HAL open science

Etudes structurales et fonctionnelles des lectines solubles de Burkholderia cenocepacia

Emilie Lameignere

To cite this version:

Emilie Lameignere. Etudes structurales et fonctionnelles des lectines solubles de Burkholderia cenocepacia. Cristallographie. UNIVERSITE JOSEPH FOURIER, 2009. Français. NNT : . tel-03678262

HAL Id: tel-03678262

https://hal.science/tel-03678262v1

Submitted on 25 May 2022

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.

UNIVERSITE DE GRENOBLE I – JOSEPH FOURIER
ECOLE DOCTORALE CHIMIE ET SCIENCES DU VIVANT

THESE

Pour l'obtention du diplôme de
DOCTEUR DE L UNIVERSITE JOSEPH FOURIER

Discipline : Biologie structurale et nano-biologie

Présentée publiquement le 2 février 2009 par

Emilie Lameignère

Etudes structurales et fonctionnelles des lectines solubles de *Burkholderia cenocepacia*

Jury

Markus AEBI	**Rapporteur**
Nushin AGHAJARI	**Rapporteur**
Ina ATREE	**Examinateur**
Anne IMBERTY	**Directeur de thèse**
Miguel VALVANO	**Examinateur**

Thèse préparée au CERMAV – CNRS – UPR5301

Sommaire

1. *Burkholderia cenocepacia*: une bactérie opportuniste qui aime les sucres 1

1.1. Le complexe Burkholderia cepacia…….. 1
1.1.1. Une taxonomie en constante évolution 1
1.1.1.1. Les bactéries du genre Burkholderia
1.1.1.2. Le complexe Burkholderia cepacia
1.1.2. Un génome inhabituel 4
1.1.3. *Burkholderia cepacia* : une bactérie utile 5
1.1.3.1. Les bactéries du complexe BCC dans l'environnement naturel
1.1.3.2. Aspects bénéfiques pour l'agriculture
1.1.3.3. Utilisation des B. cepacia comme agent de remédiation
1.1.3.4. Limitations actuelles de l'utilisation des *B. cepacia*
1.1.3.5. *B. cepacia* : un pathogène opportuniste

1.2. La mucoviscidose……*9*
1.2.1. Historique 9
1.2.2. Une maladie génétique 10
1.2.3. Défaillance de la clairance mucociliaire 11
1.2.4. Importance des sucres dans la mucoviscidose 12
1.2.5. Infections bactériennes 17
1.2.5.1. Histoire des infections à B. cepacia
1.2.5.2. Epidémiologie des infections à B. cepacia

1.3. Virulence de B. cepacia……*20*
1.3.1. Comportement social de *Burkholderia*. 20
1.3.1.1. Le Quorum sensing
1.3.1.2 Formation de biofilms
1.3.1.3. Interaction avec *P. aeruginosa*
1.3.2. Facteurs de virulence 24
1.3.2.1. Les lipopolysaccharides
1.3.2.2. Les exopolysaccharides
1.3.2.3. Les sidérophores
1.3.2.4. Les systèmes de sécrétion
1.3.2.5. Les Flagelle et pili
1.3.2.6. Les lectines solubles

2. Les lectines de bactéries pathogènes……... 31

2.1. Les lectines associées à des toxines……*32*

2.2. Les lectines de surface……*34*

2.3. Les lectines solubles……*36*
2.3.1. Lectine en β-propeller de *Ralstonia solanacearum* 36
2.3.2. β-sandwich à 1 calcium 37

2.3.3. β -sandwich à 2 calciums……38

3. Identification des lectines solubles de *B. cenocepacia*
et objectifs de ce travail de thèse…… *40*

3.1. Recherche et alignement des séquences homologues à PA-IIL……40
3.1.1. Recherche des séquences homologues à PA-IIL par BLAST 40
3.1.2. Alignement des séquences 41

3.2. Description des protéines "PA-IIL-like"…….. 41

3.3. Lectines PA-IIL-like de B. cenocepacia…… 44

3.4. Objectif de ce travail de thèse……46

4. Outils pour l'étude des interactions protéine – sucre……. 48

4.1. Les tests d'hémagglutination……..48

4.2. Les puces à sucres…… 49
4.2.1. Les glycan arrays 50
4.2.2. Les printed arrays 51

4.3. Le microcalorimétrie isotherme de titration……52
4.3.1. Notions de thermodynamique 53
4.3.2. Thermodynamique des interactions non covalentes 54
4.3.3. Principe de l'ITC 56

4.4. La résonnance plasmonique de surface……59

4.5. La cristallographie aux rayons X……..60
4.5.1. La cristallogenèse des protéines 61
4.5.2. Principe de la diffraction aux rayons X 65
4.5.3. Collecte des données des données expérimentales 69
4.5.4. Indexation, intégration et mise à l'échelle du jeu de données 69
4.5.5. Facteurs de structure et densité électronique 70
4.5.6. Obtention des phases par la méthode de remplacement moléculaire……71
4.5.7. Construction, affinement et validation du modèle 72

4.6. Le Biofilm Ring Test……..73

5. Etude de la lectine BclA…… 75
Article 1 76
Article2 89

6. Caractérisation de la lectine BclB……..107

6.1 Matériel et méthodes…… 107
6.1.1. Clonage des gènes d'intérêt 107
6.1.2. Expression des protéines recombinantes chez E. coli 110
6.1.3. Purification des protéines recombinantes 111
6.1.4. Diffusion dynamique de la lumière 115
6.1.5. Résonnance plasmonique de surface 115
6.1.6. Marquage des protéines pour les puces à sucre 117
6.1.7. Microcalorimétrie isotherme de titration 118
6.1.8. Tests d'Hémagglutination 118
6.1.9. Cristallogenèse *119*

6.2. Résultats et discussion…… 119
6.2.1. Production des protéines recombinantes 119
6.2.2. Etat d'oligomérisation de la lectine et de ses domaines en solution 124
6.2.3. Etude de la spécificité de BclB pour les oligosaccharides 128
6.2.4 Mesure de l'affinité de BclB pour l'αMeFuc 132
6.2.5 Etudes fonctionnelle du domaine *N*-terminal 134
6.2.6. Cristallogenèse 137

6.3. Conclusions…… 138

7. Localisation et fonction des lectines solubles de *B. cenocepacia*……. 139

7.1. Matériel et méthodes…….. 140
7.1.1. Clonages 140
7.1.2. Localisation cellulaire des lectines 145
7.1.3. Etude fonctionnelle des lectines 148

7.2. Résultats……150
7.2.1. Expression et immnunogénicité des lectines BclA, BclB et BclC 150
7.2.2. Localisation des lectines solubles de *B. cenocepacia* 152
7.2.3. Préparation des bactéries recombinantes 154
7.2.4. Etude de l'affinité de BclA et BclB pour des fragments lipopolysaccharidiques 156
7.2.5. Etude du rôle des lectines chez *B. cenocepacia* K56-2 158

7.3. Conclusions et perspectives……..161

8. Conclusion et perspectives……162

8.1 Etude de la lectine BclA…….. 162
8.2 Etude de la lectine BclB…….. 164
8.3. Rôle des lectines de surface chez B. cenocepacia……..165

Bibliographie……..166

Annexes……..182

Liste des principales abréviations

AHL:	N-acyl-homoserine-lactone
BCC:	*Burkholderia cepacia* complex
BclA:	*Burkholderia cepacia* lectin A
BclB:	*Burkholderia cepacia* lectin B
BclC:	*Burkholderia cepacia* lectin C
BclD:	*Burkholderia cepacia* lectin D
CF:	Cystic Fibrosis
CFTR:	Cystic Fibrosis Transmembrane conductance Regulator
CV-IIL:	*Chromobacterium violaceum* lectin II
DLS:	**D**ynamic Light Scattering
DTT:	Dithio-Threitol
E. coli:	*Escherichia coli*
EDTA:	Acide Ethylène Diamine TétraAcétique
Fuc:	L-fucose
Gal, GalNAc:	D-galactose, N-acethylgalactosamine
Glc, GLCNAC:	D-glucose, N-acétylglucosamine
HPLC:	High Pressure Liquid Chromatography
IPTG:	Iso-Propyl-β-d-Thiogalactopyranoside
ITC:	Isothermal Titration Calorimeter
LPS:	Lipopolysaccharides
Man:	D-mannose
Me- :	methyl-
NeuAc:	acide sialique
PA-IL :	*Pseudomonas aeruginosa* lectine I
PA-IIL :	*Pseudomonas aeruginosa* lectine II
PEG:	poly(ethylène glycol)
QS:	Quorum Sensing
RS-L:	*Ralstonia solanacearum* lectine I
RS-IIL:	*Ralstonia solanacearum* lectine II
SDS-PAGE:	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SelMet:	Selenométhionine
SPR:	Surface Plasmon Resonance
TCEP:	tris-(2-carboxy-. ethyl) phosphine

Chapitre 1

Burkholderia cenocepacia : une bactérie opportuniste qui aime les sucres

1.1. Le complexe Burkholderia cepacia

B. cepacia est un ensemble de bactéries à gram négatif qui comprend plusieurs espèces très similaires sur le plan phénotypique. Extrêmement versatiles, on les trouve dans l'environnement naturel où elles peuvent se révéler être à la fois néfastes et bénéfiques pour les plantes et les plantations. Initialement décrites comme phytopathogènes, leurs capacités métaboliques présentent des intérêts en biotechnologie et en bioremédiation. Cependant, au cours des vingt dernières années, cet organisme a également été décrit comme un pathogène humain opportuniste responsable d'épidémies, en particulier chez les patients atteints de la mucoviscidose.

1.1.1. Une taxonomie en constante évolution

1.1.1.1. Les bactéries du genre Burkholderia

Le genre *Pseudomonas* est décrit pour la première fois il y a plus de 100 ans. La bactérie, à gram négatif, *Pseudomonas cepacia*, est identifiée en 1950 par le phytopathologiste W.H. Burkholder, Université de Cornell, comme étant responsable de la pourriture molle de l'oignon ; *cepa* : oignon (Bukholder 1950).

La position taxonomique de cette bactérie a ensuite subi une évolution constante. En 1973, l'hétérogénéité taxonomique du genre *Pseudomonas* est révélée par le travail de Palleroni et al. (Palleroni, *et al.* 1973).Au vu de résultats d'hybridation ARNr-ADN, le genre

Pseudomonas est divisé en 5 groupes. En 1992, Yabuuchi *et al*., se basent sur la séquence des ARNr 16S, sur les homologies ADN-ADN, sur la composition des lipides et sur des caractères phénotypiques pour transférer *P. cepacia* et quatre autres espèces, *P. solanacearum*, *P. pickettii*, *P. gladioli*, *P. mallei*, *P. pseudomallei et P. caryophylli* dans le nouveau genre, *Burkholderia* (Yabuuchi, *et al.* 1992). Celui-ci est placé dans la famille des Burkholderiaceae (ordre des Burkholderiales, classe des Betaproteobacteria, phylum des "Proteobacteria"). Il est donc phylogénétiquement éloigné du genre *Pseudomonas* situé dans la classe des Gammaproteobacteria. Quelques années plus tard, *P. solanacearum* et *P. pickettii* sont déplacées dans le nouveau genre, *Ralstonia* (Gillis, *et al.* 1995, Yabuuchi, *et al.* 1995).

Dans les années qui suivent, l'intérêt porté à *Burkholderia* est grandissant et de nombreuses nouvelles espèces, principalement d'origine environnementale, sont détectées et décrites. Depuis sa création il y a seulement quarante ans, le nombre de bactéries appartenant à ce genre est en constante augmentation et atteint aujourd'hui les 63 espèces (http://www.bacterio.cict.fr/b/burkholderia.html)

1.1.1.2. Le complexe Burkholderia cepacia

Au milieu des années 90, les études taxonomiques décrivent une certaine hétérogénéité à l'intérieur même de l'espèce *B. cepacia*. Les techniques utilisées auparavant pour étudier cette bactérie manquaient de sensibilité et de spécificité. Pour y remédier, Peter Vandamme et ses collaborateurs de l'Université de Ghent en Belgique, ont développé une approche de taxonomie polyphasique basée sur différentes méthodes génotypiques et phénotypiques (hybridation ADN-ADN, analyse des homologies de séquence de l'ARNr 16S et du gène recA, tests biochimiques et différentes méthodes d'empreinte de l'ADN comme l'AFLP: Polymorphisme de longueur de fragments amplifiés). Cette équipe pionnière en taxonomie, publie en 1997, les résultats d'une première étude faite sur plus de 80 isolats d'origine environnementale et clinique. Cette étude montre que *B. cepacia* est en réalité un complexe composé d'au moins cinq espèces génomiques ou genomovars. Dans les cinq années suivantes, quatre autres génomovars ont été identifiées (*Tableau 1.1*). Le terme genomovar désigne des espèces phénotypiquement similaires mais génotypiquement différentes. Toutes les espèces du complexe de *Burkholderia cepacia* ou BCC présentent une variabilité phénotypique considérable (Vandamme, *et al.* 1997). Par exemple, une simple modification de la source de carbone peut modifier la morphologie des colonies des espèces de BCC

cultivées sur agar. Une telle capacité à modifier leur métabolisme et leur profil biochimique rend l'identification de ces bactéries par les seules méthodes phénotypiques très difficile.

