stenkamp, R.E. (2010). Donor strand exchange and conformational changes during *E. coli* fimbrial formation. *J. Struct. Biol.* 172, 380–388.

## U

Uematsu, J., Koyama, A., Takano, S., Ura, Y., Tanemura, M., Kihira, S., Yamamoto, H., Kawano, M., Tsurudome, M., O'Brien, M., et al. (2012). Legume lectins inhibit human parainfluenza virus type 2 infection by interfering with the entry. *Viruses* 4, 1104–1115.

## V

Vallet, I., Diggle, S.P., Stacey, R.E., Cámara, M., Ventre, I., Lory, S., Lazdunski, A., Williams, P., and Filloux, A. (2004). Biofilm formation in *Pseudomonas aeruginosa*: fimbrial cup gene clusters are controlled by the transcriptional regulator MvaT. *J. Bacteriol.* 186, 2880–2890.

Vallet, I., Olson, J.W., Lory, S., Lazdunski, A., and Filloux, A. (2001). The chaperone/usher pathways of *Pseudomonas aeruginosa*: identification of fimbrial gene clusters (cup) and their involvement in biofilm formation. *Proc. Natl. Acad. Sci. U.S.A.* 98, 6911–6916.

Vallet-Gely, I., Donovan, K.E., Fang, R., Joung, J.K., and Dove, S.L. (2005). Repression of phase-variable cup gene expression by H-NS-like proteins in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 11082–11087.

Vallet-Gely, I., Sharp, J.S., and Dove, S.L. (2007). Local and global regulators linking anaerobiosis to cupA fimbrial gene expression in *Pseudomonas aeruginosa*. *J. Bacteriol.* 189, 8667–8676.

Vera, A., González-Montalbán, N., Arís, A., and Villaverde, A. (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnology and Bioengineering* 96, 1101–1106.

Vergani, P., Lockless, S.W., Nairn, A.C., and Gadsby, D.C. (2005). CFTR channel opening by ATP-driven tight dimerization of its nucleotide-binding domains. *Nature* 433, 876–880.

Verma, A., Schirm, M., Arora, S.K., Thibault, P., Logan, S.M., and Ramphal, R. (2006). Glycosylation of b-Type flagellin of *Pseudomonas aeruginosa*: structural and genetic basis. *J. Bacteriol.* *188*, 4395–4403.

Volkan, E., Ford, B.A., Pinkner, J.S., Dodson, K.W., Henderson, N.S., Thanassi, D.G., Waksman, G., and Hultgren, S.J. (2012). Domain activities of PapC usher reveal the mechanism of action of an *Escherichia coli* molecular machine. *Proc. Natl. Acad. Sci. U.S.A.* *109*, 9563–9568.

## W

Waksman, G., and Hultgren, S.J. (2009). Structural biology of the chaperone-usher pathway of pilus biogenesis. *Nat. Rev. Microbiol.* *7*, 765–774.

Walter, T.S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I., and Grimes, J.M. (2006). Lysine methylation as a routine rescue strategy for protein crystallization. *Structure* *14*, 1617–1622.

Wang, Y., Ha, U., Zeng, L., and Jin, S. (2003). Regulation of membrane permeability by a two-component regulatory system in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* *47*, 95–101.

Weldhagen, G.F., Poirel, L., and Nordmann, P. (2003). Ambler class A extended-spectrum beta-lactamases in *Pseudomonas aeruginosa*: novel developments and clinical impact. *Antimicrob. Agents Chemother.* *47*, 2385–2392.

Wellens, A., Lahmann, M., Touaibia, M., Vaucher, J., Oscarson, S., Roy, R., Remaut, H., and Bouckaert, J. (2012). The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. *Biochemistry* *51*, 4790–4799.

Williams, B.J., DehnbosTEL, J., and Blackwell, T.S. (2010). *Pseudomonas aeruginosa*: Host defence in lung diseases. *Respirology* *15*, 1037–1056.

Wright, K.J., and Hultgren, S.J. (2006). Sticky fibers and uropathogenesis: bacterial adhesins in the urinary tract. *Future Microbiol* *1*, 75–87.

Wu, A.M., Wu, J.H., Singh, T., Liu, J.-H., Tsai, M.-S., and Gilboa-Garber, N. (2006). Interactions of the fucose-specific *Pseudomonas aeruginosa* lectin, PA-IIL, with mammalian glycoconjugates bearing polyvalent Lewis(a) and ABH blood group glycotopes. *Biochimie* *88*, 1479–1492.

## Y

Young, N.M., and Oomen, R.P. (1992). Analysis of sequence variation among legume lectins. A ring of hypervariable residues forms the perimeter of the carbohydrate-binding site. *J. Mol. Biol.* *228*, 924–934.



# **Annexe**



# A Lectin from *Platypodium elegans* with Unusual Specificity and Affinity for Asymmetric Complex *N*-Glycans<sup>\*[5]</sup>

Received for publication, April 25, 2012, and in revised form, June 7, 2012. Published, JBC Papers in Press, June 12, 2012, DOI 10.1074/jbc.M112.375816

Raquel Guimarães Benevides<sup>†S1</sup>, Géraldine Ganne<sup>‡</sup>, Rafael da Conceição Simões<sup>S</sup>, Volker Schubert<sup>¶</sup>, Mathäus Niemiets<sup>¶</sup>, Carlo Unverzagt<sup>¶</sup>, Valérie Chazalet<sup>‡</sup>, Christelle Breton<sup>‡</sup>, Annabelle Varrot<sup>‡</sup>, Benildo Sousa Cavada<sup>‡2</sup>, and Anne Imberty<sup>‡3</sup>

From the <sup>†</sup>Centre de Recherche sur les Macromolécules Végétales-CNRS (affiliated with Université Joseph Fourier and Institut de Chimie Moléculaire de Grenoble), 38041 Grenoble, France, the <sup>S</sup>Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, CE, 60.440-970, Brazil, and the <sup>¶</sup>Bioorganische Chemie, Gebäude NW1, Universität Bayreuth, 95440 Bayreuth, Germany

**Background:** Legume lectins with new glycan specificity can be discovered in the *Dalbergieae* tribe.

**Results:** We determined the sequence, specificity, and crystal structures of a new Man/Glc lectin (PELa) from *Platypodium elegans* seeds.

**Conclusion:** The unusual affinity of PELa for asymmetrical complex *N*-glycans is related to the extended binding site and conformational constraints on oligosaccharides.

**Significance:** *Dalbergieae* lectins are of interest for biotechnological applications.

Lectin activity with specificity for mannose and glucose has been detected in the seed of *Platypodium elegans*, a legume plant from the *Dalbergieae* tribe. The gene of *Platypodium elegans* lectin A has been cloned, and the resulting 261-amino acid protein belongs to the legume lectin family with similarity with *Pterocarpus angolensis* agglutinin from the same tribe. The recombinant lectin has been expressed in *Escherichia coli* and refolded from inclusion bodies. Analysis of specificity by glycan array evidenced a very unusual preference for complex type *N*-glycans with asymmetrical branches. A short branch consisting of one mannose residue is preferred on the 6-arm of the *N*-glycan, whereas extensions by GlcNAc, Gal, and NeuAc are favorable on the 3-arm. Affinities have been obtained by microcalorimetry using symmetrical and asymmetrical Asn-linked heptasaccharides prepared by the semi-synthetic method. Strong affinity with  $K_d$  of 4.5  $\mu\text{M}$  was obtained for both ligands. Crystal structures of *Platypodium elegans* lectin A complexed with branched trimannose and symmetrical complex-type Asn-linked heptasaccharide have been solved at 2.1 and 1.65 Å resolution, respectively. The lectin adopts the canonical dimeric organization of legume lectins. The trimannose bridges the binding sites of two neighboring dimers, resulting in the formation of infinite chains in the crystal. The Asn-linked heptasaccharide binds with the 6-arm in the primary binding site with extensive additional contacts on both arms. The GlcNAc on the 6-arm is bound in a constrained conformation that may rationalize the higher affinity observed on the glycan array for *N*-glycans with only a mannose on the 6-arm.

Legume lectins compose a large family of homologous carbohydrate-binding proteins that are found primarily in the seeds of legume plants. Despite similarities in their physicochemical properties and their relatively conserved primary sequences, *Leguminosae* lectins display considerable diversity in their carbohydrate binding specificities, not only in terms of the recognition of monosaccharides but also in terms of differential binding to complex carbohydrates (1). Legume lectins are central to the study of the molecular basis and specificity of protein-carbohydrate interactions (2), and they also have technological implications for the understanding of cell-cell recognition, adhesion, tumor spread, bacterial and viral infection, and inflammation (3). However, the function of plant lectins remains elusive (4), although they are generally proposed to play a role in defense mechanisms against pathogens and invertebrate or vertebrate predators (5) or in the establishment of symbiosis (6). The mechanism of action still remains unclear even though induced response mechanisms are proposed for many pathogen-mediated injuries in plants. The direct interference with viruses and microorganisms is rather exceptional, but the deleterious effects of plant lectins on both predatory invertebrates and animals are well documented (5). The existence of hydrophobic sites within the structure of leguminous lectins, which bind phytohormones, suggests a possible role in certain aspects of hormonally regulated plant growth and development (7).

Leguminous plant lectins, although differing in carbohydrate-binding specificity, have similar amino acid sequences (1) and similar fold at the monomer level (8). The architecture of the so-called “legume lectin fold” is structurally related to a jelly roll motif characterized by a  $\beta$ -sandwich of 25–30 kDa that contains a carbohydrate-combining site, a tightly bound  $\text{Ca}^{2+}$ , and a transition metal ion, usually  $\text{Mn}^{2+}$ . Variability occurs both at the level of the quaternary fold, with a variety of dimeric and tetrameric arrangement (9), and at the level of the binding site. The floor of the binding site consists of few conserved key

\* This work was supported in part by CNRS and the French Ministry of Research and the German Research Foundation (DFG).

[5] This article contains supplemental Figs. S1–S4 and Tables S1–S3. The atomic coordinates and structure factors (codes 3ZVX and 3ZYR) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

<sup>1</sup> Recipient of CAPES and CNPq grants from the Brazilian government.

<sup>2</sup> To whom correspondence may be addressed. E-mail: bscavada@ufc.br.

<sup>3</sup> To whom correspondence may be addressed. E-mail: imberty@cervat.cnrs.fr.

amino acids, but its perimeter depends on five loops with very variable size and sequence that control the carbohydrate specificity (8, 10, 11).

Lectins isolated from legume seeds have been regarded as a paradigm for investigating the structural basis and thermodynamics of selective sugar recognition (12, 13). The interaction between legume lectin and carbohydrate is generally characterized by a low affinity for monosaccharides (mM range). Longer oligosaccharides exhibit micromolar affinity, generally by generating a multivalent effect either by binding several sites of a multivalent lectin or by cross-linked neighboring lectins in solution (14, 15).

The large majority of leguminous lectins have been isolated and characterized from plants belonging to the *Phaseoleae* and *Vicieae* tribes of the *Papilionoideae* subfamily of Leguminosae (16). In this context, it would be of interest to characterize the structure and fine specificity of other lectins in different tribes to extend our knowledge on protein-carbohydrate interactions. The *Dalbergieae* tribe is of particular interest because only one lectin has been structurally characterized, *i.e.* the agglutinin from *Pterocarpus angolensis* seed (PAL).<sup>4</sup> Crystallography and thermodynamic studies demonstrated high affinity for both oligomannose *N*-glycans and for biantennary complex type, correlated with the presence of an unusually extended binding region close to the primary binding site (17–20).

In this study, we characterize the lectin from *Platypodium elegans* Vogel, another member of the *Dalbergieae* tribe. The genus *Platypodium* comprises two species of leguminous trees. *P. elegans* Vogel is a natural species found in the Brazil cerrado and in transition areas between cerrado and seasonal forest (21). We demonstrate here the presence of a mannose/glucose lectin (PELa) in *P. elegans* seeds. The cloning of the gene has been performed, resulting in the amino acid sequence determination of PELa. From the recombinant lectin, a complete study of the structure-specificity relationship has been performed, and we describe here the molecular basis of the unusual affinity of PELa for the asymmetric complex *N*-glycans.

## EXPERIMENTAL PROCEDURES

### Materials

*P. elegans* seeds were purchased from Reserva de Reflorestamento Flora Tietê (São Paulo, Brazil) and were stored at room temperature or at  $-70^{\circ}\text{C}$ . Monosaccharides, disaccharides, sulfated polysaccharides (Fucoidan, Carrageenan), and glycoprotein (pork stomach mucin) were purchased from Sigma.

### Synthesis of Asn-linked *N*-Glycans H, E, and C

**Synthesis of H**—Nonasaccharide-Asn **1** was prepared from egg yolk by enzymatic hydrolysis of its disialylated derivative (22). 10.0 mg (5.7  $\mu\text{mol}$ ) of nonasaccharide-Asn **1** were dissolved in 450  $\mu\text{l}$  of phosphate buffer (100 mM, pH 6.8, 30 mM

$\text{MgCl}_2$ , 100 mM  $\beta$ -mercaptoethanol). To the solution were added 50  $\mu\text{l}$  of buffered BSA (10 mg BSA/ml, 20 mM  $\text{KH}_2\text{PO}_4$ , 50 mM NaCl, 0.1 mM EDTA, pH 7). To this mixture 1 unit of lyophilized  $\beta$ -galactosidase (EC 3.2.1.23) was added. The mixture was incubated at  $37^{\circ}\text{C}$  for 3 days (TLC: 2-propanol, 1 M ammonium acetate, 1.5:1). After lyophilization, the residue was dissolved in 500  $\mu\text{l}$  of 0.1 M  $\text{NH}_4\text{HCO}_3$  and purified by gel filtration (Superdex 30,  $1.6 \times 60$  cm, 0.1 M  $\text{NH}_4\text{HCO}_3$ , 1 ml  $\text{min}^{-1}$ ). The peak eluting at 80 min was lyophilized and desalted by gel filtration (Sephadex G-25,  $2.5 \times 70$  cm, 5% ethanol in water, 0.5 ml  $\text{min}^{-1}$ ). The peak eluting at 427 min was lyophilized yielding 5.2 mg of heptasaccharide-Asn **H** (64.1%). The purity was confirmed by 360 MHz  $^1\text{H}$  NMR in  $\text{D}_2\text{O}$ .

**Synthesis of E**—The Fmoc-protected octasaccharide-Asn **2** was isolated by HPLC according to Kajihara (22). 15.0 mg (8.3  $\mu\text{mol}$ ) of Fmoc-Asn(octasaccharide) **2** were dissolved in 320  $\mu\text{l}$  of HEPES buffer (50 mM, pH 6.0, 0.1 mg  $\text{ml}^{-1}$  BSA). 17  $\mu\text{l}$  (0.9 units) of a suspension of jack bean  $\beta$ -*N*-acetylglucosaminidase in 2.5 M  $(\text{NH}_4)_2\text{SO}_4$ , pH 7.0, were dissolved in HEPES buffer (50 mM, pH 6.0, 0.1 mg  $\text{ml}^{-1}$  BSA). The solutions of oligosaccharide **2** and the enzyme were combined, and the mixture was incubated at  $37^{\circ}\text{C}$  for 3 days. After completion (LC-MS, C18, 20–35% MeCN (0.1% HCOOH)) the reaction was lyophilized. The residue was dissolved in 1 ml of 10% MeCN (0.1% HCOOH) and purified by RP-HPLC (NUCLEOGEL RP 100–10 ( $2.5 \times 30$  cm), 20–35% MeCN, 10 ml  $\text{min}^{-1}$ ) The peak eluting at 29.3 min was lyophilized to yield 10.3 mg of Fmoc-Asn(heptasaccharide) **3** (77.3%). The purity was confirmed by LC-MS (C18, 20–35% MeCN (0.1% HCOOH)).

9.0 mg (5.6  $\mu\text{mol}$ ) of Fmoc-Asn(heptasaccharide) **3** were dissolved in a mixture of water, methanol, and acetonitrile (10:3:5). Subsequently, 89  $\mu\text{l}$  (44.5  $\mu\text{mol}$ ) of sodium hydroxide (0.5 M) were added, and the solution was stirred for 3 h at ambient temperature. After completion (TLC: 2-propanol, 1 M ammonium acetate, 1.5:1), 850  $\mu\text{l}$  of ammonium bicarbonate (0.1 M) were added, and the mixture was purified by gel filtration (Superdex 30,  $1.6 \times 60$  cm, 0.1 M ammonium bicarbonate, 0.75 ml/min). The peak eluting at 108.2 min was lyophilized and desalted by gel filtration (Sephadex G-25 ( $2.5 \times 71$  cm), 5% ethanol in water, 0.5 ml  $\text{min}^{-1}$ ). The fractions eluting at 422 min were lyophilized. Yield was 7.1 mg of **E** (5.1  $\mu\text{mol}$ , 91.4%);  $R_f$  (**E**) = 0.31 (2-propanol, 1 M ammonium acetate, 1.5:1);  $[\alpha]_D^{23} = +1.0$  (0.1, water);  $\text{C}_{52}\text{H}_{87}\text{N}_5\text{O}_{38}$  (1390.26), ESI-MS negative mode:  $M_{\text{calc}} = 1389.50$ ;  $M_{\text{found}} = 1387.92$  (M-H) $^-$ . The purity was confirmed by 360-MHz  $^1\text{H}$  NMR in  $\text{D}_2\text{O}$ .

**Synthesis of C**—11.0 mg (7.7  $\mu\text{mol}$ ) of heptasaccharide-Asn **H** were dissolved in 380  $\mu\text{l}$  of HEPES buffer (50 mM, pH 6.0, 0.4 mg  $\text{ml}^{-1}$  BSA). 15.2  $\mu\text{l}$  (0.81 units) of a suspension of jack bean  $\beta$ -*N*-acetylglucosaminidase (EC 3.2.1.52) in 2.5 M  $(\text{NH}_4)_2\text{SO}_4$ , pH 7.0, were dissolved in 50  $\mu\text{l}$  of HEPES buffer (50 mM, pH 6.0, 0.4 mg  $\text{ml}^{-1}$  BSA). The solution of oligosaccharide **H** was added to the buffered enzyme, and the mixture was incubated at  $37^{\circ}\text{C}$  for 5 days. After completion (ESI-MS), the reaction was lyophilized. The residue was dissolved in 0.3 ml of 0.1 M  $\text{NH}_4\text{HCO}_3$  and purified by gel filtration (Superdex 30 ( $2.6 \times 60$  cm), 0.1 M  $\text{NH}_4\text{HCO}_3$ , 1.5 ml  $\text{min}^{-1}$ ). The peak eluting at 149 min was lyophilized and desalted by gel filtration (Sephadex G-25 ( $2.5 \times 71$  cm), 5% ethanol in water, 0.5 ml  $\text{min}^{-1}$ ). The elution at 451

<sup>4</sup> The abbreviations used are: PAL, *P. angolensis* lectin; ConA, concanavalin A; DGL, *D. grandiflora* lectin; ITC, isothermal titration calorimetry; LacNAC, *N*-acetyl-lactosamine; PELa, *P. elegans* lectin A; PSL, *P. sativum* lectin; Tri-Man, trimannose  $\alpha\text{Man}1-3(\alpha\text{Man}1-6)\text{Man}$ ; RACE, rapid amplification of cDNA ends; Bicine, *N,N*-bis(2-hydroxyethyl)glycine; Fmoc, *N*-(9-fluorenyl)methoxycarbonyl.

## N-Glycan Binding by *Platypodium elegans* Lectin

min was lyophilized to yield 3.7 mg of pentasaccharide-Asn C (46.8%). The purity was checked by 360 MHz  $^1\text{H}$  NMR in  $\text{D}_2\text{O}$ .

### Hemagglutination Tests on Seed Extracts

The seeds were ground and later defatted in the presence of *n*-hexane. A protein extract was prepared from 1 g of the ground seeds with 10 ml of 0.15 M NaCl at room temperature for 3 h. The homogenate was centrifuged at  $16,000 \times g$  for 20 min at 4 °C, and the precipitate was discarded. The hemagglutinating activity of the lectin was determined by a 2-fold serial dilution procedure using native rabbit or human erythrocytes as described previously (23). The hemagglutination titer was defined as the reciprocal of the highest dilution exhibiting observable hemagglutination. Inhibition of agglutination by sugar was assayed by serially diluting the solution of sugar in the microtiter plate, followed by the addition of four agglutinin units of the lectin and, after 30 min, of erythrocyte suspension. The lowest concentration of sugar that visibly decreased the extent of agglutination was defined as the minimum inhibitory concentration.

### Cloning of PELa and Sequencing

For RNA isolation, 1 g of seeds was ground to a powder with a pestle under liquid nitrogen. Total cellular RNA was isolated using the Concert<sup>TM</sup> plant RNA reagent (Invitrogen), according to the manufacturer's protocol. Using the 3'-RACE kit from Invitrogen, mRNAs were converted into cDNA using reverse transcriptase and an oligo(dT) adapter primer. Partial cDNA of *P. elegans* seed lectin (PELa) was then obtained by PCR using a gene-specific primer (PLAT2, 5'-CAYATHGGNATHGAYGTNAA-3') in combination with the abridged adapter primer of the 3'-RACE kit. PLAT2 was designed for minimal degeneration from a highly conserved peptide region (HIGIDVN) corresponding to the metal-binding site of legume lectins. The partial cDNA obtained was cloned into pGEM<sup>®</sup>T-easy (Promega) and sequenced. Complete coding region of PELa was obtained using the 5'-RACE kit (Invitrogen). mRNAs were first converted into cDNA using the antisense gene-specific primer PLAT3 (5'-TGATATAGAGATCCCTCTGGA-3'), which annealed downstream from the stop codon of PELa, and cDNA was then amplified using PLAT3 and the anchor primer of the 5'-RACE kit, after a TdT tailing step, and as described by the manufacturer. The PCR fragment was purified, cloned into the pGEM<sup>®</sup>T-easy, and further sequenced. All PCRs were performed using the GoTaq<sup>®</sup> DNA polymerase (Promega). The complete coding DNA sequence of PELa was then subcloned into the pET-29a plasmid (Invitrogen). PCR was performed using the 5'-RACE product as template and the sense PLAT4 (5'-GGAATTCATATGCTACTGAATAGGGCAAAC-3') and antisense PLAT5 (5'-GGCAAGCTTTCACATATCACGTGCAAGATAC-3') primers, containing NdeI and HindIII restriction sites, respectively (underlined sequences). After digestion with NdeI and HindIII, the amplified fragment was introduced into the expression plasmid. The resulting plasmid was named pET-29a\_PELa.

### Expression and Purification of Recombinant PELa

*Escherichia coli* BL21(DE3) cells containing the plasmid pET-29a\_PELa were cultured in LB broth medium at 37 °C. When the culture reached an  $A_{600}$  of 0.5–0.6, isopropyl  $\beta$ -D-thiogalactoside was added to a final concentration of 0.5 mM. Cells were harvested after 4 h of incubation at 30 °C, washed, and suspended in buffer A (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM  $\text{CaCl}_2$ , 5 mM  $\text{MnCl}_2$ ). The cells were broken by cell disruption (Constant Cell Disruption System) at 1700 bars. After centrifugation at  $20,000 \times g$  for 30 min at 4 °C and filtration, the pellet was washed twice with 50 mM Tris-HCl, pH 8.0, 4 M urea, 0.5 M NaCl, 1 mM EDTA, 0.5% Triton X-100. The final precipitate was denatured and solubilized in 50 mM Tris-HCl, pH 8.0, 8 M urea, 10 mM DTT and, after centrifugation at  $20,000 \times g$  for 30 min at 4 °C, had its protein concentration adjusted to  $1 \text{ mg ml}^{-1}$ . 50 ml of the denatured protein solution at  $1 \text{ mg ml}^{-1}$  was slowly diluted in 950 ml of 50 mM Tris-HCl, pH 7.5, 240 mM NaCl, 2 mM  $\text{MnCl}_2$ , 2 mM  $\text{CaCl}_2$ , 10 mM KCl, 1 mM DTT. After 1 h at 10 °C and filtration, the fresh refolded protein was submitted to the purification step.