En 2002, Vermis et al. proposent d'ajouter une dixième espèce : *B. ubonensis* (Vermis, *et al.* 2002). La présence de celle-ci au sein du complexe a récemment été confirmée par la technique de Multilocus Sequence Type (MLST) (Vanlaere, *et al.* 2008). Cette méthode, est fondée sur l'étude de la séquence des fragments d'ADN de sept gènes ancillaires bien conservés (housekeeping genes). La capacité de cette technique à différencier les espèces existantes au sein du complexe est supérieure à l'analyse du gène recA seul et montre une excellente corrélation avec les multiples méthodes de taxonomie polyphasique utilisées pour caractériser pleinement ces bactéries. La MLST, est donc probablement la méthode du futur pour l'identification de ces bactéries. Enfin dans cette même étude, E. Vanlaere et al. ont démontré la présence de cinq nouvelles espèces, *B. latens* sp. nov., *B. diffusa* sp. nov., *B. arboris* sp. nov., *B. seminalis* sp. nov., et *B. metallica* sp. nov au sein du complexe (*Tableau 1.1*).

Bactéries appartenant au complexe BCC	
Espèce	**Génomovar**
B. cepacia (sensu stricto)	*B. cepacia* genomovarI
B. multivorans	*B. cepacia* genomovarII
B. cenocepacia	*B. cepacia* genomovarIII
B. stabilis	*B. cepacia* genomovarIV
B. vitenamiensis	*B. cepacia* genomovarV
B. dolosa	*B. cepacia* genomovarVI
B. ambifaria	*B. cepacia* genomovarVII
B. anthina	*B. cepacia* genomovarVIII
B. pyrrocinia	*B. cepacia* genomovarIX
B. ubonensis	*B. cepacia* genomovarX
Bactéries en attente de classification taxonomique	
Espèce	
B. latens sp. nov	
B. diffusa sp. nov.	
B. arboris sp. nov	
B. seminalis sp. nov	
B. metallica sp. nov	

Tableau 1.1. Vue d'ensemble des bactéries appartenant au complexe *Burkholderia cepacia*

1.1.2. Un génome inhabituel

Les avancées dans le séquençage du génome des différentes souches de *Burkholderia* ont contribué à une meilleure compréhension de la versatilité de ces organismes (http://www.ncbi.nlm.nih.gov/, http://genome.jgi-psf.org/, http://www.broad.mit.edu/seq/msc/ ou http://pathema.tigr.org/Burkholderia/beta/). La souche LB400 de *B. xenovorans*, une bactérie de l'environnement qui peut dégrader les composés polluants de polychlorobiphényles, a été le premier génome d'une bactérie du genre *Burkholderia* à être séquencé. Aujourd'hui, de nombreux génomes complets sont disponibles comme par exemple celui de *B. pseudomallei* et *B. mallei* (Holden, *et al.* 2004, Nierman, *et al.* 2004), celui de la souche G4 de *B. vietnamiensis*, un bon agent de bio-remédiation (Lessie, *et al.* 1996), ainsi que celui de souche 383 de *B. cepacia*,, une bactérie décrite dans le sol en 1966 (Stanier, *et al.* 1966). Une étude du métagénome de la mer des Sargasses a également permis d'identifier un génome quasiment complet de *Burkholderia*, désigné par *Burkholderia* SAR-1 (Venter, *et al.* 2004).

Le génome des espèces du genre *Burkholderia* est exceptionnel pour deux raisons. Tout d'abord, alors que la plupart des bactéries disposent d'un chromosome circulaire unique, *B. cepacia* possède un génome complexe constitué de plusieurs chromosomes circulaires éventuellement associés à un ou plusieurs plasmides permettant ainsi une grande souplesse dans l'acquisition ou la perte de l'expression d'un gène (Mahenthiralingam, *et al.* 2005). La taille du génome de ces bactéries est de 6 à 9MB, soit deux fois la taille de celui d'*Escherichia coli*, quatre fois celui d'*Haemophilus influenza* et à peine moins que celui du plus petit eucaryote, la levure. Ensuite, ce génome contient de nombreux fragments d'ADN mobiles transposables sous forme de séquences d'insertion (IS) et d'îlots génomiques (région du génome récemment acquise par la bactérie). Environ 10% des 7000 gènes codés par ce génome seraient acquis par transfert horizontal de gènes à partir d'ADN étranger : îlots génomique, phages intégrés au génome ou plasmides (Mahenthiralingam, *et al.* 2005). Ceci contribue à la plasticité génomique de cette bactérie par réarrangement du génome, activation de l'expression de certains gènes et recrutement de gènes étrangers et explique probablement la polyvalence nutritionnelle (utilisation de nombreux substrats comme source unique de carbone) et les capacités d'adaptation à des environnements très divers de la bactérie (*Figure 1.1*).

Le génome de la souche J2315 de *B. cenocepacia* auquel appartient le clone épidémique et virulent, ET12, a été séquencé et est en cours d'annotation (Govan, *et al.* 1993, Govan and Deretic 1996). La taille du génome de cette bactérie est de 8.056 Mb, avec un pourcentage moyen de bases GC de 66,9%, similaire à celui observé pour *Pseudomonas species* (Yabuuchi, *et al.* 1992). Trois grands chromosomes et un petit plasmide de 92kb ont été identifiés. La plupart des gènes essentiels au métabolisme sont codés par le chromosome 1. Environ 10% de l'ensemble du génome, soit 800 kb, est sous forme d'îlots génomiques.Un bactériophage de type Mu, connu sous le nom de BcepMu, a également été identifié dans le génome de J2315 (Summer, *et al.* 2004). Les séquences correspondant à ce bactériophage ont été détectées dans toutes les souches de la lignée ET12 à l'exception de K56-2, mais elles sont absentes dans les autres souches de *B. cenocepacia* ainsi que dans les autres espèces du complexe.

1.1.3. Burkholderia cepacia : une bactérie utile

Des capacités d'adaptation remarquables ainsi qu'une aptitude à métaboliser des sources de carbone très variées permettent aux bactéries du complexe *Burkholderia* de coloniser des niches écologiques diversifiées.

1.1.3.1. Les bactéries du complexe BCC dans l'environnement naturel

Depuis 1990, de nombreuses espèces de *Burkholderia* ont été découvertes dans la rhizosphère des cultures à travers le monde. Parmi les bactéries du complexe BCC, *B. vietnamiensis* est présente dans la rhizosphère du riz au Vietnam (Tran Van, *et al.* 1994), *B. ambifaria* dans celle du pois ou du blé au Royaume-Uni (Compant, *et al.* 2008). *B. multivorans*, *B. cenocepacia*, *B. stabilis*, *B. dolosa* et *B. pyrrocinia* ont également été mises en évidence dans la rhizosphère du maïs (Dalmastri, *et al.* 1999, Payne, *et al.* 2006).

En 2004, J. Craig Venter publie une étude novatrice explorant la diversité microbienne de la mer des Sargasses, par séquençage à haut-débit de dizaines de milliers de fragments d'ADN présents dans un mètre cube d'eau de mer (Venter, *et al.* 2004). Ce séquençage systématique de fragments d'ADN clonés a conduit à la création d'un métagénome de la mer des Sargasses, c'est-à-dire à l'ensemble des génomes des populations bactériennes de ce milieu. Dans celui-ci, l'identification d'une bactérie désignée *Burkholderia* SAR-1, est très surprenante car ces organismes ne sont pas connus comme étant capables de vivre en eau

salée. L'équipe de Mahenthiralingam a suggéré que cet isolat serait associé à une contamination par une *B. cepacia* (*sensu stricto*) du groupe K (Mahenthiralingam, *et al.* 2006).

1.1.3.2. Aspects bénéfiques pour l'agriculture

Certaines espèces du genre *Burkholderia* telles que *B. glumae* et *B. plantari* sont des phytopathogènes problématiques pour les cultures de riz. Comme initialement décrites par Walter Burkholder, les bactéries du complexe *B. cepacia* sont également des pathogène des plantes, mais elles peuvent aussi avoir une action favorable pour l'agriculture (*Figure 1.1*).

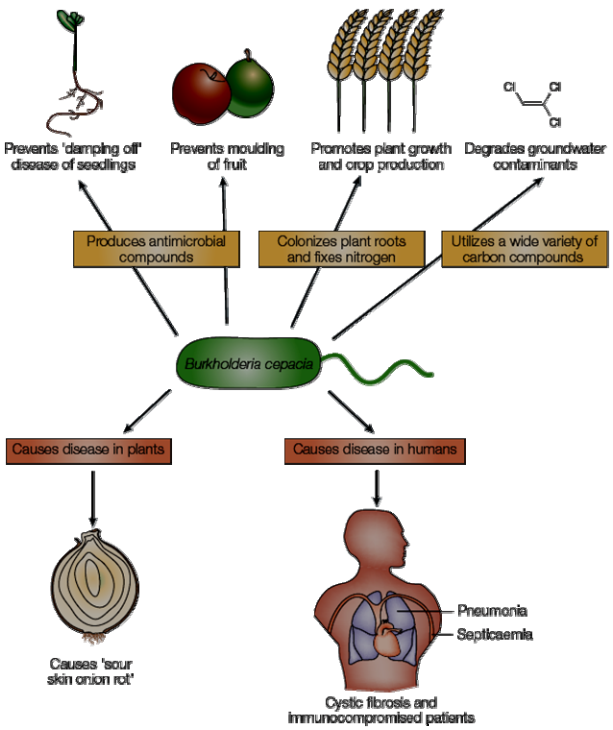


Figure 1.1. Effets néfastes et bénéfiques des bactéries du complexe BCC.

Les bactéries du complexe BCC, initialement décrites comme des phytopathogènes responsables de la pourriture de l'oignon, peuvent aussi avoir des effets bénéfiques sur l'environnement. Elles sont capables de produire des composés antimicrobiens qui protègent les plantes, ainsi que des facteurs favorisant leur croissance. Elles ont la capacité d'établir une symbiose fixatrice d'azote et sont susceptibles de dégrader une grande variété de polluants. Mais certaines bactéries de ce groupe sont responsables d'infections nosocomiales chez les patients immunodéprimés ainsi que chez ceux atteints de la mucoviscidose.

D'après (Mahenthiralingam, et al. 2005)

Tout d'abord, certaines espèces de *Burkholderia* : *B. cepacia, B. cenocepacia* ou *B. ambifaria*, produisent des antifongiques naturels. Elles peuvent ainsi protéger les cultures contre des maladies telles que la pourriture des racines, causée par *Aphanomyces euteiches*, et l'infections des semences par des fongiques nuisibles tels que *Pythium sp*, et *Rhizoctonia solani* (Parke and Gurian-Sherman 2001). La souche MC17 de *B. ambifaria* protège le maïs contre le phytopathogène *Fusarium verticillioides*. L'utilisation de ces bactéries comme biopesticides permettrait de limiter l'utilisation de pesticides chimiques toxiques et hautement persistants tels que le Captan (Parke and Gurian-Sherman 2001). L'aptitude de certaines

espèces de *Burkholderia* à augmenter le rendement des cultures agricoles et forestières par production d'antifungiques a amené l'agence de protection de l'environnement américaine (EPA) à enregistrer cinq souches de *Burkholderia* pour permettre leur utilisation en agriculture. Trois d'entres elles ont plus tard été répertoriées comme *B. ambifaria* et une comme *B. cenocepacia*

Certaines *Burkholderia* sont également capables d'établir une symbiose fixatrice d'azote de l'air qui permettrait de limiter l'utilisation d'engrais azotés et d'améliorer le rendement des cultures agricoles. Comme pour le genre *Rhizobium*, cette association s'effectue dans les nodules localisés au niveau des racines ou plus rarement sur les tiges de la plante. En 2005, l'induction de façon efficace de la formation de nodules sur une légumineuse tropicale sauvage de Madagascar, *Dalbergia spp* par une souche appartenant au complexe *B. cepacia* est signalée pour la première fois (Rasolomampianina*, et al.* 2005). En 2007, les capacités de formation de nodules par deux souches de *B. tuberum* (STM678 et DUS833), chez la légumineuse *Cyclopia sp*. en microscopie électronique et par fluorescence sont confirmées (Elliott*, et al.* 2007). Des bactéries du genre *Burkholderia* ont récemment été décrites comme des symbiotes efficaces associés aux racines de *Fabaceae* et aux feuilles de *Rubiaceae* (Compant*, et al.* 2008).