The purification of PELa was performed on a D-mannose-Sepharose column (Sigma). PELa refolded was allowed to bind to the immobilized mannose in equilibrating buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM  $\text{CaCl}_2$ , 5 mM  $\text{MnCl}_2$ ). After washing with buffer A, PELa was eluted using buffer A supplemented with 200 mM D-mannose. The purified protein was washed and concentrated as desired using a Vivaspin (Sartorius) ultrafiltration device with a cutoff of 10 kDa and stored at 4 °C. The protein fractions were analyzed on 12% SDS-PAGE, and the protein was stained with Brilliant Blue staining (Sigma).

### Isothermal Titration Microcalorimetry Analysis

ITC experiments were performed with a VP-ITC isothermal titration microcalorimeter (Microcal; GE Healthcare). Experiments were carried out at  $25 \pm 0.1$  °C. Protein was prepared in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM  $\text{CaCl}_2$ , 5 mM  $\text{MnCl}_2$ , and the same buffer was used to prepare the oligosaccharide solutions. Protein (monomer) concentration used for ITC was 0.06 mM. Injections of 10  $\mu\text{l}$  of each oligosaccharide solution at various concentrations (from 0.4 to 0.8 mM) were added automatically to the protein solution present in the calorimeter cell. Control experiments performed by injection of buffer into the protein solution yielded insignificant heat of dilution. Integrated heat effects were analyzed by nonlinear regression using a single site binding model (Origin 7.0) after subtracting the saturation heat evaluated at the end of the titration. The experimental data fitted to a theoretical titration curve gave the association constant ( $K_a$ ) and the enthalpy of binding ( $\Delta H$ ). The other thermodynamic parameters such as changes in free energy ( $\Delta G$ ) and entropy ( $\Delta S$ ) were calculated from the equation  $\Delta G = \Delta H - T\Delta S = -RT \ln K_a$ , where  $T$  is the absolute temperature, and  $R$  is the molar gas constant ( $8.314 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ ). All experiments were performed with  $c$  values of  $10 < c < 100$ , with  $c$  being the Wiseman constant (24) defined as  $c = nK_a [M]$  ( $[M]$  receptor concentration in calorimeter cell).

TABLE 1

Data collection and refinement statistics

	TriMan	Compound H
Beamline (wavelength)	Id14EH1/0.933 Å	BM30A/0.9797 Å
Space group	P2 <sub>1</sub> 2 <sub>1</sub>	P2 <sub>1</sub> 2 <sub>1</sub>
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i>	42.62 80.68 148.76 Å	51.40 76.98 125.59 Å
$\alpha$ , $\beta$ , $\gamma$	90.0, 90.0, 90.0°	90.0, 90.0, 90.0°
Resolution (outer shell)	42.60 to 2.10 (2.21 to 2.10 Å)	47.57 to 1.65 (1.74 to 1.65 Å)
Measured/unique reflections	99,535/28,894	186,416/55,022
Average multiplicity	3.4 (2.4)	3.4 (2.6)
$R_{\text{merge}}$	0.082 (0.330)	0.054 (0.250)
Completeness	93.7% (70.1%)	91.4% (65.8%)
Mean $I/\sigma I$	9.3 (2.0)	13.6 (3.9)
Wilson B	24.7	14.00
$R_{\text{cryst}}/R_{\text{free}}$	19.5/25.2	15.5/18.8
r.m.s.d. bonds <sup>a</sup>	0.015 Å	0.015 Å
r.m.s.d. angles	1.54°	1.59°
r.m.s.d. chiral	0.106 Å <sup>3</sup>	0.098 Å <sup>3</sup>
Protein atoms	1935/1926	2004/1986
Bfac	23.5/23.7 Å <sup>2</sup>	12.4/14.0 Å <sup>2</sup>
Water molecules	177/149	295/223
Bfac	29.0/28.3 Å <sup>2</sup>	25.7/26.6 Å <sup>2</sup>
Ligand atoms	34	89/92
Bfac	22.6 Å <sup>2</sup>	14.2/16.6 Å <sup>2</sup>
Hetero atoms	18/8	30/12
Bfac, Å <sup>2</sup>	32.4/33.5	24.7/25.7
Protein Data Bank code	3ZVX	3ZYR

<sup>a</sup> r.m.s.d. is root mean square deviation.

### Glycan Microarray Analysis

Purified PELA sample was labeled with Alexa Fluor 488 (Invitrogen) according to the manufacturer's instructions and re-purified on a D-Salt polyacrylamide desalting column (Pierce). Alexa-labeled protein was used for glycan array screening with a standard procedure of the protein-glycan interaction core (H) of the Consortium for Functional Glycomics.

### Crystallization and Structure Determination

PELA was concentrated in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 mM CaCl<sub>2</sub>, 5 mM MnCl<sub>2</sub> to 10 mg ml<sup>-1</sup> with addition of TriMan or compound H at the final concentration of 1 mM. Crystallization conditions for PELA were screened using the hanging-drop vapor diffusion method with a commercially available crystallization solution (Hampton Research Screens I and II, Hampton Research, Riverside, CA) at 20 °C. Drops were prepared by mixing a 1.0- $\mu$ l sample solution and 1.0- $\mu$ l reservoir solution, followed by equilibration against a 500- $\mu$ l reservoir solution.

Crystals of the PELA-TriMan and PELA-heptasaccharide complexes were obtained using crystallization conditions 42 (1.5 M ammonium sulfate, 0.1 M Tris-HCl, pH 8.5, 12% v/v glycerol) and 48 (100 mM Bicine, pH 9, 10% PEG 20K and 2% dioxane) of the Crystal Screen II kit (Hampton Research), respectively. Crystals were frozen in liquid nitrogen (100 K) after soaking them for a time as short as possible in 25 and 30% (v/v) glycerol in precipitant solution, respectively. Data sets to 2.2 and 1.6 Å were collected for the PELA-TriMan and PELA-heptasaccharide complexes at European Synchrotron Radiation Facility (ESRF), Grenoble, France, on beamlines ID14-2 and BM30A, respectively, using an ADSC Q4R CCD (charge-coupled device) detector (Quantum Corp.). The statistics for both data collections are given in Table 1.

Phasing information was obtained using a monomer of *P. angolensis* lectin (Protein Data Bank code 1N3O, with removal

TABLE 2

Inhibition of hemagglutination activity of *P. elegans* seeds crude extract by various sugars

Sugars	Minimum inhibitory concentration
D-Mannose	12.5 mM
Methyl- $\alpha$ -D-mannopyranoside	6.12 mM
$\alpha$ Man1-2 Man	1.56 mM
$\alpha$ Man1-3 Man	1.56 mM
$\alpha$ Man1-6 Man	12.5 mM
D-Glucose	50 mM
Methyl- $\alpha$ -D-glucopyranoside	12.5 mM
Maltose	50 mM
N-Acetyl-D-glucosamine	25 mM
D-Galactose	— <sup>a</sup>
N-Acetyl-D-galactosamine	—
L-Fucose	—
D-Arabinose	—
Lactose	—
Lactulose	—
Fuicoidan	—
Carrageenan	—
Pork stomach mucin	—

<sup>a</sup> No inhibition was observed at concentration of 100 mM.

of water molecules) as a probe for molecular replacement for the orthorhombic crystal form of the PELA-TriMan complex using the program PHASER (25). For the PELA-heptasaccharide complex, the crystal structure of the PELA-TriMan complex was used as template. Crystallographic refinement was carried out with the program REFMAC (26), and manual model building was achieved using Coot (27). Refinement statistics of both structures are given in Table 1. Coordinates have been deposited at the Protein Data Bank under code 3ZVX and 3ZYR, respectively. Molecular drawings were prepared using PyMOL Molecular Graphics System (DeLano Scientific, Palo Alto, CA) and LIGPLOT (28).

## RESULTS AND DISCUSSION

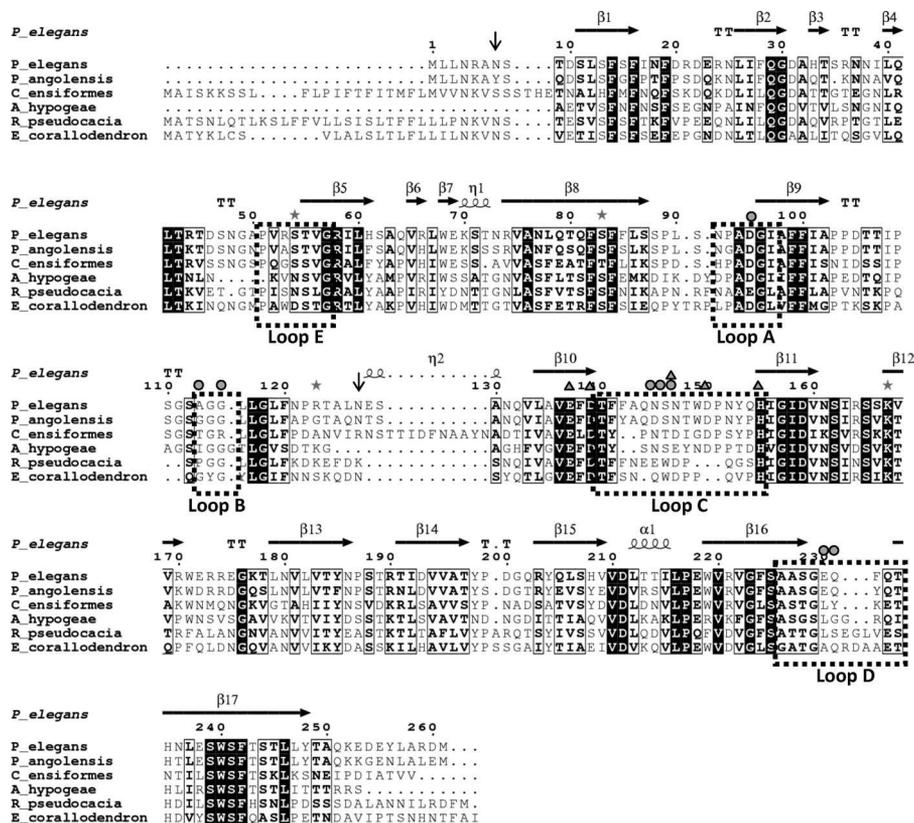
### Identification of Lectin Activity in *P. elegans* Seed Extracts

Crude extract from *P. elegans* seeds agglutinated rabbit erythrocytes with a titer of 2098 hemagglutination units ml<sup>-1</sup>. Human erythrocytes with blood group A, B, and O were also tested, but no hemagglutination was observed. These results indicate the expression of at least one lectin in seeds with specificity for the oligosaccharides present in rabbit erythrocytes. Several monosaccharides and disaccharides were tested for their ability to inhibit PELA-induced rabbit erythrocyte agglutination as shown in Table 2. Glucose and its derivatives  $\alpha$ -methylglucoside and GlcNAc demonstrated some inhibition, albeit at a rather large concentration. The most efficient compound is mannose, which could inhibit the agglutination at a concentration of 12 mM. Methyl- $\alpha$ -mannoside- and  $\alpha$ Man-containing disaccharides are better ligands. These results indicate the presence of a mannose/glucose-binding lectin in the crude extract from *P. elegans* seeds.