1.1.3.3. Utilisation des B. cepacia comme agent de remédiation

Les bactéries du complexe BCC ont la capacité de métaboliser une large gamme de composés organiques incluant des composés aromatiques, chlorés et autres polluants. Cette aptitude pourrait en faire un des biopesticides intéressants et de bons agents de remédiation des sols et nappes phréatiques pollués par l'utilisation d'herbicides et de pesticides (Lessie*, et al.* 1996, Parke and Gurian-Sherman 2001). Par exemple, la souche G4 de *B. vietnamiensis* peut dégrader les polluants aromatiques comme le toluène et les solvants chlorés tels que le trichloréthylène. La souche ENV435 de *B. vietnamiensis* (derivée de G4) a été testée sur le terrain et a permis de réduire la quantité de solvants chlorés d'un aquifère sableux de plus de 70% après 2 jours (Steffan*, et al.* 1999)

1.1.3.4. Limitations actuelles de l'utilisation des B. cepacia

Les perspectives d'utilisation de souches environnementales du genre *Burkholderia* en agriculture ou en bioremédiation ne sont pas sans inquiéter les professionnels de la santé, dans la mesure où les relations entre ces souches et les souches cliniques ne sont pas totalement élucidées. En 2002, Lipuma et al. a constaté que le clone du district de Philadelphie-Columbia (PHDC), responsable d'épidémies aux États-Unis et en Europe, est également présent sur les oignons du sol de l'État de New York (Lipuma 2002). De plus, *B. cepacia* est naturellement résistantes à de nombreux antibiotiques. Etant donné que le risque pour l'homme ne peut encore être évalué, un moratoire a donc été décidé par la communauté scientifique internationale quand à l'utilisation à grande échelle de *Burkholderia* en agriculture en attendant une meilleure compréhension de cette bactérie et du risque qu'implique son utilisation. En 2003, en réponse aux pétitions de la « Cystic Fibrosis Foundation » et sur la base des données scientifiques, l'EPA (U.S. Environmental Protection Agency) reconsidère l'utilisation de produits commerciaux contenant ces bactéries et annule l'enregistrement de tous les pesticides contenant *Burkholderia* (voir le registre fédéral de 2003, http://www.epa.gov/fedrgstr/EPA-TOX/2003/June/Day-13/t15010.htm).

1.1.3.5. B. cepacia : un pathogène opportuniste

Les bactéries du complexe *Burkholderia cepacia* ne font pas partie de la flore commensale de l'homme et sont responsables d'épidémies nosocomiales occasionnelles. Une étude récente rapporte des cas de septicémie à *B. stabilis* dans un hôpital en Turquie. Malgré un traitement antibiotique approprié, tous les patients sont décédés entre un et huit jours après l'identification de la bactérie dans leur sang (Otag, *et al.* 2005). Les infections graves à *B. cepacia* peuvent aussi être une complication de la granulomatose (Chronic Granulomatous Disease, CGD). Les patients atteints de cette maladie présentent des leucocytes polymorphonucléaires défectueux. L'analyse des données provenant d'un registre de 368 patients atteints de CGD a montré que chez ces individus, les infections à *B. cepacia* représentent la deuxième cause de décès (Johnston 2001). Mais les bactéries du complexe BCC sont aussi et surtout responsables d'infections pulmonaires chez les patients atteints de mucoviscidose. Cette bactérie n'est connue comme un agent responsable d'infection chez ces patients que depuis le début des années 80. Même si elle est moins connue que d'autres

organismes tels que *S. aureus*, *H. influenza* ou *P. aeruginosa*, elle reste très inquiétante en raison de sa virulence, de sa résistance aux antibiotiques et de sa forte transmissibilité de patient à patient.

1.2. La mucoviscidose

La mucoviscidose est définie comme une maladie pulmonaire obstructive chronique accompagnée d'une insuffisance pancréatique exocrine et d'une élévation de la concentration en sodium et en chlore dans la sueur. Un transport anormal du chlore causé par le dysfonctionnement du canal transmembranaire CFTR (Cystic Fibrosis Transmembrane conductance Regulator) est à l'origine de ces symptômes. Il s'agit de la maladie génétique létale la plus fréquemment rencontrée en Europe et en Amérique du nord alors qu'elle reste rare dans les populations africaine et asiatique Chaque année, en France, près de deux cent enfants naissent atteints de cette maladie. Ainsi, plus de deux millions de personnes (soit une sur trente) seraient porteuses du gène de la mucoviscidose sans le savoir. Pour les enfants qui naissent en 2008, l'espérance de vie est de quarante six ans, alors qu'elle n'était que de sept ans en 1965, mais l'âge moyen de décès de l'ensemble des patients n'est encore que de vingt quatre ans (http://www.vaincrelamuco.org). Il n'existe, à l'heure actuelle pas de traitement curatif de cette maladie et les traitements qui soulagent les symptômes sont lourds et contraignants.

1.2.1. Historique

Dans le folklore d'Europe du nord un vieil adage disait « Malheur à l'enfant chez qui un baiser sur le front a un goût salé. Il est ensorcelé et doit bientôt mourir. » (Busch 1990).Connue depuis le Moyen Âge, la maladie n'a pourtant été mise en évidence que récemment. C'est en 1936, dans une thèse présidée par le pédiatre suisse Guido Fanconi, que la maladie est décrite pour la première fois chez des enfants supposés atteints de maladie cœliaque, sous le nom de « fibrose kystique du pancréas et bronchectasies ». Les travaux d'Andersen mettent pour la première fois en évidence la nature génétique et autosomale de la maladie. Le terme de mucoviscidosis, crée à partir des termes « mucus » et « visqueux », est proposé pour la première fois en 1943 par le docteur Sydney Farber, médecin-chef au Children's Hospital de Boston.

En 1953, Di Sant'agnese décrit un excès de chlorure de sodium (NaCl) dans la sueur des enfants atteints de mucoviscidose (Di Sant'Agnese, *et al.* 1953). Peu après, cette découverte conduit, à la mise au point du "test de la sueur". Dans les années 80, des études menées sur des conduits sudoripares ont permis à Paul Quinton, lui-même atteint de la mucoviscidose, d'identifier un transporteur au chlore comme étant à la base de la pathologie.

En 1989, le gène impliqué dans la mucoviscidose est caractérisé selon une démarche de clonage positionnel par les équipes nord américaines de Lap-Chi Tsui, Collins et Riordan (Rommens, 1989 #12; Riordan, 1989 #11; Kerem, 1989 #14). L'anomalie génétique à l'origine de la maladie est enfin découverte, il s'agit du gène CFTR, localisé sur le bras long du chromosome7 au locus q31.

1.2.2. Une maladie génétique

Le gène *CFTR* dont les mutations sont responsables de la maladie de la mucoviscidose code pour une glycoprotéine transmembranaire composée de 1480 acides aminés (Kerem, *et al.* 1989, Riordan, *et al.* 1989, Rommens, *et al.* 1989). Cette protéine forme un canal à ions chlorure localisé au pôle apical des cellules épithéliales glandulaires des voies aériennes, intestinales, reproductrices, ainsi que sur les glandes de sécrétion externe telles que les glandes sudorales, pancréatiques et salivaires (*Figure 1.2*).

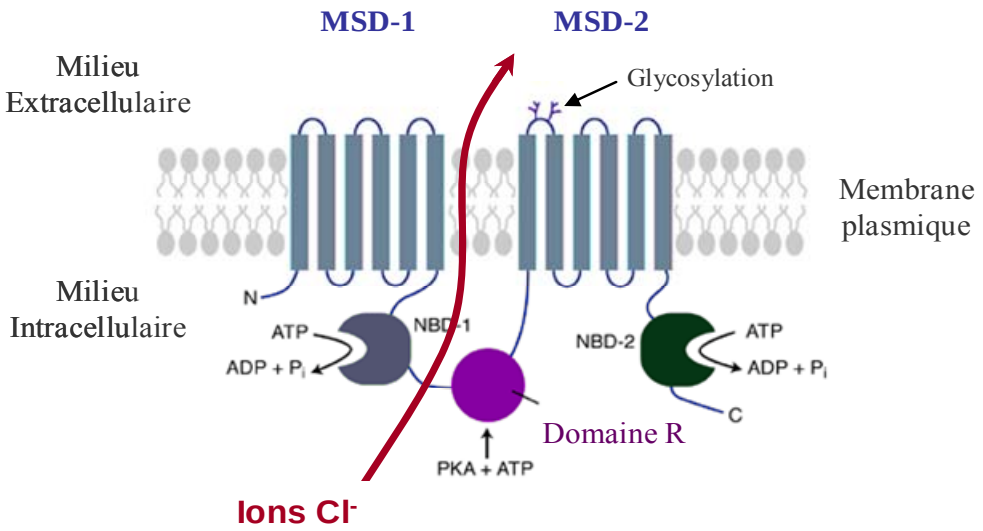


Figure 1.2. Représentation schématique de la protéine CFTR (Cystic Fibrosis Transmembrane conductance Regulator) (Davidson and Dorin 2001).

Elle appartient à la superfamillle des ABC transporteurs et contient deux domaines susceptibles de lier l'ATP (NBD ou Nucleotide Binding Domain) et deux domaines transmembranaires (TMD). Ces derniers sont reliés par un domaine cytoplasmique de régulation (domaine R) contenant de nombreux résidus chargés et la majorité des sites potentiels de phosphorylation.

Depuis le clonage du gène en 1989, 1571 mutations réparties tout au long du gène ont été identifiées et sont recensées par un consortium international : le « Cystic Fibrosis Genetic Analysis Consortium » (http://www.genet.sickkids.on.ca./cftr/app). La délétion de la phénylalanine en position 508, notée ΔF508, est la plus fréquente et représente 70% des mutations observées chez les patients mucoviscidosiques (patients CF) (Kerem, *et al.* 1989). La structure aux rayons X de la protéine CFTR humaine reste à déterminer. Néanmoins, la structure cristallographique du domaine de liaison à l'ATP de la souris, NBD1, a été résolue (Lewis, *et al.* 2004). Plus récemment, des données obtenues par microscopie électronique (résolution de 2 nm) sur la protéine CFTR recombinante humaine intégrale, montrent une forte homologie structurale avec un transporteur ABC de cellules eucaryotes, la P-glycoprotéine.

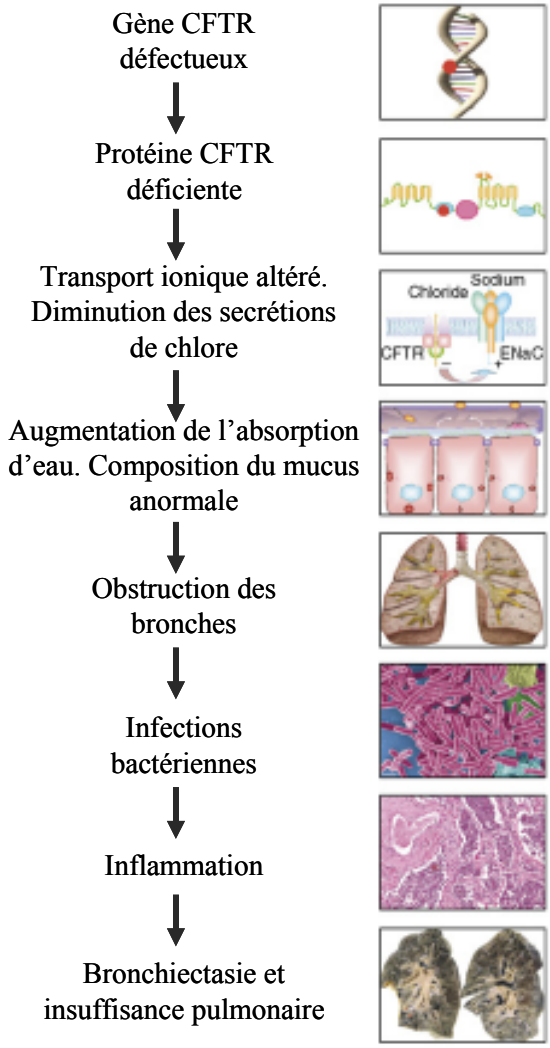


Figure 1.3. La mucoviscidose

Cascade du gène défectueux à l'insuffisance pulmonaire.*D'apres (Amaral and Kunzelmann 2007)*

1.2.3. Défaillance de la clairance mucocilliaire

Il existe chez l'individu sain, un mécanisme de défense dans les voies pulmonaires faisant intervenir la clairance mucociliaire d'une part et l'activité antimicrobienne exercée par le liquide de surface pulmonaire d'autre part. En battant de façon coordonnée, les cils des cellules

épithéliales mettent en mouvement une couche de liquide bistratifiée, appelée liquide de surface pulmonaire ou ASL (Airway Surface Liquid). L'ASL est composée d'une couche supérieur visqueuse formée de mucines secrétées, et d'une phase aqueuse profonde, la PCL (Periciliary Liquid Layer) dans laquelle battent les cils de la face apicale des cellules épithéliales (Boucher 2002, Matsui, *et al.* 1998). Lorsqu'une bactérie est inhalée, elle se fixe à la couche de mucus grâce à des interactions avec les mucines. De part leur grande diversité de chaînes glucidiques, ces macromolécules peuvent interagir avec une large gamme de particules inhalées (Knowles and Boucher 2002). Le battement ciliaire permet alors de mettre en mouvement la couche de mucus et d'éliminer les bactéries des voies respiratoires vers le naso-pharynx afin que celles-ci soit expectorées. En parallèle, les cellules respiratoires secrètent des peptides et des protéines ayant des propriétés protectrices tels que des anticorps, des antiprotéases ainsi que des cellules phagocytaires du système immunitaire telles que les neutrophiles polymorphonucléaires (PMNs).