### RNA Isolation and cDNA Cloning

Total mRNAs were extracted from the seeds of *P. elegans*, and amplification of cDNA ends (3'-RACE and 5'-RACE) was performed using degenerated and specific primers. The specific degenerated primers were designed based on conserved peptide sequences known to be involved in binding to calcium and

# N-Glycan Binding by *Platypodium elegans* Lectin



**FIGURE 1. Aligned amino acid sequences of legume lectins.** Multiple alignment and secondary structure predictions were achieved using the webESPrpt program (40). The sequence data were obtained from the Entrez protein sequence data base. ID numbers are as follows: AEK69351 (*P. elegans* lectin), CAD19804 (*P. angolensis* lectin), 2PEL\_A (*A. hypogaea* lectin), CAA25787 (*C. ensiformes* lectin), BAA36415 (*Robinia pseudocacia* lectin), and CAA36986 (*E. coraliodendron* lectin). Symbols are as follows:  $\alpha$ -helices ( $\alpha$ ),  $3_{10}$  helices ( $\eta$ ),  $\beta$ -strands ( $\beta$ ), strict  $\alpha$ -turns (TTT),  $\beta$ -turns (TT), sugar binding residues (circles), metal-binding residues (triangles), and putative *N*-glycosylation sites ( $\downarrow$ ).

manganese in legume lectins (supplemental Figs. 1S and 2S and supplemental Table 1S).<sup>5</sup>

The amino acid sequence deduced from the cDNA sequence indicates that the protein includes 261 amino acids with two predicted *N*-glycosylation sites (Asn-7 and Asn-127). PELa sequence shows similarities with other Man/Glc-specific legume lectins, such as those from *P. angolensis* (72% identity), *Pterocarpus rotundifolius* (82%), and *Arachis hypogaea* (70%) (ID numbers CAD19804, AAT57665, and AAA74572, respectively). Multiple alignment with a set of legume lectins with known crystal structures is displayed in Fig. 1. Sequence identities vary from 72% for the Man/Glc-specific lectin from *P. angolensis* to 39% for the Gal/GalNAc lectin from *Erythrina coraliodendron*. The five loops that have been described to build the carbohydrate-binding sites (8, 10) are highlighted in Fig. 1. Sequences in these loops present strong similarities with *P. angolensis* lectin, but weaker similarities with other lectins, in agreement with the large differences in monosaccharide specificities.

## Expression and Purification of Recombinant PELa

Production of recombinant PELa in *E. coli* BL21(DE3) strain resulted in the accumulation of the protein in inclusion bodies. A renaturation protocol, including solubilization by 8 M urea,

was used to obtain the protein in the soluble form (supplemental Fig. 3S). Active PELa was obtained by affinity purification on a mannose-agarose column with a final yield of 5 mg of PELa/liter of cell culture. The soluble lectin displays a molecular mass of 29 kDa, in agreement with the theoretical one (29.287 kDa). The activity of the lectin was checked at each step of production and purification by hemagglutination assay.

## Specificity of PELa

The lectin specificity was analyzed by glycan array (Consortium for Functional Glycomics) that allows the screening of a broad spectrum of natural and synthetic glycans. The printed array version 4.1 containing 465 natural and synthetic glycans was used, and PELa was observed to bind to a limited number of oligosaccharides (Fig. 2A). Short oligosaccharides are not bound with the exception of linear or branched oligomannose fragments (compounds A and B). The lectin is specific for *N*-glycans. Oligomannose-type *N*-glycans with terminal nonreducing  $\alpha$ -mannose are high affinity ligands (compounds C and D). However, the lectin displays a clear preference for asymmetrical complex-type *N*-glycans with a short 6-arm, containing only the  $\alpha$ Man residue, and long 3-arm with *N*-acetyl-lactosamine (LacNAc) (compound E) or sialylated LacNAc (compound F) on the  $\alpha$ Man residue. Symmetrical complex *N*-glycans are not bound in the same way. The truncated complex *N*-glycan terminated by two GlcNAc residues (compound H) is recognized, albeit more weakly than the asymmetrical

<sup>5</sup> RNA and protein sequences have been deposited GenBank with codes JN133278 and AEK69351, respectively.

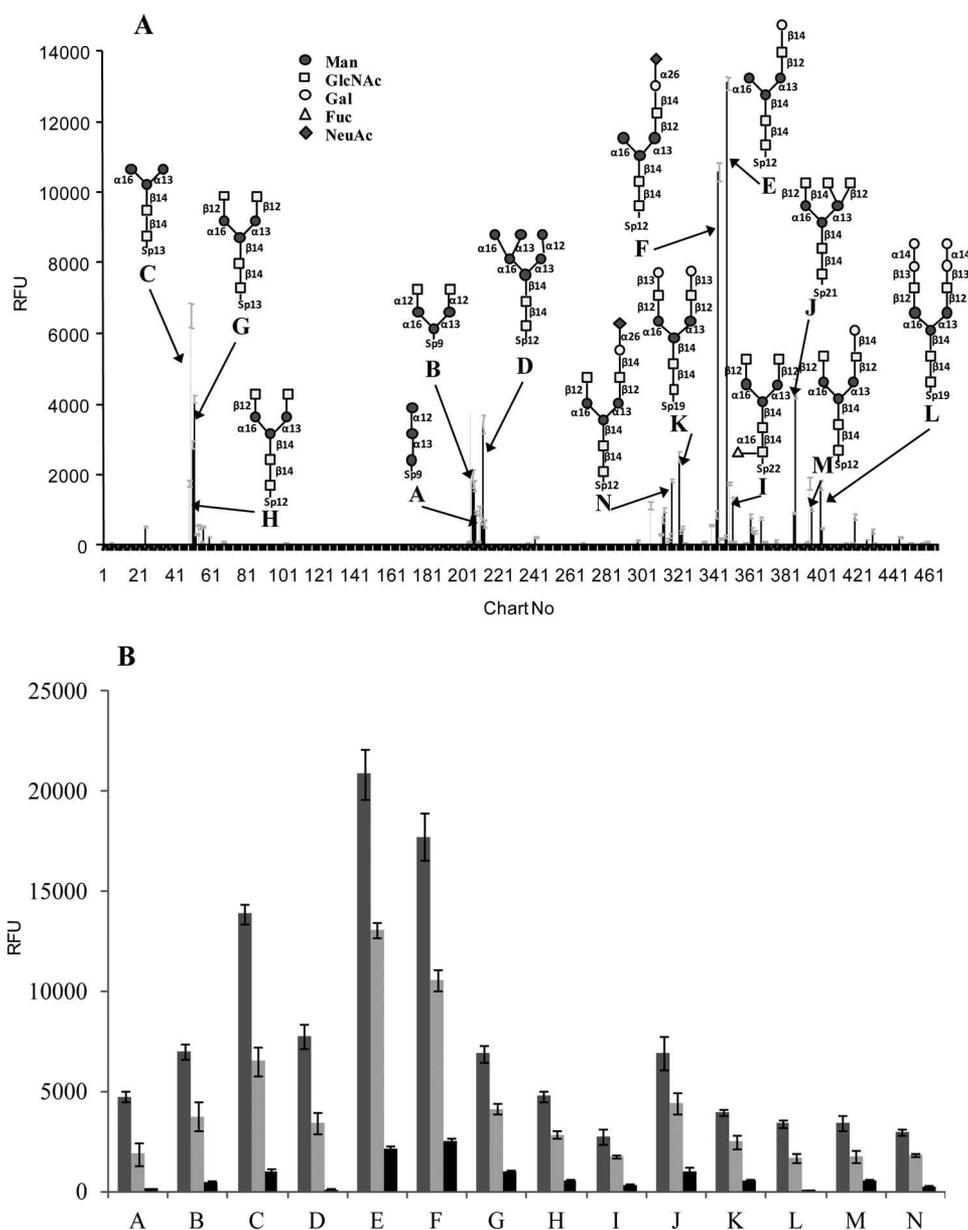


FIGURE 2. **Glycan array data.** *A*, binding of labeled PELA ( $10 \mu\text{g ml}^{-1}$ ) to all carbohydrates of the glycan array version 4.1. *B*, comparison for binding at three lectin concentrations (medium gray,  $50 \mu\text{g ml}^{-1}$ ; light gray,  $10 \mu\text{g ml}^{-1}$ ; dark gray,  $1 \mu\text{g ml}^{-1}$ ).

compounds. In contrast, the more complete complex *N*-glycans terminating with *N*-acetylglucosamine (LacNAc,  $\beta\text{Gal1-4GlcNAc}$ ) or sialyl-LacNAc are not recognized. Interestingly, the same compounds with a  $\beta\text{1-3}$  linkage between Gal and GlcNAc are bound (compounds **K** and **L**), albeit less strongly than the asymmetrical compounds.

Glycan array results allowed a semi-quantitative analysis by reporting the intensities of fluorescence at different protein concentrations. As displayed in Fig. 2*B*, the preference for asymmetrical complex *N*-glycans is conserved for the three PELA concentrations tested ( $50$ ,  $10$ , and  $1 \mu\text{g ml}^{-1}$ ) of lectin. The influence of the terminal sialic acid on the 3-arm seems to be moderate because oligosaccharides **E** and **F** have a rather equivalent response at low protein concentration. Comparing the concentration-dependent fluorescence response between oligosaccharides **E** and **M**, and also **F** and **N**, clearly confirms

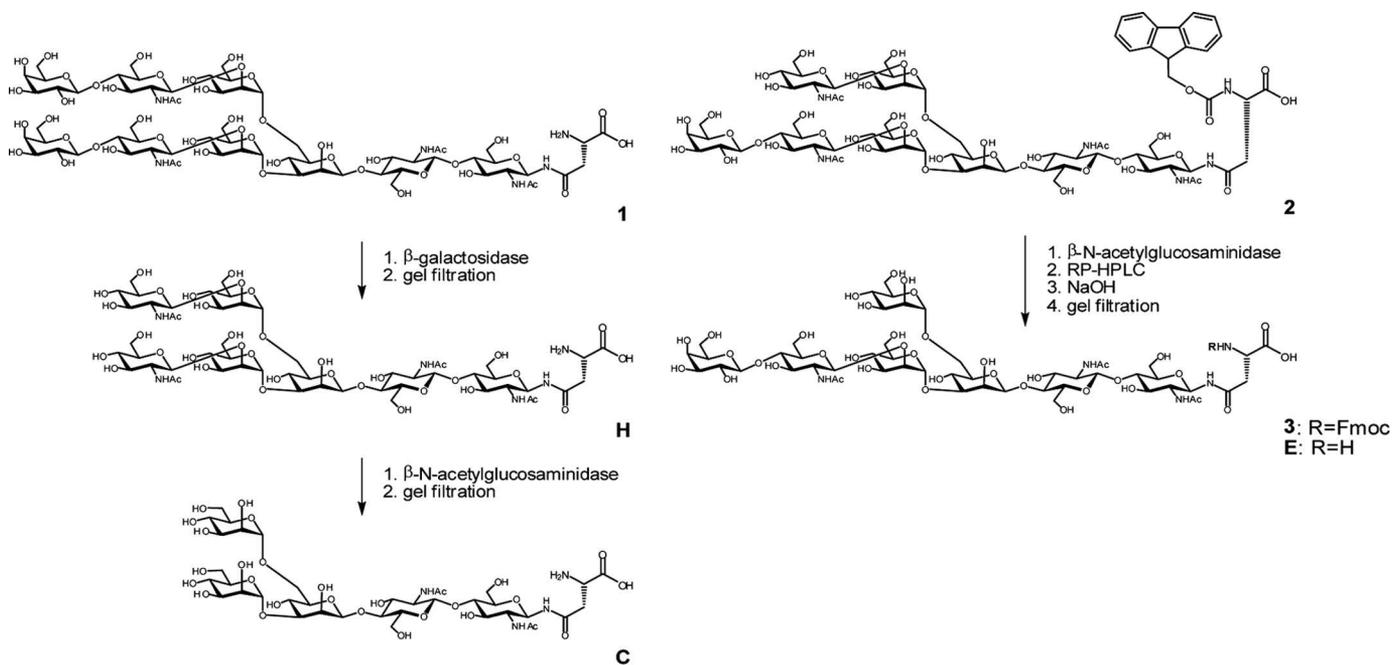
that the addition of one GlcNAc on the 6-arm significantly decreases the lectin affinity.

Glycan array data for other legume lectins are available online. The classical Glc/Man legume lectins such as ConA from *Canavalia ensiformis* and lectin from *Pisum sativum* bind preferentially to oligomannose-type *N*-glycans, more weakly to complex *N*-glycan, and not much to asymmetrical *N*-glycans. These data therefore demonstrate that the unusual specificity of PELA for asymmetrical glycans is clearly different from the classical Glc/Man legume lectins.

#### Synthesis of Symmetrical and Asymmetrical Asn-linked *N*-Glycans

The *N*-glycan asparagines **H** and **E** were both obtained by chemical and enzymatic conversion of a sialoglycopeptide isolated from egg yolk (Scheme 1) (29). The nonasaccharide-Asn **I**

## N-Glycan Binding by *Platypodium elegans* Lectin



SCHEME 1. Schematic representation of glycopeptide synthesis.

(22) was prepared from the Fmoc-protected precursor according to a previously disclosed procedure (30). Removal of both terminal galactosyl moieties was accomplished by digestion with *E. coli*  $\beta$ -galactosidase followed by purification and desalting via gel filtration. The heptasaccharide-Asn **H** (22, 31) was obtained in good yield and high purity. For the preparation of asymmetrical compound **E**, the purification of a suitable precursor was carried out using a mixture of Fmoc-protected *N*-glycan asparagines. According to Kajihara (22), the separation of Fmoc octasaccharide-Asn **2** (terminal galactoside on the 3-arm) from its isomer (galactoside on the 6-arm) can be performed by preparative HPLC. Compound **2** was digested with jack bean  $\beta$ -*N*-acetylglucosaminidase and purified by RP-HPLC yielding as an intermediate the Fmoc-protected heptasaccharide-Asn **3**. Alkaline cleavage of the Fmoc group led to the target compound **E**. The pentasaccharide-Asn **C** was obtained from heptasaccharide-Asn **H** using jack bean  $\beta$ -*N*-acetylglucosaminidase. The  $^1\text{H}$  NMR resonances of the sugar part of compounds **1**, **2**, **3**, **H**, **E**, and **C** (supplemental Fig. 4S) were in accordance with the published data (22). Depending on the pH and on counter ions, the resonances of the Asn moiety may vary.

### Affinity Measurements by Titration Microcalorimetry

The affinity between PELa and the three Asn-linked oligosaccharides **C**, **E**, and **H** was estimated by titration microcalorimetry by injecting aliquots of ligands in the cell containing the protein, and the resulting thermograms are displayed in Fig. 3. In all cases, strong exothermic peaks are observed indicating a favorable enthalpy of binding as classically observed in protein-carbohydrate interactions (13). No precipitation was observed at the end of titration. Using a simple one site model, oligosaccharides **E** and **H** bind with similar affinities, with  $K_d$  of  $\sim 5 \mu\text{M}$ . Such affinity is unusually strong for legume lectins and has been obtained only for trimannoside in multivalent binding to ConA

or *Dioclea grandiflora* lectin (DGL) (14, 15). The shorter glycopeptide **C** (pentasaccharide-Asn) presents a lower affinity ( $K_d = 185 \mu\text{M}$ ) in agreement with glycan array data. In the case of the symmetrical glycan **H**, a better fit is obtained using a model with two sites indicating that each branch of the heptasaccharide can bind to the lectin but with different affinities. However, no quantitative data can be safely extracted when using a two sites model for fitting procedure. In contrast, the unsymmetrical glycan **E** fits very well with the one site model.

As displayed in Fig. 3, the heat released by oligosaccharides **H** and **E** when binding to PELa are very different, as well as the shape of the peaks. Longer time is required for oligosaccharide **H** to come to equilibrium. Although both compounds have approximately the same affinity for PELa, and therefore similar free energy of binding  $\Delta G$  of about  $-30 \text{ kJ mol}^{-1}$  (Table 3), the enthalpy and entropy contributions are very different. The symmetrical compound **H** has a medium size enthalpy of binding ( $\Delta H = -23 \text{ kJ mol}^{-1}$ ) and a favorable entropy that contributes to  $\sim 20\%$  of the binding energy ( $-T\Delta S = -7.2 \text{ kJ mol}^{-1}$ ). The asymmetrical compound **E** has a much higher enthalpy of binding ( $\Delta H = -54 \text{ kJ mol}^{-1}$ ) and a strong entropy barrier with an unfavorable contribution of  $-T\Delta S = 23 \text{ kJ mol}^{-1}$ . Compound **C** (pentasaccharide-Asn) binding is also enthalpy-driven (Table 3). In relation to the presence of an additional GlcNAc on the 6-arm, compound **H** therefore has a different behavior, with longer equilibrium time related to favorable entropy of binding.

### Crystal Structures of PELa in Complex with Oligosaccharides

**Overall Description**—Crystals for different complexes were obtained by the hanging drop vapor diffusion method. The protein at  $10 \text{ mg ml}^{-1}$  was incubated with  $1 \text{ mM}$  oligosaccharide prior to testing crystal growth. Co-crystallization with the branched trimannose  $\alpha\text{Man1-3}(\alpha\text{Man1-6})\text{Man}$  (TriMan) leads to crystals diffracting to  $2.1 \text{ \AA}$  in space group  $P2_12_12_1$

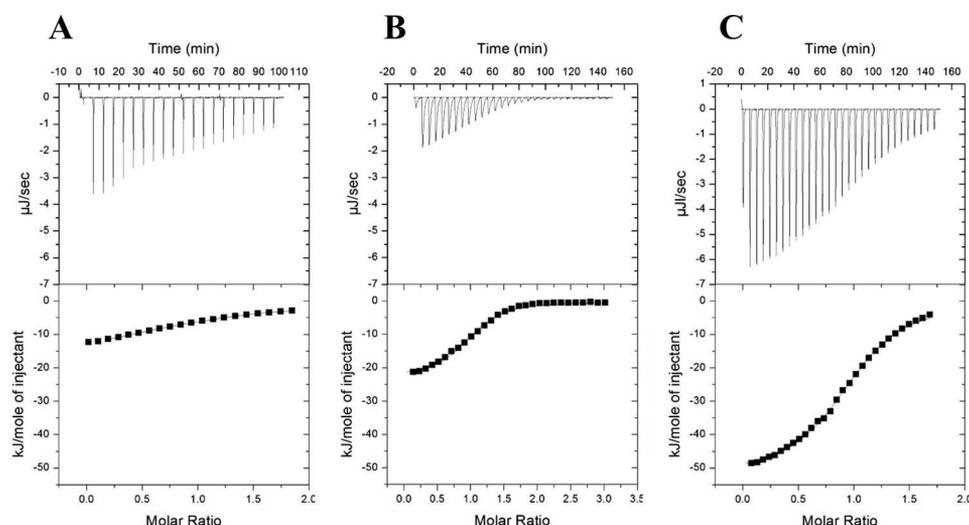


FIGURE 3. **Microcalorimetry data.** ITC plot obtained from the titration of PELA (60  $\mu\text{M}$ ) by Asn-linked glycans **C** (A), **H** (B), and **E** (C) measured on ITC200 (A) and VP-ITC (B and C) from Microcal. The plots in the *lower panel* show the total heat released as a function of total ligand concentration for the titration shown in the *upper panels*. The *solid line* represents the best least squares fit to the experimental data using a one site model.

**TABLE 3**

**Microcalorimetry titration data for the binding of the oligosaccharides E and H to PELA**

All measured values are averaged over two experiments except for compound C. Standard deviations are indicated in parenthesis.

Ligand	$K_d$	$n$	$-\Delta G$	$-\Delta H$	$-T\Delta S$
	$\mu\text{M}$		$\text{kJ/mol}$	$\text{kJ/mol}$	$\text{kJ/mol}$
Compound H	4.6 (0.3)	1.06 (0.06)	30.4	23.2 (0.3)	-7.2
Compound E	4.5 (0.2)	0.98 (0.06)	30.5	53.9 (1.8)	23.4
Compound C <sup>a</sup>	133	1.03	22.1	18.7	-3.4

<sup>a</sup> Performed on ITC200.

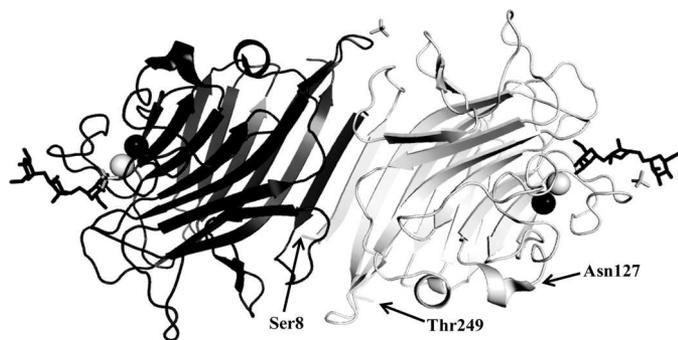


FIGURE 4. **Overall structure of a dimer of PELA complexed with trimannoside.** Peptide chains are represented as ribbon colored in *light gray* (chain A) and *dark gray* (chain B). Oligosaccharides are represented by *sticks* and calcium and manganese ions by *white* and *black spheres*, respectively.

(Table 1). The structure was solved by molecular replacement using the coordinates of PAL (Protein Data Bank code 1N3O). A dimer of lectin and one molecule of ligand are observed in the asymmetric unit. Crystals were also obtained with the symmetrical glycan **H**, diffracting to 1.6 Å in the same space group but with a different unit cell. The structure was solved by molecular replacement using the coordinates of the PELA-TriMan complex. A lectin dimer and two oligosaccharides are observed in the asymmetric unit.

In both structures, each monomer of PELA adopts the classical  $\beta$ -sandwich fold observed in all legume lectins, with the presence of one calcium and one manganese ion (Fig. 4) (9). From the 261 residues of PELA, all could be identified in the

electron density map, except for the first 7 or 8 at the N-terminal extremity, and the last 11 to 12 at the C-terminal extremity, depending on the chain. The two monomers in the asymmetric unit associate by parallel arrangement of their back  $\beta$ -sheet, resulting in the “canonical legume lectin dimer” that is classically observed in both dimeric and tetrameric legume lectins (9). The two potential glycosylation sites Asn-7 and Asn-127 are located in the disordered N-terminal region and in the loop at the surface of the protein, respectively (Fig. 4). Both positions are therefore likely to be glycosylated in plants.

**Crystal Structure of PELA-Trimannoside Complex**—The complex of the lectin with TriMan contains very clear density for the trisaccharide bridging two symmetry related PELA dimers, resulting in formation of lectin chains in the crystal (Fig. 5, A and B). The two  $\alpha$ -linked mannose residues are bound in facing monomers, the 1–3 linked mannose in monomer A and the 1–6 linked one in monomer B. The binding modes of the two mannose residues are very similar, each establishing five direct hydrogen bonds with side chains of Asp-95, Asn-147, and main chain nitrogen of Gly-115 and Glu-230 (Fig. 5C). Hydrophobic contacts involve Ala-94 and Phe-141. This interaction corresponds to the classical mannose binding mode observed in PAL (20) and ConA (32). A conserved water bridge was also observed at both sites between O2 and the carbonyl oxygen of Ala-113.