Au cours de la mucoviscidose, les poumons qui constituent normalement un environnement stérile, sont continuellement envahis par les pathogènes inhalés. L'augmentation de la viscosité du mucus entraîne une défaillance de la clairance mucociliaire (Boucher 2002, Matsui, *et al.* 1998). La modification ionique observée à la surface des poumons inhibe l'activité des peptides anti-microbiens. Smith et al ont montré que l'ASL collectée in vitro à partir de cellules épithéliales CF en culture ont une activité bactéricide diminuée qui peut revenir à la normale après diminution du contenu en sel de ce liquide (Smith, *et al.* 1996). Ces disfonctionnement créent un environnement propice à la colonisation des poumons par des bactéries qui sont alors capables de se multiplier et d'induire une infection chronique. Finalement les infections pulmonaires sont la principale cause de maladie et de mortalité chez les patients atteints de mucoviscidose (*Figure 1.3*).

1.2.4. Importance des sucres dans la mucoviscidose

Les glycoconjugés jouent un rôle important dans les phénomènes de reconnaissance moléculaire entre l'hôte et le pathogène. Ce sont des molécules constituées d'une partie glucidique (mono- ou oligo- saccharide) liée de manière covalente à une partie non glucidique (aglycone). Selon la nature de la partie aglycone, on peut mettre en évidence trois classes de glycoconjugués : les glycolipides, glycoprotéines et peptidoglycanes. Les glygoconjugués de la surface cellulaire, forment le glycocalyx, une structure filamenteuse légère, maillée qui

recouvre la surface cellulaire (*Figure 1.4*). Ces sucres sont autant de marqueurs et de récepteurs qui confèrent une spécificité et une identité à la cellule.

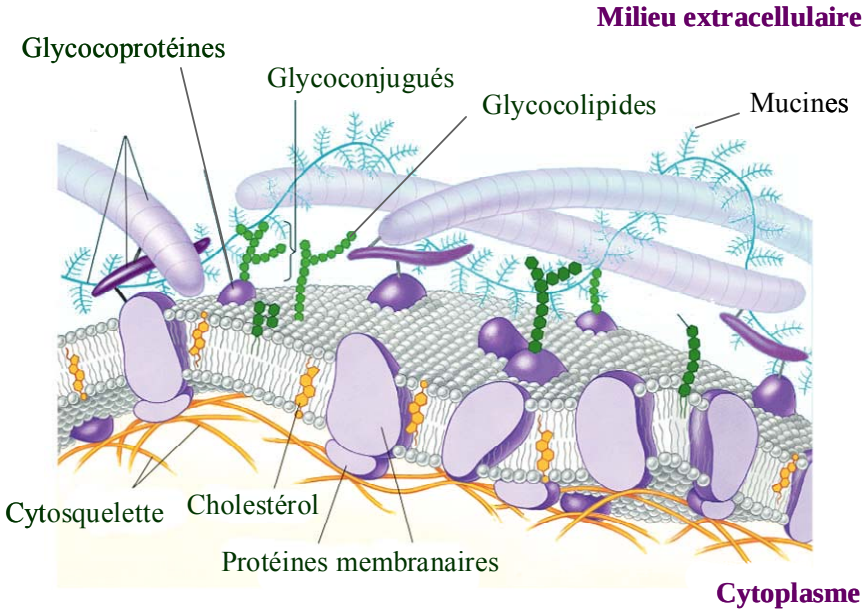


Figure 1.4. Représentation de la surface cellulaire.

D'après Copyright© 2002 Pearson Education, Inc., publishing as Benjamin Cummings

Le mucus est un gel très hydraté (95 % d'eau) qui tapisse la lumière des organes creux de l'organisme comme le tractus digestif, les voies respiratoires ou l'appareil génital. Chez l'individu sain, un film de mucus, recouvre l'épithélium respiratoire depuis les voies aériennes supérieures jusqu'aux bronchioles terminales neutralisant ainsi les toxines, polluants et microorganismes inhalés. Le mucus contient de l'eau et des sels mais aussi des lipides, des glycoprotéines : les mucines. On y trouve également des peptides et des protéines ayant des propriétés protectrices tels que des peptides antimicrobiens (lysozyme et défensines) capables de détruire la paroi bactérienne ainsi que des lactoferrines qui chélatent le fer nécessaire au métabolisme bactérien. La famille des mucines se divise en deux classes, les mucines sécrétées qualifiées de « vraies mucines » et les mucines membranaires désignées par « mucin-like » par opposition aux mucines solubles du mucus.

Ancrées à la membrane cytoplasmique, les mucines membranaires, sont produites par des cellules épithéliales non spécialisées de l'organisme. En raison de leur conformation rigide et étirée, ces molécules chargées négativement se projettent loin à l'extérieur de la

cellule et forment des structures de 1.2 à 2µm de hauteur qui dépassent largement l'ensemble des molécules du glycocalyx (épaisseur estimée à environ 10nm) (*Figure 1.5*).

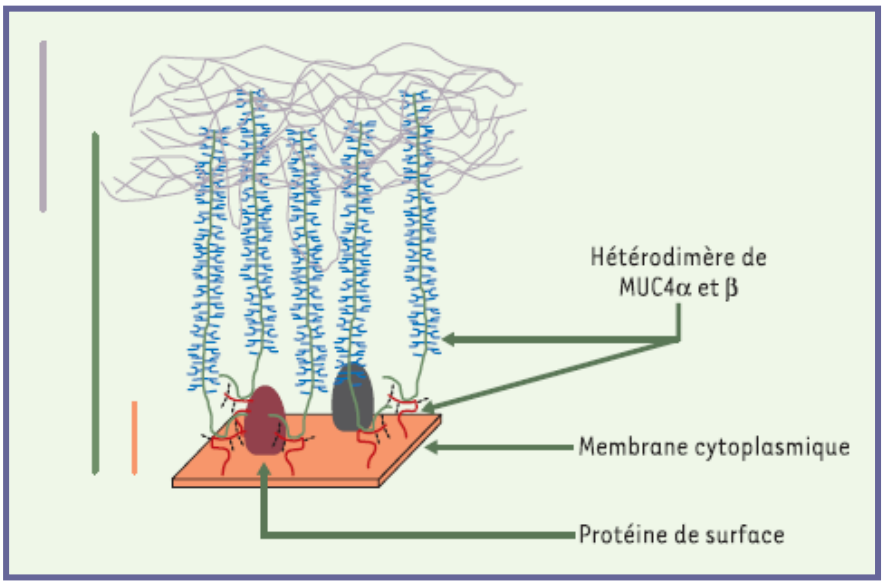


Figure 1.5. Mucines membranaires et protéction cellulaire.

Ce schéma montre une première ligne de défense (trait vertical gris) constituée par le gel de mucus (mucines secrétées) Une deuxième ligne de défense (trait vertical vert) est constituée par les mucines membranaires. MUC4 est représentée à la surface cellulaire sous la forme d'une structure hétéro-dimérique comportant un domaine Muc4α (en vert) et un domaine MUC4β associé à la membrane, représenté (en rouge). Ces deux domaines sont liés de façon non covalente. La taille de MUC4 varie de 1.2 à 2µm selon les individus. Le trait vertical orange délimite le glycocalyx d'une épaisseur estimée à 10nm. *http://id.erudit.org/iderudit/008425ar*

Les mucines solubles, sécrétées par les cellules caliciformes de l'épithélium bronchique et par les cellules muqueuses des glandes sous-muqueuses, sont le constituant macromoléculaire majoritaire du mucus. Ce sont de longues glycoprotéines filamenteuses (quelques nanomètres à quelques micromètres), flexibles, de très haut poids moléculaire. Chaque mucine possède un cœur protéique : l'apomucine codée par les gènes MUC, sur lequel se greffent des chaînes glucidiques (Lamblin, *et al.* 1991).. Les apomucines des voies respiratoires humaines sont codées par au moins huit gènes, MUC1, MUC2, MUC4, MUC5B, MUC5AC, MUC7, MUC8 et récemment caractérisé MUC19 (Porchet and Aubert 2004). Les régions terminales, de ces protéines, riches en cystéines sont probablement impliquées dans l'établissement de ponts disulfure intra et intermoléculaires. La région centrale, riche en acides aminés de type proline, sérine et thréonine, est formée par la répétition "en tandem" de motifs

plus ou moins conservés sur lesquelles viennent se greffer des chaînes glucidiques (*Figure 1. 6*).

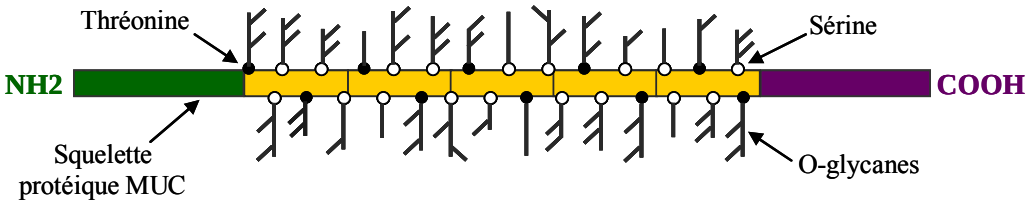


Figure 1.6. Représentation schématique d'une mucine sécrétée. Description du squelette protéique MUC et des O-glycans.

Le squelette protéique MUC est typiquement constitué d'un domaine -NH2 terminal (en vert), d'un domaine central avec un nombre élevé d'unités de répétition, TR (en jaune) et enfin d'un domaine -COOH terminal (en violet). Un grand nombre de O-glycans sont attachés à des résidus thréonine (ronds noirs) ou sérine (ronds blancs) au niveau des TR. La taille des O-glycans présentent une hétérogénéité en raison d'une biosynthèse incomplète au cours de l'élongation et de la terminaison des O-Glycans ou d'une différence de préférence de substrat des glycosyl transférases variées pour les sites spécifiques sérine et thréonine *(Rose and Voynow 2006)*

L'attachement de ces structure sur l'apomucine se fait par liaison O-glycosydique entre un résidu N-acetylgalactosamine (GalNAc) et un acide aminé hydroxylé (sérine ou thréonine). Les chaînes latérales glucidiques qui représentent classiquement 80-85% du poids moléculaire de la glycoprotéine mature sont composées de 2 à 20 monosaccharides. Elles sont principalement constituées de galactose, fucose, N-acétylgalactosamine, N-acétylglucosamine et d'acide sialique. On trouve au niveau des extrémités non réductrices de ces chaines les épitopes glucidiques des groupe sanguins A, B, H ou O et H-sulfaté, les épitopes Lewis a, Lewis b, Lewis x ou Lewis y ainsi que les sialyl-et / ou sulfo- Lewis a et Lewis x. La biosynthèse de ces glycoconjugués complexes nécessite un processus multi-étapes impliquant de nombreuses glycosyl- ou sulfotransférases.

De nombreuses études rapportent une modification de la glycosylation (mucines, glycoprotéines ou glycolipides) dans les epithelia étudiés de patients atteints de mucoviscidose.. En 1959, une augmentation du rapport fucose / acide sialique dans les mucines intestinales de patients CF a été observée. Cette observation a ensuite été confirmée par plusieurs méthodes sur différents types cellulaires : fibroblastes, intestins, poumons, glandes salivaires… (Rhim, *et al.* 2000, Scanlin and Glick 1999). L'augmentation de la fucosylation apparaît à la fois dans les glycoconjugués des cellules épithéliales et dans les

mucines pulmonaires des patients CF. Le fucose est le plus souvent lié en position 3 ou 4 de la N-AcétylGlucosamine (Fucα1,3/4GlcNAc), correspondant respectivement aux structures Lewis x et Lewis a. Une augmentation significative de la quantité de Fucα1,6GlcNAc a également été observée. Par contre la quantité de fucose lié en α1,2 au galactose (épitope groupe sanguin O/H) est diminuée dans les cellules des voies aériennes des patients CF. La sialylation est elle aussi modifiée et des études ont montré une diminution d'acide sialique sur les glycoprotéines des cellules épithéliales CF (Dosanjh, *et al.* 1994, Rhim, *et al.* 2000). On observe deux fois plus de composants sialylés et sulfatés dans les mucines CF que dans les non-CF (Lamblin, *et al.* 2001, Xia, *et al.* 2005). Cependant, certaines équipes ont noté une absence de modification de la O-glycosylation dans les cellules CF (Hill, *et al.* 1997, Leir, *et al.* 2005).

Différentes hypothèses ont été proposées pour expliquer ce phénotype glucidique altéré dans la mucoviscidose. Des études suggèrent que ce phénotype est modulé par l'expression de la protéine CFTR. En effet, la quantité normale de Fucα1,3/4GlcNac Fucα1,2Gal et d'acide sialique peuvent être rétablies par transfection de cellules épithéliales CF par l'ADNc codant pour la protéine CFTR sauvage. Mais des résultats contradictoires ont également été obtenus (Jiang, *et al.* 1997). Il a été proposé que la protéine CFTR pourrait influencer directement la glycosylation des mucines et des glycoconjugués épithéliaux. En effet, un mauvais fonctionnement du canal à chlorure entraînerait une acidification du compartiment Trans Golgi. L'activité cinétique des glycosyl- et sulpho-transférases sensibles au pH serait alors perturbée (Barasch, *et al.* 1991). Cependant d'autres équipes montrent que le pH de ce compartiment n'est pas influencé par l'expression de CFTR mutant ou sauvage (Gibson, *et al.* 2000, Seksek, *et al.* 1996). Brockhausen et al. ont observé l'activité de glycosyl- et sulfo-transferase participant à la synthèse des mucines et aucune différence d'activité n'a été observée entre les cellules exprimant ou non CFTR WT ou encore Δ508CFTR (Brockhausen, *et al.* 2001).