The central core mannose residue establishes bonds with both monomers, involving Asn-145, Ser-146, and Gln-231 of chain A and Glu-230 of chain B. The bridging TriMan results in rather close contact of the two facing protein monomers. Protein-protein contacts are created with the participation of Arg-123B side chain hydrogen bonded to Thr-148-A and Asp-150-A main chain. Also, the Asn-145 residue of each monomer interferes with the binding site of the other monomer and participates in hydrogen bonding to mannose O2 for Asn-145B or in hydrophobic contact to C2 for Asn-145A (Fig. 5, B and C). Because of the different environments, Asn-145 side chains of the two monomers have different orientations, and the back-

## N-Glycan Binding by *Platypodium elegans* Lectin

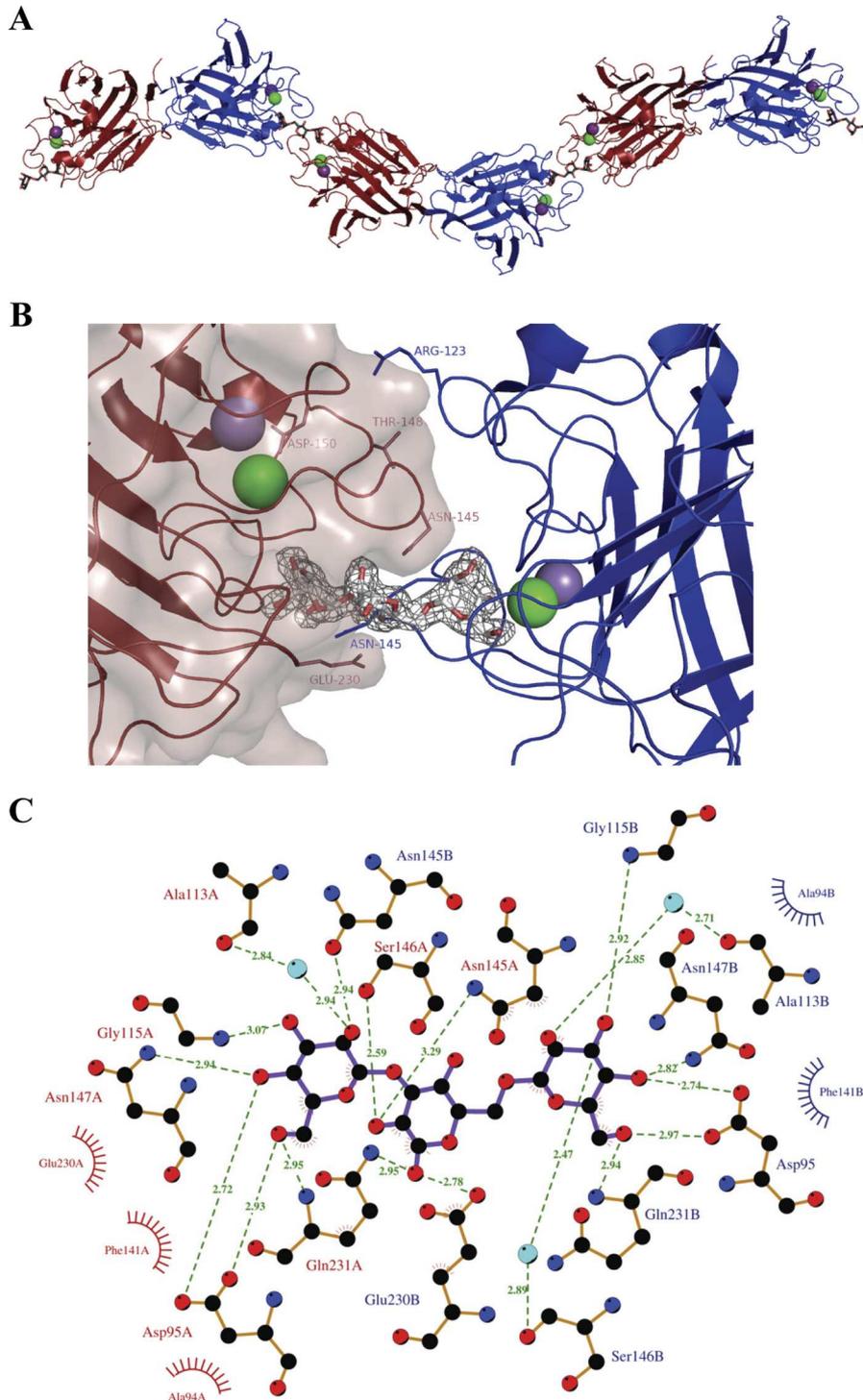
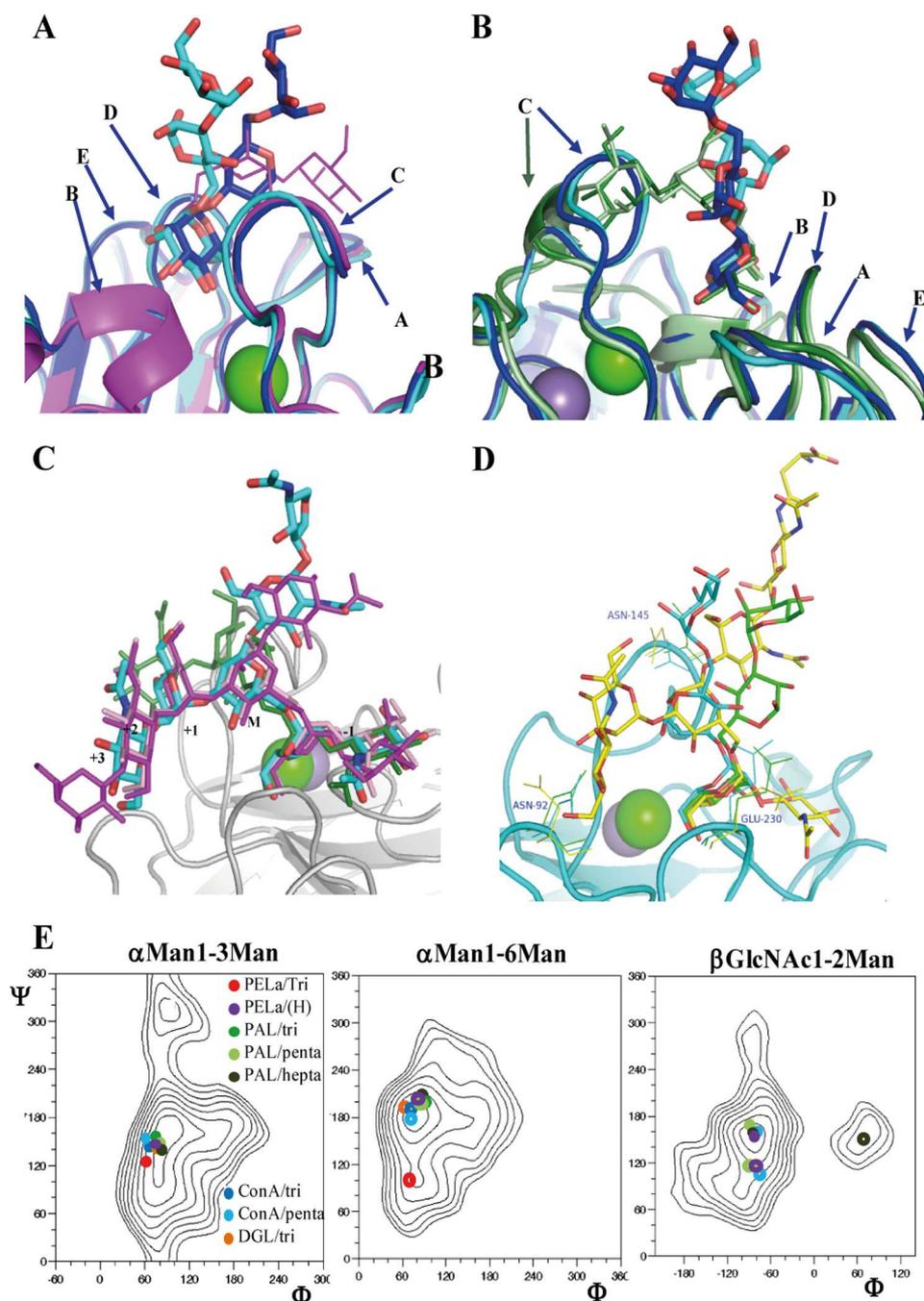


FIGURE 5. **Crystal structure of PELA-trimannoside complex.** *A*, chains of PELA lectins bridged by trimannosides represented as ribbons colored in *red* (chain A) and *blue* (chain B). *B*, close view of the interface with the trimannoside represented together with its  $2mF_o - DF_c$  electron density maps contoured at  $1\sigma$  (0.32 electrons per  $\text{\AA}^3$ ). Chain A is represented by its accessible surface (*red*) and chain B as ribbon (*blue*). *C*, LIGPLOT representation of the hydrogen bonds and van der Waals contact around the trimannoside.

bone of loop C also displays slightly different conformations (Fig. 6A).

As a result of its interaction with facing monomers, the Tri-Man shape is very extended and differs from what has been previously observed in other complexes with legume lectins such as PAL (20), ConA (33), and DGL (34). Comparison between PELA and PAL confirms the strong similarity between

the binding sites, with same conformation of the five loops, labeled A–E (Fig. 6A). PAL binds the trimannoside with the  $\alpha$ Man residue of the 6-arm in the primary binding site, as in monomer B of PELA. Nevertheless, the observed oligosaccharide conformation is very different because TriMan folds around the phenyl residue in the binding site of PAL, and the 3-linked mannose is hydrogen bonded to an asparagine residue



**FIGURE 6. Comparison of PELA-oligosaccharide complexes with related ones from other legume lectins together with labeling of loops from the binding site and analysis of the conformation of oligosaccharides bound to PELA.** *A*, superimposition of chain A (blue) and chain B (cyan) from the PELA-trimannoside complex together with PAL-trimannoside structure (magenta, Protein Data Bank code 1Q8V) (20). *B*, orthogonal view for chains A and B of PELA together with trimannoside complexed with ConA (dark green, 1ONA) (33) and with *D. grandiflora* lectin (light green, 1DGL) (34). *C*, superimposition of chain B (cyan) from PELA-Asn-linked heptasaccharide complex with complex oligosaccharides in the crystal of PAL (magenta, 2ARX; salmon, 2A6X) (17) and ConA (green, 1TEI) (36). *D*, comparison of the three binding modes observed in crystal structures as follows: TriMan bound through the  $\alpha$ Man1–3 (cyan) and by the  $\alpha$ Man1–6 (green) residues and heptasaccharide-Asn (yellow). *E*, energy maps are from 3D-cernav. Conformation at each linkage of all complexes has been superimposed on the corresponding energy map. Open and filled circles represent conformations of the 6-arm and 3-arm, respectively.

in loop A. The difference in conformation of TriMan in PELA and PAL is due to a  $90^\circ$  shift in the  $\Psi$  dihedral angle of the  $\alpha$ Man1–6Man linkage, the unusual value of  $105^\circ$  observed in PELA complex (Fig. 6E and Table 4) resulting in the extended shape of TriMan. Comparison with lectins from the *Dioclea* tribe is displayed in Fig. 6B. Binding site loops are well conserved between *Dioclea* (ConA and DGL) and *Dalbergiæae* (PAL and PELA), with the exception of loop C that is shorter in

ConA/DGL than in PAL-PELA (14 amino acids instead of 16) resulting in a more open binding site, as described previously (20). As a result, the TriMan conformation in ConA and DGL is different; it folds back on the protein surface and the 1–3-linked  $\alpha$ -Man makes additional contact with amino acids of loop C. The conformation observed for TriMan in ConA could not be possible in PELA due to the different conformation of loop C that would result in a steric conflict.