D'autre part, il a été démontré que l'inflammation pourrait induire des modifications de glycosylation dans les poumons, comme la surexpression des épitopes sialyl-Lewis x sur mucines bronchiques (Davril, *et al.* 1999). De plus, les cytokines-α pro-inflammatoires TNF peuvent induire une augmentation de l'expression et l'activité de certains sialyl-, et fucosyl - sulfotransférases d'explants bronchique et respiratoires humains et de lignée cellulaires glandulaires (MM-39) (Delmotte, *et al.* 2002). L'incubation de fragments de mucus bronchique humain avec les Interleukines 6 ou 8 (IL-6, IL-8) augmentent significativement la quantité de Sialyl-Lewis x et de 6-sulfo-sialyl-Lewis x (Groux-Degroote, *et al.* 2008). Les

deux principaux agents pathogènes des patients CF, *P. aeruginosa* et *Hemophilus influenza,* possèdent des lectines qui reconnaissent les extremité Fucα1-3 et les asialoglycoconjugués.. L'altération de la glycosylation chez les patients CF pourrait donc favoriser la colonisation des voies respiratoires par ces microorganismes. En effet, les asialo gangliosides contiennent une séquence tetrasaccharidique, Galβ1-3GalNAcβ1-4Galβ1-4Glc, qui constitue le site d'ancrage de la piline et de la flagelline de *P. aeruginosa*. L'implication de aGM1 dans l'adhérence de *P. aeruginosa* est plus importante au niveau de l'épithélium respiratoire CF que non-CF (Comolli, *et al.* 1999, de Bentzmann, *et al.* 1996).

1.2.5. Infections bactériennes

La présence de secrétions très visqueuses et peu mobiles au niveau des voies aériennes ainsi qu'un ensemble mucocilliaire défaillant constituent donc un environnement propice aux infections bactériennes. L'obstruction progressive des petites voies respiratoires induit une réponse inflammatoire importante et favorise les surinfections microbiennes. Pratiquement constantes et évoluant sous forme d'exacerbations sur fond chronique, ces poussées infectieuses surviennent très tôt dans la vie du patient et conditionnent le pronostic vital. La progression de l'atteinte des poumons est progressive et inéluctable. La structure même des poumons s'altère, des dilatations des bronches et une fibrose (cicatrisation anormale) apparaissent. A ce stade là, seule la transplantation d'organe peut sauver la vie du patient.

Avant la naissance, les poumons sont remplis de liquide amniotiques et les tissus semblent, histologiquement normaux (Sturgess and Imrie 1982). La colonisation par des germes a ensuite lieu très tôt dans la vie du patient et les espèces bactériennes impliquées varient au cours de la vie du malade (*Figure 1.7*).

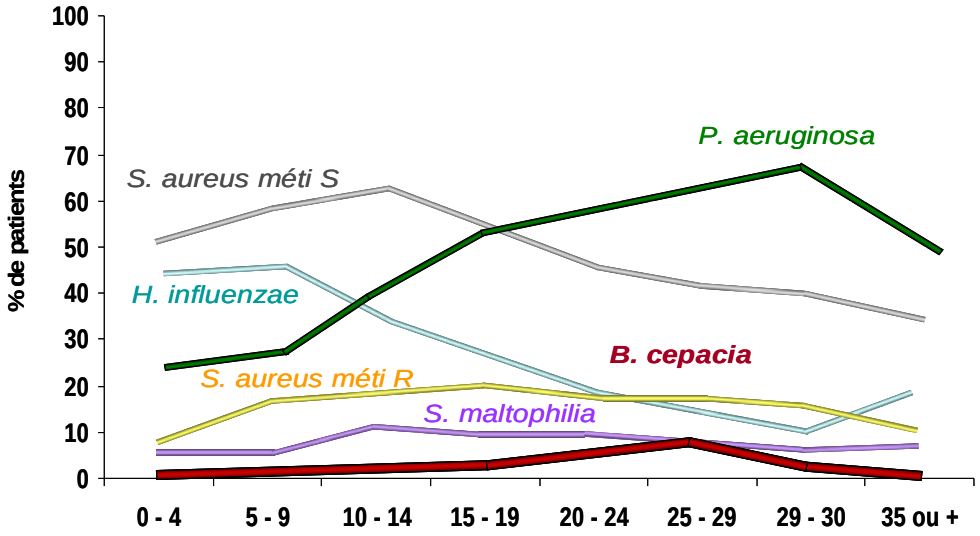


Figure 1.7. Prédominance selon l'âge, des bactéries infectant les voies aériennes des patients atteints de mucoviscidose en 2004. Vaincre *la Mucoviscidose et Ined, Paris, 2006].*

Traditionnellement, les germes le plus souvent retrouvés sont *Staphylococus aureus et Haemophilus influenza* au stade précoce de la maladie. Plus tard, *P. aeruginosa* devient majoritaire et infecte environ 80 % des patients, et ce, de façon définitive (Aebi et al., 1995). La détérioration préalable de l'épithélium respiratoire par Staphylococus aureus et *H. influenzae* faciliterait l'adhésion de *P. aeruginosa* (Lyczak*, et al.* 2002).

Certaines bactéries du complexe *Burkholderia cepacia* (BCC), représentent aujourd'hui un grave danger pour les patients mucoviscidosiques, du fait de leur virulence, leur potentiel de transmission et de leur résistance exceptionnelle aux antibiotiques. La plupart des études portent sur ces quatre microorganismes mais d'autres germes tels que *Stenotrophomonas maltophilia, Achromobacter xylosoxidans, Alcaligenes xylosoxidans et Klebsiella spp. Streptococcus pneumoniae, Mycobacterium spp., Aspergillus fumigatus, Pandoraea apista, Ralstonia spp., Inquilinus limosus* et une variété d'autres organismes rares ont également été retrouvés dans les crachats des patients CF. La diversité de cette communauté de microorganisme a augmenté depuis les dix dernières années en raison de l'augmentation de l'espérance de vie et de l'amélioration des méthodes de détection de ces bactéries (Rogers*, et al.* 2004). Les virus pathogènes comme le virus respiratoire syncytial (VRS), l'adénovirus, la grippe ainsi que des champignons tels que Candida et Aspergillus (Haase*, et al.* 1991)sont également fréquents.

1.2.5.1. Histoire des infections à B. cepacia

Il y a environ 30 ans, les bactéries du complexe BCC émergent en tant qu'agents pathogènes problématiques pour les patients mucoviscidosiques. En 1984, le premier rapport important concernant les infections à *P. cepacia* chez les patients CF est publié (Isles*, et al.* 1984). Dans l'année qui suit, il est confirmé que les infections par ces bactéries posent un problème dans les centres pour patients atteints de la mucoviscidose. Ces premiers rapports décrivent aussi la nature virulente des infections à *B. cepacia*, qui peut conduire à une détérioration rapide, incontrôlable et fatale de l'état clinique chez environ 10% des patients. Au milieu des années 1990, plusieurs études rapportent la prévalence d'un clone épidémique au cours des infections chez les patients CF Canadiens et Britanniques. Il sera évoqué sous plusieurs noms. En 1994, il est initialement désigné d'Electrophorétique Type 12, ou ET12, sur la base d'analyses électrophorétiques des variants d'une enzyme, l'estérase (Johnson*, et al.* 1994). Il est aussi décrit sous le clone câble pili en raison de la présence du gène cblA

codant pour un pili de classe II dans son génome (Sun, *et al.* 1995) ou sous un clone « Random Amplified Polymorphic DNA » de type II (Mahenthiralingam, *et al.* 1996). Il est probable que les isolats décrits à Toronto dès 1971 étaient déjà des clones épidémiques et virulents de type ET12 (Isles, *et al.* 1984). Plusieurs autres clones épidémiques ont été décrits, et la majorité d'entre eux appartiennent à l'espèce *B. cenocepacia*. La transmission de patient à patient (Govan, *et al.* 1993) et la virulence des infections ainsi que le taux élevé de mortalité post-infection (Mahenthiralingam, *et al.* 2001)ont été associés à plusieurs clones de *B. cenocepacia* IIIA. Aux États-Unis, d'autres clones épidémiques comme celui du district de Philadelphie-Columbia (PHDC) ou celui du « Mid-West » ont également été identifiés et classés dans le sous-groupe phylogénétique *B. cenocepacia* IIIB (Chen, *et al.* 2001). En plus de sa prévalence dans la population CF Américaine, la souche PHDC a aussi été trouvée chez les patients européens (Coenye, *et al.* 2004).

1.2.5.2 Epidémiologie des infections à B. cepacia

Même si tous les membres du complexe BCC ont été retrouvés dans les poumons de patients atteints de la mucoviscidose, *B. cenocepacia* et *B. multivorans* sont les deux principales bactéries responsables d'infections chez ces malades. Chez 153 patients CF français infectés par une bactérie du complexe, 45% et 52% des patients étaient infectés par *B. cenocepacia* et *B. multivorans* respectivement (Brisse, *et al.* 2004). La proportion de patients infectés par *B. cenocepacia* approche les 80% au Canada (Speert, *et al.* 2002). Les infections à *B. cenocepacia*, et en particulier par le clone ET12, sont associées à une diminution significative des chances de survie du patient, alors que les conséquences des infections à *B. multivorans* sont similaires à celles observées pour *P. aeruginosa* (Jones, *et al.* 2004). Aux Etats-Unis, une étude faite sur 1218 patients CF a montré que, *B. multivorans* et *B. cenocepacia* représentent à elles deux environ 85% des infections à *Burkholderia* (Reik, *et al.* 2005). Une étude récente montre que depuis 2000, les infections à *B. multivorans* sont plus fréquentes que celles causées par *B. cenocepacia* dans la population CF du Royaume-Uni (Govan, *et al.* 2007). Aujourd'hui, les chiffres laissent penser que la propagation des souches transmissibles de *B. cenocepacia* ainsi que d'autres souches épidémiques du complexe a été réduite grâce à l'application des mesures strictes de contrôle des patients infectés par ces bactéries (Saiman and Siegel 2004).

Globalement, les infections à *B. cepacia* sont associées à un diagnostic plus difficile, à de plus longues hospitalisations et à un risque accru de mortalité pour les patients CF. Les

infections à *B. cepacia* peuvent aller du simple portage à l'infection chronique, similaire à celles observées pour les autres agents pathogènes typiques de ces infections. Certains patients subissent une détérioration rapide et incontrôlable de leur état clinique, avec pneumonie nécrosante et septicémie aboutissant à une mort prématurée (Isles, *et al.* 1984); c'est ce qu'on appelle le « syndrome cepacia ». La bactériémie, associée à ce syndrome ne se produit que très rarement au cours des infections par d'autres agents pathogènes comme *P. aeruginosa* (Banerjee and Stableforth 2000). Une fois l'infection chronique établie, le traitement est très difficile, en raison de la résistance naturelle de ces bactéries aux antibiotiques (Nzula, *et al.* 2002). En plus de la nature virulente de ces infections, les études épidémiologiques montrent que les bactéries du complexe peuvent être transmises de patient à patient (Govan, *et al.* 1993, LiPuma, *et al.* 1990). La menace que représente cette bactérie a conduit le personnel médical à appliquer des mesures sanitaires drastiques afint de limiter cette transmission interhumaine (ségrégation des patients colonisés, port de masque …).

1.3. Virulence de B. cepacia

1.3.1. Comportement social de Burkholderia

1.3.1.1. Le Quorum sensing

La nécessité de parvenir à une densité de population suffisante pour surmonter les défenses immunitaires de l'hôte avant d'établir l'infection est un élément crucial pour les bactéries. Beaucoup de pathogènes à gram-négatif régulent la production de leurs facteurs de virulence extracellulaire par un mécanisme de communication, appelé le Quorum Sensing (QS). Ce mécanisme reposant sur la densité cellulaire implique la production de petites molécules signales diffusibles : les N-acyl-homosérine-lactones (acyl-HSL) et permet de contrôler l'expression de facteurs de virulence en activant leur transcription. Ces HSL sont synthétisées par une HSL synthase, qui est codée par un gène de type « I ». Lorsque leur concentration atteint un certain seuil, les HSLs sont capables d'activer un régulateur transcriptionnel de type « R » qui déclenche l'expression de gènes cibles. Ce mécanisme permet aux bactéries de rester « invisible » vis-à-vis du système immunitaire de l'hôte jusqu'à ce que la densité cellulaire soit suffisante pour établir l'infection.

Certaines bactéries du complexe BCC ont la particularité de posséder deux systèmes de régulation de ce type: le système CepI/CepR (CepIR) et le système BviI/BviR (BviI/R).
Le premier système, CepIR, a été décrit pour la première fois chez *B. cenocepacia* K56-2 (Lewenza, *et al.* 1999). Ce système est composé d'un régulateur transcriptionnel (CepR) et d'une AHL synthase (CepI) responsable de la synthèse des N-octanoyl homosérine-lactones (C8-HSL) et des N-hexanoyl homosérine-lactones (C6-HSL) dans les proportions 10 :1 (Gotschlich, *et al.* 2001). Si la densité de population est faible, les bactéries produisent les molécules signal d'AHLs à un niveau basal. Lorsque cette densité augmente, les molécules d'AHLs s'accumulent dans le milieu. Si la concentration en HSL atteint un seuil critique, témoin d'une forte population bactérienne, les HSLs forment un complexe avec CepR, ce qui active la transcription de plusieurs gènes de virulence de manière synchrone dans toute la population bactérienne (*Figure 1.8*).

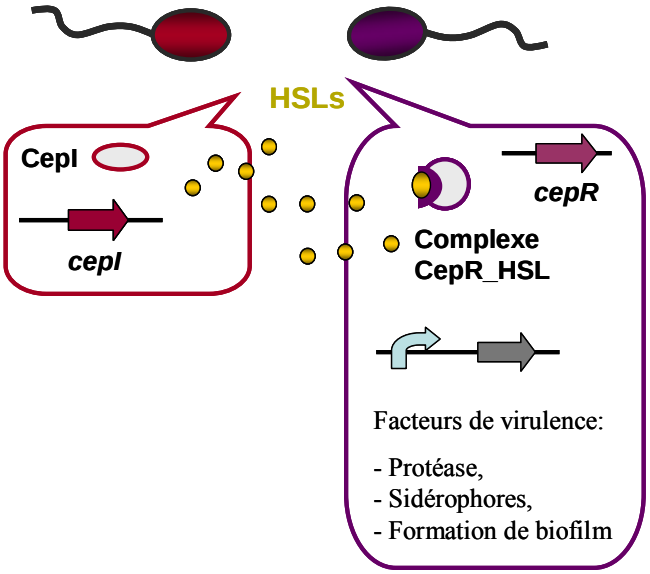


Figure 1.8. Représentation schématique du système de Quorum Sensing de type CepI/CepR des bactéries du complexe *Burkholderia*.

Chez les bactéries du complexe, différents facteurs extracellulaires comme les toxines, protéases, lipases et sidérophores, mais aussi la mobilité de type 'swimming » et la formation de biofilm sont sous le control de ce système (Venturi, *et al.* 2004). Des études faites chez la souris et le rat (Sokol, *et al.* 2003) ainsi que chez *Caenorhabditis* (Kothe, *et al.* 2003) ont montré que le système CepIR du QS joue un rôle crucial dans la virulence de la bactérie. Le système CepIR a été retrouvé chez les 9 espèces du complexe (Lutter, *et al.* 2001).

Certaines bactéries telles que *B. vietnamiensis*, possèdent un second système comprenant les gènes *bviI* et *bviR* qui codent respectivement pour une enzyme BviI nécessaire à la synthèse des N-decanoyl homosérine-lactones (C10-AHL) et pour la protéine régulatrice(Conway and Greenberg 2002). L'interaction éventuelle entre les systèmes BviI /

BviR et CepI / CepR reste à démontrer. Chez la souche virulente ET-12 de *B. cenocepacia*, un troisième système de QS a été identifié dans un îlot de pathogénicité (Baldwin*, et al.* 2004).. Il est encore difficile de savoir si ce système, composé de l'AHL synthase CciI et du récepteur CciR, fonctionne indépendamment du système CepIR.

1.3.1.2 Formation de biofilms

Le traitement antibiotique des infections pulmonaires est rendu d'autant plus difficile que certaines souches bactériennes se développent au sein d'un biofilm. Le biofilm est une communauté bactérienne, adhérée à une surface (biotique ou abiotique) souvent hétérogène, structurée, dans laquelle les bactéries sont protégées d'un environnement hostile. La matrice extracellulaire de cette structure est généralement composée d'exopolysaccharides, de protéines et d'acide nucléiques. Cette architecture ajoute encore à l'obstruction et à la viscosité des poumons CF. Trois étapes peuvent être distinguées dans la formation du biofilm: une étape initiale d'attachement à la surface des muqueuses, une étape de colonisation-prolifération aboutissant à la formation de micro-colonies et enfin une étape de croissance et de structuration du biofilm (*Figure 1.9*).

Il a été démontré que *B. cenocepacia* est capable de croître sous forme de biofilm dans les poumons des patients CF (Mathee*, et al.* 1999). On sait aujourd'hui que toutes les bactéries du complexe sont capables de former de telles structures *in vitro*, et que ce processus est associé au Quorum Sensing (Conway and Greenberg 2002). Une telle organisation confère un avantage important à ces bactéries pathogènes. Tout d'abord, le biofilm représente une barrière physique et la pénétration des antibiotiques y est très limitée. Par exemple, les bactéries sessiles de *B. cepacia* sont quinze fois plus résistantes à la ceftazidime et à la ciprofloxacine que les bactéries planctoniques (Desai*, et al.* 1998). Cette organisation en biofilm confère également aux bactéries une résistance vis-à-vis du système immunitaire de l'hôte (anticorps et macrophages).

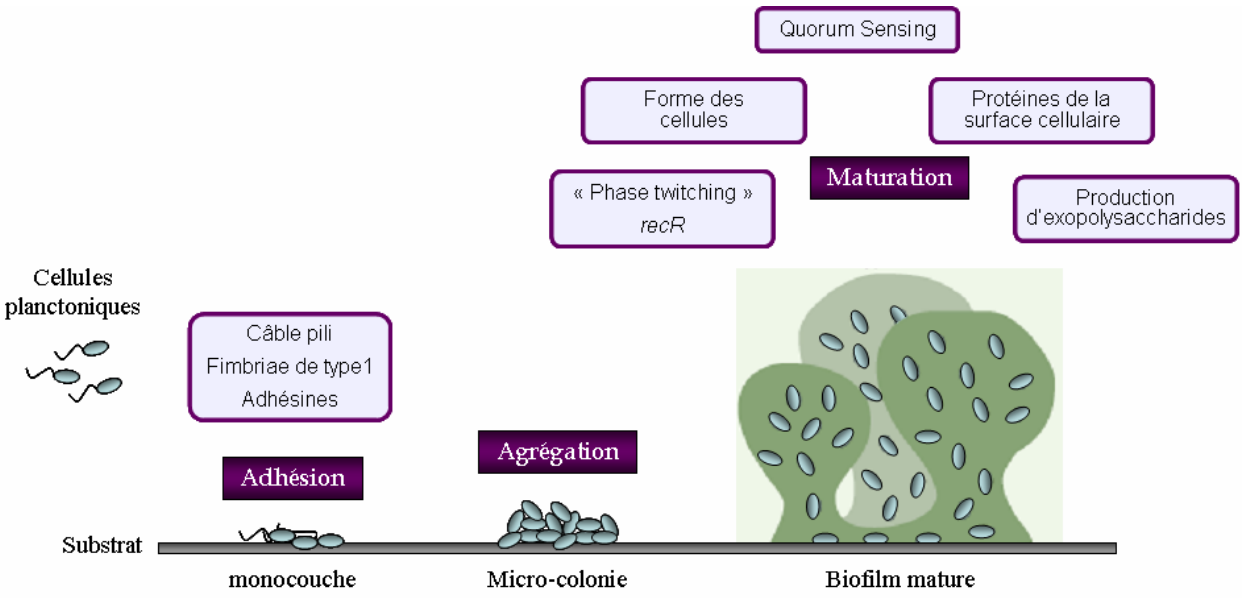


Figure 1.9. Modèle de formation d'un biofilm par les bactéries *du genre Burkholderia*.

Les bactéries planctoniques s'attachent sur une surface en réponse à un des stimuli environnementaux et nutritionnels. L'approche de cette surface nécessite la mobilité flagellaire. Puis l'attachement initial à la surface fait intervenir un certain nombre de structures exposées à la surface telles que des adhésines, le cable pili et les fimbriae de type1. Finalement, les micro-colonies évoluent en un biofilm mature dans lequel elles s'entourent d'une épaisse matrice d'exopolysaccharides. Les gènes impliqués dans la maturation du biofilm appartiennent à différentes classes : (i) gènes affectant la forme des cellules, (ii) gènes impliqués dans la production d'exopolysaccharides, (iii) gènes codant pour les protéines de la surface cellulaire, (iv) gènes impliqués dans le changement de phase, (v) gènes impliqués dans la communication intercellulaire (Quorum Sensing). La structuration des micro-colonies en biofilm mature implique la communication entre les cellules via les molécules signal du QS, les acyl-Homosérine Lactone (acyl-HSL). *Modifié d'après Burkholderia,Molecular Microbiology and Genomics, T. Coenye & P. Vandamme, 2007*

1.3.1.3. Interaction avec P. aeruginosa.

Au cours des infections, les bactéries du complexe *BCC* et *P. aeruginosa* sont capables de former des biofilms mixtes dans les poumons des patients CF (Saiman, *et al.* 1990). Les troubles pulmonaires causés par la co-colonisation par ces deux organismes sont plus graves que ceux causés par la seule bactérie *P. aeruginosa* (Tomlin, *et al.* 2001). Les deux bactéries utilisent le Quorum Sensing et les molécules signal d'HSLs pour coordonner l'expression de leurs facteurs de virulence avec la formation du biofilm. L'addition du surnageant d'une culture de *P. aeruginosa* au milieu utilisé pour la croissance de *B. cenocepacia* permet d'augmenter la production de sidérophores, lipases et protéases de cette dernière (McKenney, *et al.* 1995). Plus récemment, il a été démontré que les molécules d'HSLs de *P. aeruginosa* stimulent l'expression de protéases et de *cepI* chez *B. cenocepacia* K56-2 (Lewenza, *et al.*

2002). Au sein des biofilms mixtes, les bactéries du complexe BCC sont capables de détecter les molécules signal d'AHL sécrétées par *P. aeruginosa* alors que l'inverse n'est pas vrai (Riedel, *et al.* 2001). Ces observations sont en accord avec le fait que la plupart des infections à *B. cenocepacia* interviennent chez des patients déjà colonisés par *P. aeruginosa*. Les deux bactéries pourraient donc coordonner la régulation de leurs facteurs de virulence et influencer la progression des troubles pulmonaires qui leur sont attribuables.

1.3.2. Facteurs de virulence

Les facteurs de virulence permettant d'expliquer les capacités de certaines bactéries du complexe BCC à coloniser les tissus pulmonaires sont encore mal connus. De nombreuses études récentes portent sur les mécanismes de pathogénicité et les facteurs de virulence putatifs des *Burkholderia*. (*Figure1.10*).

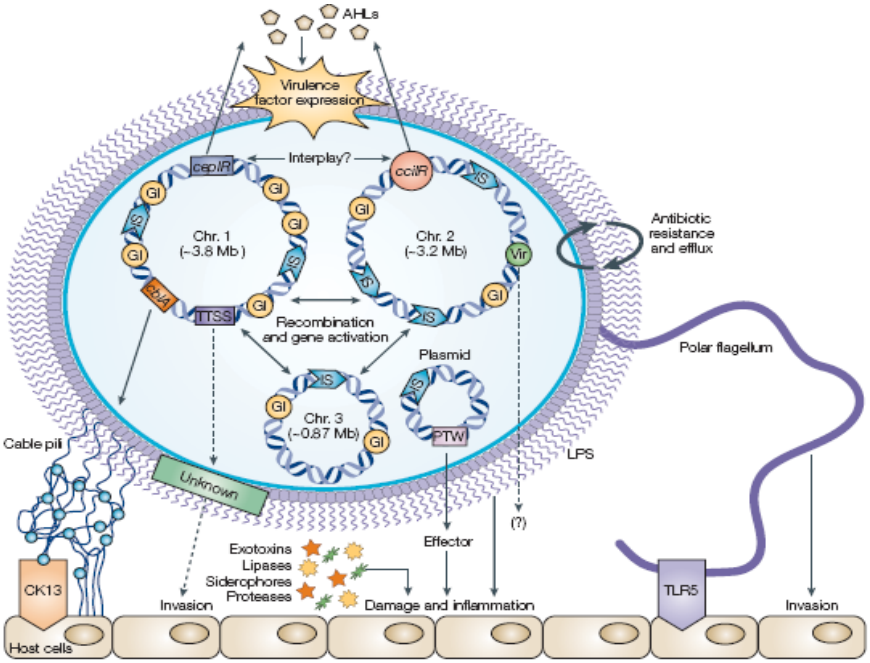


Figure 1.10. Structure génomique et principaux facteurs de virulence de la souche ET12 de *B. cenocepacia*.