## N-Glycan Binding by *Platypodium elegans* Lectin

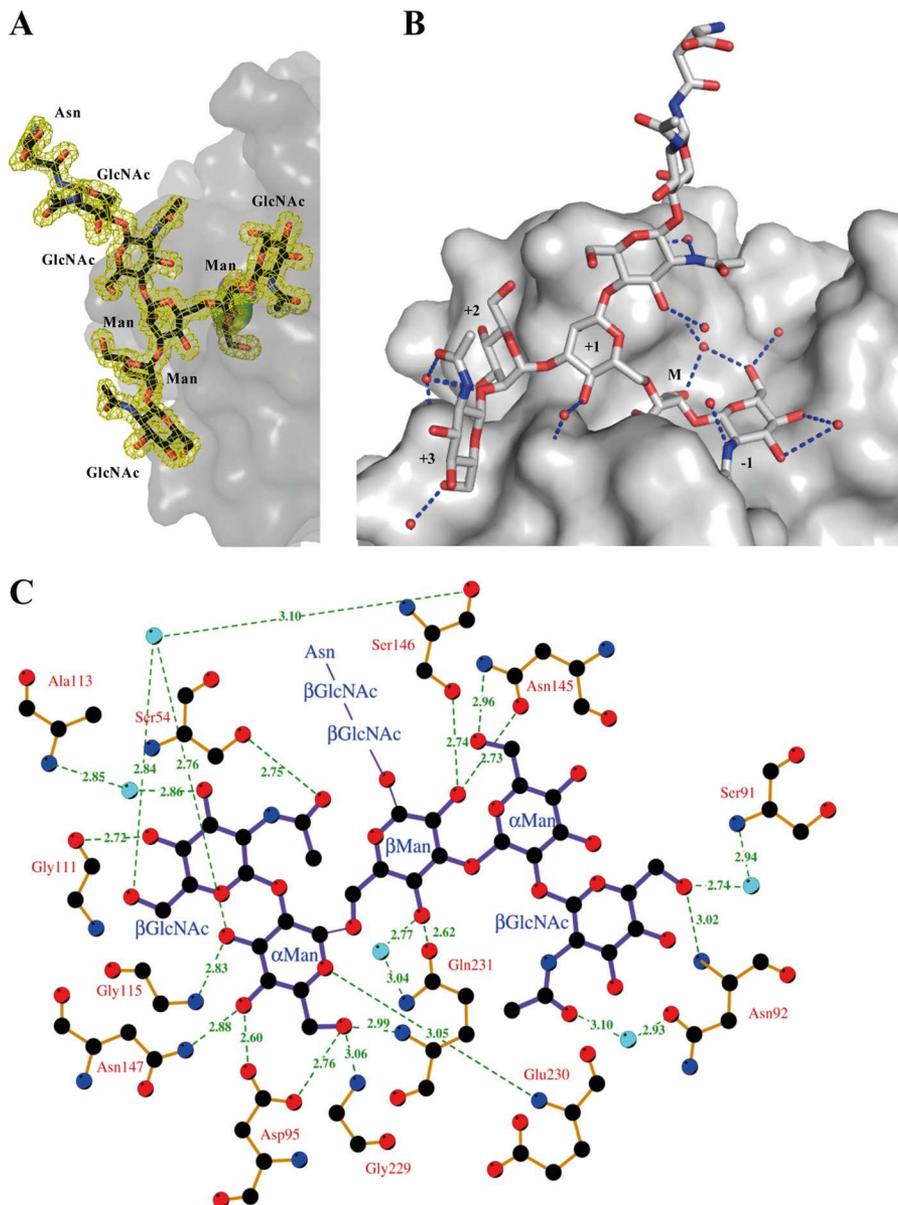
**TABLE 4**

Torsion angles ( $\Phi, \Psi$ ) or ( $\Phi, \Psi, \omega$ ) at selected glycosidic linkages in the crystal structures of PELA-oligosaccharide complexes and other lectins displayed in Fig. 6

All values are given in degrees.  $\Phi = \theta(O5-C1-O1-Cx')$ ,  $\Psi = \theta(C1-O1-Cx'-Cx_{+1}')$ ,  $\omega = \theta(O1-C6-C5'-O5')$ .

	1-6 arm		1-3 arm	
	$\alpha$ Man1-6Man	$\beta$ GlcNAc1-2Man	$\alpha$ Man1-3Man	$\beta$ GlcNAc1-2Man
PELa-trimannose <sup>a</sup>	70, 105, -53		65, 123	
PAL-trimannose	92, -149, -56		70, 129	
ConA-trimannose	75, -171, -63		72, 119	
DGL-trimannose	64, -165, -46		63, 129	
PELa-heptaose-Asn (H)	90, -153, -57	-79, 115	69, 128	-85, 152
PAL-pentaose	89, -154, -55	-81, 115	70, 129	-90, 170
PAL-heptaose	84, -147, -56	76, 150	72, 130	-84, 167
ConA-pentaose	68, 175, -43	-71, 105	62, 137	-80, 157

<sup>a</sup> See legend of Fig. 6 for Protein Data Bank code. A representative chain has been selected in each structure.



**FIGURE 7. Crystal structure of PELA-compound H complex.** *A*,  $2mF_o - DF_c$  electron density maps contoured at  $1\sigma$  (0.47 electrons per  $\text{\AA}^3$ ) around the heptasaccharide-Asn in monomer *A*. *B*, representation of the oligosaccharide on the protein surface. *C*, LIGPLOT representation of the hydrogen bonds for the ligand in monomer *A*. The chitobiose and asparagine moiety do not establish hydrogen bond with the protein, and for the sake of clarity, they are not displayed in the figure.

*PELa*-Asn-linked Heptasaccharide Complex—Clear density for the whole biantennary heptasaccharide was observed in both monomers of PELA, and the linked asparagine residue

could be located in monomer *A* (Fig. 7*A*) because of its stabilization by a salt bridge with Arg-221 of a symmetry-related protein. Apart from this slight difference, the oligosaccharides in

chains A and B are very similar in terms of conformation and hydrogen bonds (details available in supplemental Tables 2S and 3S), and all descriptions below will refer to monomer A, unless otherwise indicated. The mannose of the 6-arm is located in the primary binding site (subsite M) with the same orientation and the same hydrogen bond network as described above for the TriMan ligand. With the exception of the asparagine and the first GlcNAc attached to it, all residues are interacting with the protein surface, resulting in an extended surface of contact (Fig. 7B). The second GlcNAc of the chitobiose core sits against loop C and establishes a water-bridged hydrogen bond with Asn-145 main chain oxygen. The numerous other contacts involve both arms of the biantennary glycans. The GlcNAc on the 6-arm establishes three hydrogen bonds with amino acids of site -1 (Fig. 7C and supplemental Table 3S). The core mannose in site +1 and the 3-arm in sites +2 and +3 also displays a large number of contacts. Altogether, more than 15 direct hydrogen bonds and 6 water-bridged ones are observed, rationalizing the high affinity measured by titration microcalorimetry.

The conformations of the oligosaccharides have been analyzed and compared with previously determined energy maps (35). The conformation of the  $\beta$ GlcNAc1-2Man and  $\alpha$ Man1-3Man linkages in the 3-arm are close to the global minima (Table 4 and Fig. 6E). For the other antenna, the  $\alpha$ Man1-6Man linkage is in a low energy conformation, with a GG conformation of the  $\omega$  angle, but the  $\beta$ GlcNAc1-2Man linkage is clearly distorted with a  $\Psi$  value of about  $115^\circ$ , in a region of the energy map that is  $4 \text{ kcal mol}^{-1}$  above the global minimum at  $\Psi = 150^\circ$ . It could be noted that the O3 oxygen atom of this GlcNAc is solvent-exposed and available for elongation by a Gal residue, although the O4 points toward the Ser-110-Gly-111-Ser-112 peptides close to loop B and cannot be substituted. This is in agreement with glycan array data showing no binding to symmetrical *N*-glycans with LacNAc on both arms.

The conformation of the biantennary glycan and the network of contacts are very similar to what has been described previously for the structure of PAL interacting with complex-type pentasaccharide (17), albeit with small differences at the extremities of branches (Fig. 6C). The same pentasaccharide complexed with ConA exhibits the same position of 6-arm in subsite M and -1, but a different orientation of the 3-arm, again because of the important difference in loop C conformation (36). The same distorted orientation of the  $\beta$ GlcNAc1-2Man linkage in subsite -1 was described for the ConA complex. It should be noted that in other complexes, such as PAL-heptasaccharide (17), this linkage adopts a completely different orientation with a large change in the  $\Phi$  angle (Fig. 6E).

## CONCLUSION

This study describes a very complete sequence/structure/specificity characterization of a legume lectin from the *Dalbergieae* tribe with unusual affinity for asymmetric complex *N*-glycans. The sequence and structure present similarity with the lectin from *P. angolensis*, PAL, the only *Dalbergieae* lectin that was structurally characterized previously (17–20). Some of the structural features of binding of PELA to complex *N*-glycans are

also common to both lectins, such as the occurrence of an extended binding site that results in a very large surface of contact between the oligosaccharide and the protein. Nevertheless, very important new data are revealed by this work. This is the first characterization of a lectin from this subfamily by glycan array. Interestingly, PELA binds to a subset of complex *N*-glycans, and this binding profile is very different from those already determined for other lectins with Glc/Man specificity such as ConA, DGL, and lectins from *P. sativum* and from *Lens culinaris* that all bind preferably to oligomannose type *N*-glycans. We could also determine the very high affinity of PELA for Asn-linked complex-type oligosaccharides.

The crystallization of PELA in complex with trimannose results in infinite lectin chains in the crystal with one carbohydrate ligand bridging two lectin dimers through its terminal nonreducing mannose residues (Fig. 5A). Such trimannose bridging has already been observed in the crystal of snowdrop agglutinin (37) and bacterial lectin (38) but not in legume lectins. The only other structures of cross-linked legume lectin have been obtained with galactose-specific soybean agglutinin complexed with biantennary blood group I antigen and analogs (39). In PELA, the conformation of the bridging trimannoside is very extended and drastically different from the folded shape previously observed in other legume lectin complexes (20, 33, 34) as displayed in Fig. 6E and Table 4. Interestingly, the conformational change is not due to a rotation around the  $\omega$  angle, which always adopts a *gauche-gauche* conformation in the different legume lectins, but around the  $\Psi$  angle that is rotated by  $90^\circ$  (Fig. 6, D and E).

The preference for asymmetrical *N*-glycans with a short 6-arm has been confirmed by a semi-quantitative analysis of the glycan array data performed at several lectin concentrations. No crystals could be obtained to date for the complex between PELA and the asymmetric compound E, but the complex with the symmetric compound H (heptasaccharide-Asn) gives some clues about this preference. The GlcNAc on the 3-arm establishes contacts in the extended sites +2 and +3 while maintaining a low energy conformation of the linkages involved. In contrast, the binding of the GlcNAc of the 6-arm in site -1 induces a strong conformational distortion with energy cost. In addition, oxygen O4 is not accessible for elongation by a galactose. Oligosaccharides with long 3-arm and short 6-arm would therefore be energetically favored for binding. There is therefore an excellent agreement between glycan array and crystallographic data. However, solution data from ITC are not concordant because both the asymmetric and symmetric glycan bind with the same strong affinity to PELA. Nevertheless, the enthalpy profiles are different, and the asymmetric glycan has a much stronger enthalpy of binding, which confirms that the presence of an extra GlcNAc on the 6-arm is not favorable. The difference in entropy contribution between the two oligosaccharides is more difficult to explain. Possible reasons for such differences in entropy could be the trapping/release of water molecules on the protein surface. This work therefore exemplifies one of the cases where the physicochemical state of the ligand, *i.e.* solution state or attached to a surface, can influence its affinity for a receptor.

**Acknowledgments**—The glycan microarray analysis was provided by the Consortium for Functional Glycomics funded by National Institutes of Health Grant GM62116 from NIGMS. We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities, and we are grateful for the assistance in using beamline BM30A.