Le schéma montre la structure du génome et les facteurs de virulence importants de *B. cenocepacia*. Le génome comprend trois grands réplicons chromosomiques (Chr. 1, Chr. 2 et Chr. 3) et un plasmide, la présence d'ilots génomiques (GI) sur chacun des trois chromosomes des séquences d'insertion (IS). Les facteurs de virulence et leur rôle ainsi que les interactions avec les cellules de l'hôte comprennent: le cable-pili et adhésine de 22 kDa (ronds bleus), les systèmes cepIR et cciIR du quorum-sensing, le système de sécrétion de type III (TTSS), la flagelline, plusieurs facteurs extracellulaires et associés à la surface, comprenant les lipopolysaccharide (LPS), exotoxines, lipases, protéases et sidérophores. D'après (Mahenthiralingam, *et al.* 2005)

1.3.2.1. Les lipopolysaccharides

Les lipopolysaccharides (LPS) sont des complexes macromoléculaires toxiques présents de manière constitutive dans la membrane externe de toutes les bactéries à Gram négatif chez qui ils constituent un facteur de virulence important. Ils présentent un certain nombre de propriétés telles que, l'immunogénicité, l'induction de cytokines pro inflammatoires, et la protection contre la phagocytose et le système du complément. Sur le plan structural, le LPS est constitué de trois éléments : le lipide A, une partie hydrophobe ancrée dans la membrane externe ; un noyau oligosaccharidique et enfin la chaîne O-spécifique, une partie O-polysaccharidiques débordant la membrane externe (*Figure 1.11*).

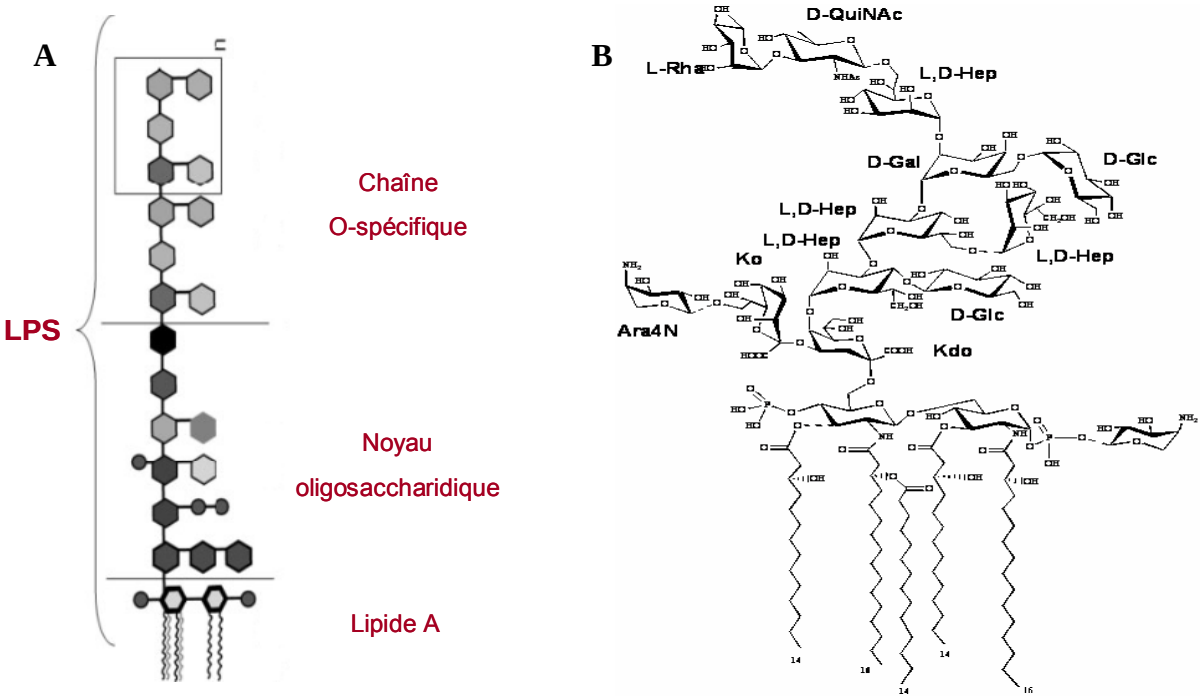


Figure 1.11. Structure du Lipopolysaccharide de *B. cenocepacia*

A. Structure générale du LPS des bactéries à gram négatif. Le LPS consiste en un domaine lipidiques, le *lipide A* enchâssé dans la membrane et lié de façon covalente à au *noyau oligosaccharidique* lui-même lié à une partie polysaccharidique : la *chaine O-spécifique.*
B. Structure complète du lipopolysaccharide de la souche J2315 de *B. cenocepacia*, clone ET-12. Les pointillés indiquent les substitution non stoechiométriques. L-RHa : L-Rhamnose. D-QuiNac : Ara4N : 4-amino-4-deoxy-4-amino-arabinopyranose. GlcN: glucosamine. Kdo : acide 3-désoxy-D-manno-2-octulosonique. (D'après De Soyza, *et al.* 2008)

Certaines souches du complexe ne possèdent pas de O-antigène (phénotype LPS-rugueux), ce qui les rend non-sérotypables. Sur un plan structural, le noyau oligosaccharidique des LPS des bactéries du genre *Burkholderia* contient moins d'acide 3-

désoxy-D-manno-2-octulosonique (Kdo) et de groupements phosphate que celui des autres bactéries à Gram-négatif (De Soyza, *et al.* 2008)

B. cenocepacia est naturellement résistante à de nombreux antibiotiques et il a été démontré que les LPS jouent un rôle important dans cette résistance. La partie 4-amino-4-deox-4-amino-arabinopyranose (Ara4N) liée aux résidus phosphate du lipide A serait impliqué dans la résistance aux peptides antimicrobiens cationiques ainsi qu'à la polymyxine (Cox and Wilkinson 1991, Shimomura, *et al.* 2003). Le cœur oligosaccharidique du LPS semble lui aussi jouer un rôle dans la résistance des bactéries aux antibiotiques comme la polymixine (Loutet, *et al.* 2006).

Des expériences d'infection faites chez le rat avec un mutant défectueux dans la formation du LPS complet, ont montré le rôle essentiel du cœur oligosaccharidique dans la survie in vivo des bactéries (Loutet, *et al.* 2006). L'activité endotoxine des LPS de *B. cepacia* est quatre à cinq fois plus importante que celle de *P. aruginosa* avec une augmentation de l'activité des neutrophiles et de la sécrétion des interleukine-8 (IL-8) des cellules épithéliales (Hughes, *et al.* 1997, Reddi, *et al.* 2003, Shaw, *et al.* 1995). Ce rôle pro-inflammatoire des LPS contribue certainement à la dégradation de l'état pulmonaire des patients atteints de mucoviscidose. L'importance de ces glycolipides dans la virulence et la résistance aux antibiotiques chez *B. cenocepacia* en fait une cible thérapeutique prometteuse.

1.3.2.2. Les exopolysaccharides

Plus de 80% des isolats de *Burkholderia* impliqués dans les infections bactériennes chez les patients CF produisent des exopolysaccharides, EPS (Richau, *et al.* 2000). Il a été montré in vitro que les EPS isolés à partir de la souche mucoïde « LPS rough » *B. cenocepacia* C9343 interfèrent avec les neutrophiles. Les EPS joueraient donc un rôle dans l'inefficacité de la clairance mucocilliaire de l'hôte (Bylund, *et al.* 2006).

1.3.2.3. Les sidérophores

Les sidérophores sont des molécules essentielles à la survie des microorganismes qui interviennent dans la captation du fer par la bactérie. Ce sont de petites molécules (500-1000 Da) non protéiques, sécrétées par la bactérie, très affines pour les ions Fe3+, qui entrent donc en compétition avec les protéines de liaison au fer de l'hôte (transferrine et lactoferrine).

Quatre types de sidérophores ont été identifiés chez *B. cepacia* : pyochéline, cépabactine, azurécheline et ornibactine. Les bactéries du complexe sont aussi capables d'utiliser des sidérophores exogènes qu'elles ne sont pas capables de synthétiser. Le rôle essentiel des sidérophores a été démontré dans des modèles d'infection chez le rat (Visser, *et al.* 2004).

1.3.2.4. Les systèmes de sécrétion

Les bactéries à gram négatif possèdent des systèmes de sécrétion leur permettant d'exporter les protéines toxiques dans l'environnement ou dans la cellule hôte. Cinq types de systèmes de sécrétion (TSS, Type Secretion System) ont été décrits (Gerlach and Hensel 2007) (*Figure 1.12*).

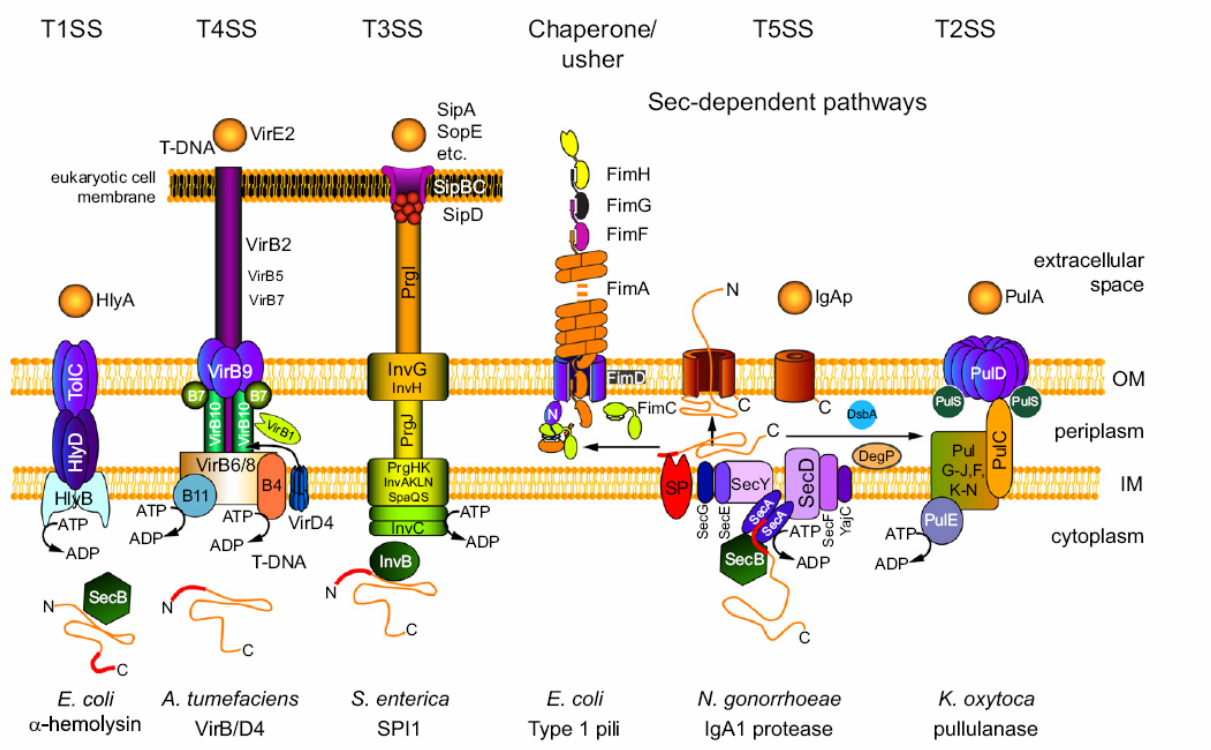


Figure 1.12. Systèmes de sécrétion des bactéries à gram négatif

La voie chaperonne / usher qui constitue l'étape finale de sécrétion des structures de surface est représentée avec l'exemple de l'assemblage du pili de type1 d'*E. coli*. Le Type I Secretion System (T1SS) est responsable de la sécrétion d'enzymes extracellulaires comme c'est le cas pour l'α-hémolysine chez *E coli*, Les T2SS, permet d'exporter des protéines possédant une séquence d'acides aminés hydrophobes, comme la pullulanase par *Klebsiella oxytoca. Le T3SS, ici représenté par le SST3 SPI-1 de Salmonella enterica* permet entre autre la sécrétion de protéines responsables de l'adhérence de la bactérie à la surface cellulaire. Le T4SS permet la sécrétion de complexes nucléoprotéiques ou protéiques à travers la membrane des bactéries à Gram négatif. Il est ici illustré par le système d'injection de l'ADN dans les cellules végétales par Agrobacterium tumefaciens. *Le T5SS qui constitue la voie privilégiée des autotransporters est représenté par le modèle de sécrétion de l'IgA1 protéase par Neisseria gonorrhoeae.* (Gerlach and Hensel 2007)

Récemment, un nouveau système de sécrétion appelé type VI (T6SS) a été décrit chez les trois bactéries *Vibrio cholerae, P. aeruginosa* et *B. mallei* (Mougous, *et al.* 2007, Pukatzki, *et al.* 2006, Schell, *et al.* 2007). Peu de données sont disponibles sur les systèmes de sécrétion de *B. cenocepacia*. Le T3SS est présent chez tous les membres du complexe excepté *B. cepacia sensu stricto* (Glendinning, *et al.* 2004).. La mutation de gènes impliqués dans le SST3 chez *B. cenocepacia* J2315 engendre une diminution de la virulence des bactéries chez la souris (Tomich, *et al.* 2003)

1.3.2.5. Les Flagelle et pili

Afin de déterminer le lieu le plus propice à la colonisation, les bactéries ont développé une sensibilité à leur environnement proche : environnement physicochimique (gradient de pH, de température, de lumière), présence de nutriments (sucres, acides aminés) ou de substances toxiques ou nuisibles ; c'est ce qu'on appelle le « chimiotactisme ». Certaines bactéries se sont aussi dotées d'appendices cellulaires. Les flagelles leur permettent de se déplacer dans le milieu, alors que les pili communs ou fimbriae (par opposition aux pili sexuels) leur donnent la capacité de sonder la surface de la cellule hôte.