### REFERENCES

- Sharon, N., and Lis, H. (1990) Legume lectins. a large family of homologous proteins. *FASEB J.* **4**, 3198–3208
- Sharon, N., and Lis, H. (1995) Lectins. Proteins with a sweet tooth. Functions in cell recognition. *Essays Biochem.* **30**, 59–75
- Rüdiger, H., Siebert, H. C., Solís, D., Jiménez-Barbero, J., Romero, A., von der Lieth, C. W., Diaz-Mariño, T., and Gabius, H. J. (2000) Medicinal chemistry based on the sugar code. Fundamentals of lectinology and experimental strategies with lectins as targets. *Curr. Med. Chem.* **7**, 389–416
- Rüdiger, H. (1998) Plant lectins. More than just tools for glycoscientists. Occurrence, structure, and possible functions of plant lectins. *Acta Anat.* **161**, 130–152
- Peumans, W. J., and Van Damme, E. J. (1995) Lectins as plant defense proteins. *Plant Physiol.* **109**, 347–352
- Hirsch, A. M. (1999) Role of lectins (and rhizobial exopolysaccharides) in legume nodulation. *Curr. Opin. Plant Biol.* **2**, 320–326
- Delatorre, P., Rocha, B. A., Souza, E. P., Oliveira, T. M., Bezerra, G. A., Moreno, F. B., Freitas, B. T., Santi-Gadella, T., Sampaio, A. H., Azevedo, W. F., Jr., and Cavada, B. S. (2007) Structure of a lectin from *Canavalia gladiata* seeds. New structural insights for old molecules. *BMC Struct. Biol.* **7**, 52
- Loris, R., Hamelryck, T., Bouckaert, J., and Wyns, L. (1998) Legume lectin structure. *Biochim. Biophys. Acta* **1383**, 9–36
- Srinivas, V. R., Reddy, G. B., Ahmad, N., Swaminathan, C. P., Mitra, N., and Suroliya, A. (2001) Legume lectin family, the “natural mutants of the quaternary state,” provide insights into the relationship between protein stability and oligomerization. *Biochim. Biophys. Acta* **1527**, 102–111
- Sharma, V., and Suroliya, A. (1997) Analyses of carbohydrate recognition by legume lectins. Size of the combining site loops and their primary specificity. *J. Mol. Biol.* **267**, 433–445
- Young, N. M., and Oomen, R. P. (1992) Analysis of sequence variation among legume lectins. A ring of hypervariable residues forms the perimeter of the carbohydrate-binding site. *J. Mol. Biol.* **228**, 924–934
- Ambrosi, M., Cameron, N. R., and Davis, B. G. (2005) Lectins. Tools for the molecular understanding of the glycode. *Org. Biomol. Chem.* **3**, 1593–1608
- Dam, T. K., and Brewer, C. F. (2002) Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.* **102**, 387–429
- Dam, T. K., Oscarson, S., and Brewer, C. F. (1998) Thermodynamics of binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to *Dioclea grandiflora* lectin. *J. Biol. Chem.* **273**, 32812–32817
- Williams, B. A., Chervenak, M. C., and Toone, E. J. (1992) Energetics of lectin-carbohydrate binding. A microcalorimetric investigation of concanavalin A-oligomannoside complexation. *J. Biol. Chem.* **267**, 22907–22911
- Cavada, B. S., Moreira, R. A., Oliveira, J. T., and Grangeiro, T. B. (1993) Primary structures and functions of plant lectins. *R. Bras. Fisiol. Veg.* **5**, 193–201
- Buts, L., Garcia-Pino, A., Imberty, A., Amiot, N., Boons, G. J., Beeckmans, S., Versées, W., Wyns, L., and Loris, R. (2006) Structural basis for the recognition of complex-type biantennary oligosaccharides by *Pterocarpus angolensis* lectin. *FEBS J.* **273**, 2407–2420
- Garcia-Pino, A., Buts, L., Wyns, L., Imberty, A., and Loris, R. (2007) How a plant lectin recognizes high mannose oligosaccharides. *Plant Physiol.* **144**, 1733–1741
- Loris, R., Imberty, A., Beeckmans, S., Van Driessche, E., Read, J. S., Bouckaert, J., De Greve, H., Buts, L., and Wyns, L. (2003) Crystal structure of *Pterocarpus angolensis* lectin in complex with glucose, sucrose, and turanose. *J. Biol. Chem.* **278**, 16297–16303
- Loris, R., Van Walle, I., De Greve, H., Beeckmans, S., Deboeck, F., Wyns, L., and Bouckaert, J. (2004) Structural basis of oligomannose recognition by the *Pterocarpus angolensis* seed lectin. *J. Mol. Biol.* **335**, 1227–1240
- Lorenzi, H. (1998) *Árvores Brasileiras: Manual de Identificação e Cultivo de Plantas Arbóreas Nativas do Brasil*, Plantarum, Nova Odessa, Brazil
- Kajihara, Y., Suzuki, Y., Yamamoto, N., Sasaki, K., Sakakibara, T., and Juneja, L. R. (2004) Prompt chemoenzymatic synthesis of diverse complex-type oligosaccharides and its application to the solid-phase synthesis of a glycopeptide with Asn-linked sialyl-undeca- and asialo-nonasaccharides. *Chemistry* **10**, 971–985
- Moreira, A., and Perrone, J. C. (1977) Purification and partial characterization of a lectin from *Phaseolus vulgaris*. *Plant Physiol.* **59**, 783–787
- Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–137
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.* **53**, 240–255
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 486–501
- Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT. A program to generate schematic diagrams of protein-ligand interactions. *Protein Eng.* **8**, 127–134
- Seko, A., Koketsu, M., Nishizono, M., Enoki, Y., Ibrahim, H. R., Juneja, L. R., Kim, M., and Yamamoto, T. (1997) Occurrence of a sialylglycopeptide and free sialylglycans in hen's egg yolk. *Biochim. Biophys. Acta* **1335**, 23–32
- Gourdine, J. P., Cioci, G., Miguet, L., Unverzagt, C., Silva, D. V., Varrot, A., Gautier, C., Smith-Ravin, E. J., and Imberty, A. (2008) High affinity interaction between a bivalve C-type lectin and a biantennary complex-type N-glycan revealed by crystallography and microcalorimetry. *J. Biol. Chem.* **283**, 30112–30120
- Mezzato, S., and Unverzagt, C. (2010) Synthesis of an Fmoc-Asn-heptasaccharide building block and its application to chemoenzymatic glycopeptide synthesis. *Carbohydr. Res.* **345**, 1306–1315
- Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb, A. J., Dodson, E. J., Papiz, M. Z., Wan, T., and Campbell, J. (1989) The structure of the saccharide-binding site of concanavalin A. *EMBO J.* **8**, 2189–2193
- Loris, R., Maes, D., Poortmans, F., Wyns, L., and Bouckaert, J. (1996) A structure of the complex between concanavalin A and methyl-3,6-di-O-( $\alpha$ -D-mannopyranosyl)- $\alpha$ -D-mannopyranoside reveals two binding modes. *J. Biol. Chem.* **271**, 30614–30618
- Rozwarski, D. A., Swami, B. M., Brewer, C. F., and Sacchettini, J. C. (1998) Crystal structure of the lectin from *Dioclea grandiflora* complexed with core trimannoside of asparagine-linked carbohydrates. *J. Biol. Chem.* **273**, 32818–32825
- Imberty, A., Delage, M. M., Bourne, Y., Cambillau, C., and Pérez, S. (1991) Data bank of three-dimensional structures of disaccharides: Part II, N-acetylglucosaminic type N-glycans. Comparison with the crystal structure of a biantennary octasaccharide. *Glycoconj. J.* **8**, 456–483
- Moothoo, D. N., and Naismith, J. H. (1998) Concanavalin A distorts the  $\beta$ -GlcNAc-(1 $\rightarrow$ 2)-Man linkage of  $\beta$ -GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -Man-(1 $\rightarrow$ 3)-[ $\beta$ -GlcNAc-(1 $\rightarrow$ 2)- $\alpha$ -Man-(1 $\rightarrow$ 6)]-Man upon binding. *Glycobiology* **8**, 173–181
- Wright, C. S., and Hester, G. (1996) The 2.0-Å structure of a cross-linked complex between snowdrop lectin and a branched mannopentaose. Evidence for two unique binding modes. *Structure* **4**, 1339–1352
- Lameignere, E., Shiao, T. C., Roy, R., Wimmerova, M., Dubreuil, F., Varrot, A., and Imberty, A. (2010) Structural basis of the affinity for oligomannosides and analogs displayed by BC2L-A, a *Burkholderia cenocepacia* soluble lectin. *Glycobiology* **20**, 87–98
- Olsen, L. R., Dessen, A., Gupta, D., Sabesan, S., Sacchettini, J. C., and Brewer, C. F. (1997) X-ray crystallographic studies of unique cross-linked lattices between four isomeric biantennary oligosaccharides and soybean agglutinin. *Biochemistry* **36**, 15073–15080
- Gouet, P., Courcelle, E., Stuart, D. I., and Métoz, F. (1999) ESPript. Analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308

*Pseudomonas aeruginosa* est un pathogène opportuniste responsable de nombreuses maladies nosocomiales chez les patients immunodéprimés ainsi que d'infections graves chez les patients atteints de la mucoviscidose (CF). La colonisation des voies respiratoires des patients CF est souvent mortelle car une fois installée, cette bactérie est difficile à éradiquer et provoque le déclin des fonctions respiratoires des patients. L'antibiothérapie devient inefficace face au développement de souches multi-résistantes et sa capacité à former un biofilm. L'étape cruciale initiant l'infection ou la formation du biofilm est l'adhésion durant laquelle des interactions spécifiques lectines/oligosaccharides permettent la fixation de la bactérie à la surface de la cellule hôte. Bloquer l'adhésion serait un moyen de lutter contre l'infection. Dans l'approche d'une thérapie anti-adhésive, plusieurs lectines impliquées dans l'adhésion et l'élaboration du biofilm sont prises comme cibles. Dans un premier temps, des lectines fimbriales putatives identifiées récemment, CupB6 et CupE6, ont fait l'objet d'une étude. Des essais d'expression, de purification et de cristallisation ont été réalisés dans l'objectif de résoudre leur structure cristallographique. Une étude visant à identifier le ligand naturel de CupB6 a également été entreprise. Puis une étude a été menée pour caractériser le potentiel d'inhibition de plusieurs molécules dérivées du galactose sur la lectine soluble, PA-IL. Certaines de ces molécules pourraient être utilisées comme glycomimétiques offrant une alternative aux antibiotiques. Une étude par microcalorimétrie et cristallographie aux rayons X a permis d'étudier la spécificité d'une lectine de légumineuse, PELa.

*P. aeruginosa* is an opportunistic pathogenic bacterium responsible for numerous nosocomial infections in immunosuppressed patients. It is the first mortal pathogen in cystic fibrosis (CF) patients. The invasion of the respiratory tract of CF patients by the bacterium is often lethal because it is hard to eradicate and it rapidly impairs the respiratory functions of the patients. None of the current antibiotherapy procedures are efficient against multiresistant, biofilm forming *P. aeruginosa*. The first step leading to infection or biofilm formation involves the initial adhesion of bacterial cells to the host pulmonary cells via specific lectin/oligosaccharid interactions. Blocking the adhesion would be a way to fight against the infection. The anti-adhesion therapy targets several bacterial lectins involved in adhesiveness and biofilm formation. In this work, the recently identified putative fimbrial lectins CupB6 and CupE6 have been studied. Expression and purification tests followed by crystallization trials have been performed. In parallel, attempts to identify the natural ligand of CupB6 were also carried out. This work also presents a systematic characterization of the inhibitory effects of various galactose-derived molecules on the PA-IL lectin. Some of these molecules could be used as glycomimetic drugs thus offering an interesting alternative to standard antibiotics. Finally, the combination of microcalorimetry together with X-ray crystallography enabled us to gain insights into the ligand specificity of PELa, a legume lectin.