Les flagelles sont des appendices protéiques filamenteux, de 6 à 15 µm de long sur 12 à 30 nanomètres d'épaisseur. Ils sont constitués de trois parties. Un filament hélicoïdal, dont la longueur peut atteindre 10 µm, formé par l'assemblage de monomères de flagelline. Un crochet, lui aussi de nature protéique, très court, incurvé et flexible. Et enfin, le corpuscule ou corps basal, correspond au lieu d'insertion du flagelle dans la cellule. Cette dernière structure est le moteur du flagelle et sa rotation est assurée par un transport de proton. Ces organes jouent un rôle important dans la diffusion des bactéries d'un site infectieux local à d'autres organes (Drake and Montie 1988). En plus de leur rôle dans la motilité, ils sont impliqués dans l'invasion des cellules épithéliales (Moens and Vanderleyden 1996), la formation de biofilm (O'Toole and Kolter 1998) et l'induction de réponses inflammatoires chez l'hôte (Eaves-Pyles, *et al.* 2001, Urban, *et al.* 2004), contribuant ainsi à la virulence des bactéries chez la souris (Chua, *et al.* 2003).

A l'exception de *B. mallei*, toutes les bactéries du complexe BCC sont des organismes mobiles pourvus d'un ou plusieurs flagelles polaires. *B. cenocepacia* possède un seul long

flagelle polaire, responsable de sa motilité par un mouvement de type « swimming ». Les connaissances sur la structure et la fonction du flagelle chez les bactéries du complexe BCC sont encore limitées. Deux types de flagellines, codées par le gène *fliC,* ont été identifiés. Elles se distinguent par leur taille : 55 kDa pour le type I et 45 kDa pour le type II ; mais aussi par le polymorphisme de longueur des fragments de restriction (RFLP) du gène *fliC* (Hales*, et al.* 1998, Montie and Stover 1983)). Deux mutants non-mobiles de la souche J2315 de B. *cenocepacia*, obtenus par mutation du gène *fliG* codant pour un composant du corps basal du flagelle ou du gène *fliI* codant pour une ATPase intervenant dans la translocation de protéines flagellaires par le système de sécrétion de type III montrent une diminution significative de l'invasion des cellules épithéliales pulmonaires A549 (Tomich*, et al.* 2002). L'inactivation du gène *fliCII* codant pour la sous unité majeure de la flagelline chez la souche K56-2 de *B. cenocepacia*, entraine une absence de mortalité chez les souris infectées par le mutant (Urban*, et al.* 2004). Ces résultats indiquent que le flagelle de *B. cenocepacia* contribue à la virulence de la bactérie.

De nombreuses bactéries à Gram négatif et exceptionnellement des bactéries à Gram positif possèdent des appendices de surface plus courts et plus fins que les flagelles, les pili, qu'on appelle aussi fimbriae. Ce sont des spicules protéiques courts de 2 à 3 µm de long, constitués d'une seule protéine polymérisée : la piline. Ils peuvent être associés à d'autres polypeptides plus minoritaires dont certains jouent un rôle d'adhésines. Les bactéries utilisent ces protéines présentes sur les pili pour reconnaître, adhérer et infecter la cellule hôte. Environ 60% des bactéries du complexe étudiées expriment spontanément des fimbriae. Chez les différentes souches du BCC, cinq classes de pili ont été décrites selon leur morphologie en microscopie électronique (Goldstein*, et al.* 1995). La souche épidémique ET12 de *B. cenocepacia* est caractérisée par la présence de pili de classe II (cable pili) associés au gène *CblA* (Sun*, et al.* 1995). Les cable-pili sont de grands appendices péritriches de 2 à 4 nm capables de se fixer aux mucines respiratoires des patients atteints de mucoviscidose (Sajjan and Forstner 1992). Les souches de *B. cenocepacia* qui expriment le cable-pili associé à l'adhésine de 22 kDa (Sajjan and Forstner 1993) utilisent comme récepteur la cytokeratin 13 (CK13), une protéine des filaments intermédiaires fortement exprimée dans les cellules épithéliales CF (Sajjan*, et al.* 2000) (*Figure 1.10*). La présence d'adhésines reconnaissant des oligosaccharides des cellules de l'hôte n'a pas encore été identifiée chez les bactéries du complexe BCC.

1.3.2.6. Les lectines solubles

Certaines bactéries du complexe BCC possèdent dans leur génome quatre séquences homologues au gène *lecA* codant pour la lectine Pa-IIL chez *P. aeruginosa*. Ces lectines solubles font l'objet de ce travail de thèse et seront présentée plus en détail dans le chapitre 3.

Chapitre 2

Les lectines de bactéries pathogènes

Les bactéries, champignons et parasites sont des organismes hétérotrophes qui dépendent du saprophytisme, de la symbiose ou encore de la pathogénicité pour se procurer leur source d'énergie. De tels modes de vie nécessitent la reconnaissance spécifique de la matière organique ou des tissus de l'hôte. L'arsenal que ces microorganismes utilisent pour reconnaître, adhérer et infecter leur hôte comprend plusieurs types de protéines dont certaines reconnaissent les glycoconjugués de surface ; ce sont les lectines. Du latin *legere*, signifiant « selectionner », les lectines sont des protéines ou glycoprotéines ne faisant pas partie de l'immunité acquise, qui se lient de façon spécifique et réversible aux sucres, et ce, sans les modifier.

A quelques exceptions prés, ces protéines n'ont pas une très grande affinité pour leurs ligands monosaccharidiques, avec une constante d'affinité K_a de l'ordre de 10^3-10^4 M^{-1}. Cette affinité moyenne est généralement compensée par la multivalence de ces protéines. Ce sont des molécules ubiquitaires que l'on trouve à la fois chez les microorganismes (virus, bactéries et parasites) (*Figure 2.1*), chez les plantes mais aussi chez les insectes et les animaux.

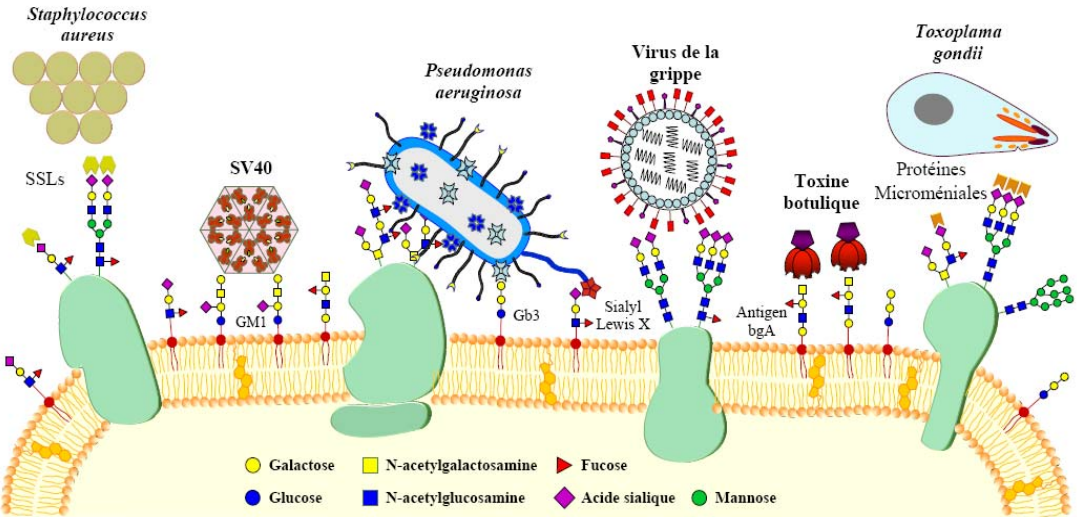


Figure 2.1. Représentation des stratégies utilisées par les pathogènes pour reconnaître et adhérer aux glycoconjugués de la cellule hôte (Imberty and Varrot 2008).

Comme le montre la *Figure 2*.2, les lectines bactériennes connues peuvent être classées dans trois familles: les lectines fimbriales, celles associées à des toxines et enfin les lectines solubles (Imberty *et al.* 2005)

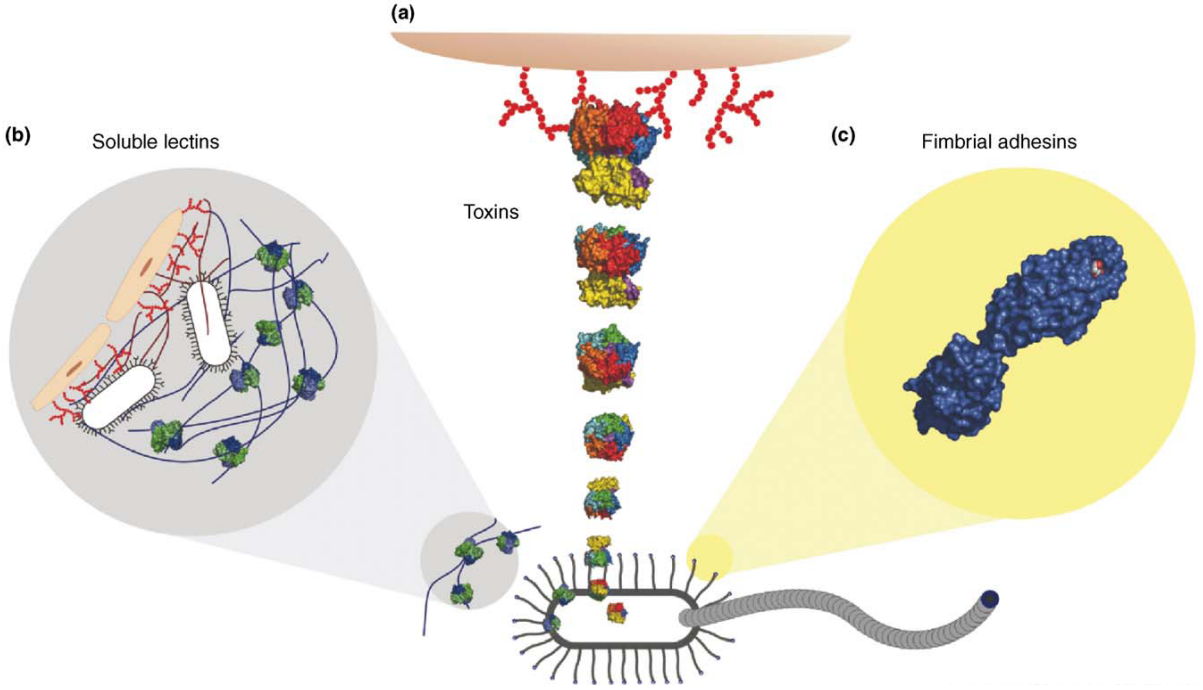


Figure 2.2. Dessin schématique d'une bactérie générique représentée avec un flagelle et plusieurs pili.

Les différentes les différentes protéines de liaison au sucre qui ont été caractérisées structurellement sont montrées avec des agrandissements. **(a)** la toxine AB-5 (code PDB 1LTS) reconnaissant les glycolipides de la surface cellulaire, **(b)** les lectines solubles (code PDB 1GZT) interagissant avec les filaments de mucine de la surface des cellules hôtes et **(c)** les adhésines fimbriales (code PDB 1KLF) (Imberty *et al.* 2005)

2.1. Les lectines associées à des toxines

De nombreuses bactéries pathogènes secrètent des toxines protéiques qui endommagent les cellules par rupture de la paroi cellulaire, inhibition de la synthèse protéique ou par activation des métabolismes secondaires. Des lectines sont parfois associées à ces toxines, exploitant ainsi l'adhésion aux glycoconjugués de la surface cellulaire de l'hôte pour entrer dans la cellule cible. Les toxines de la famille AB_5 (ADP-ribosylating toxin), sont formées d'une sous-unité A (ADP-ribosyltransférase toxique), responsable de l'activité enzymatique et

d'une sous unité pentamérique B possédant les sites de liaison au sucre (Merritt and Hol 1995). Les structures cristallographiques connues des toxines AB_5 regroupent la toxine cholérique de *Vibrio cholera* (Merritt *et al.* 1994), la toxine pertussique de *Bordetella pertussis* (Stein *et al.* 1994), ainsi que plusieurs toxines de *E. coli* enteropathogénique telles que la verotoxin ou Shiga-like toxin (Ling *et al.* 1998), la heat-labile enterotoxine (Sixma *et al.* 1992) et la subtilase cytotoxine (Byres *et al.* 2008). Ces protéines sont reconnaissent spécifiquement les gangliosides tels que le GM1, GM2 et le Gb3 présents sur la surface cellulaire des parois intestinales et pulmonaires.

Le domaine lectine de la toxine cholérique sécrétée par la bactérie *Vibrio cholerae* reconnaît spécifiquement les GM1 des cellules épithéliales humaines. La structure cristallographique des domaines lectines en complexe avec les pentasaccharide GM1 a été résolue (*Figure 2.3*)

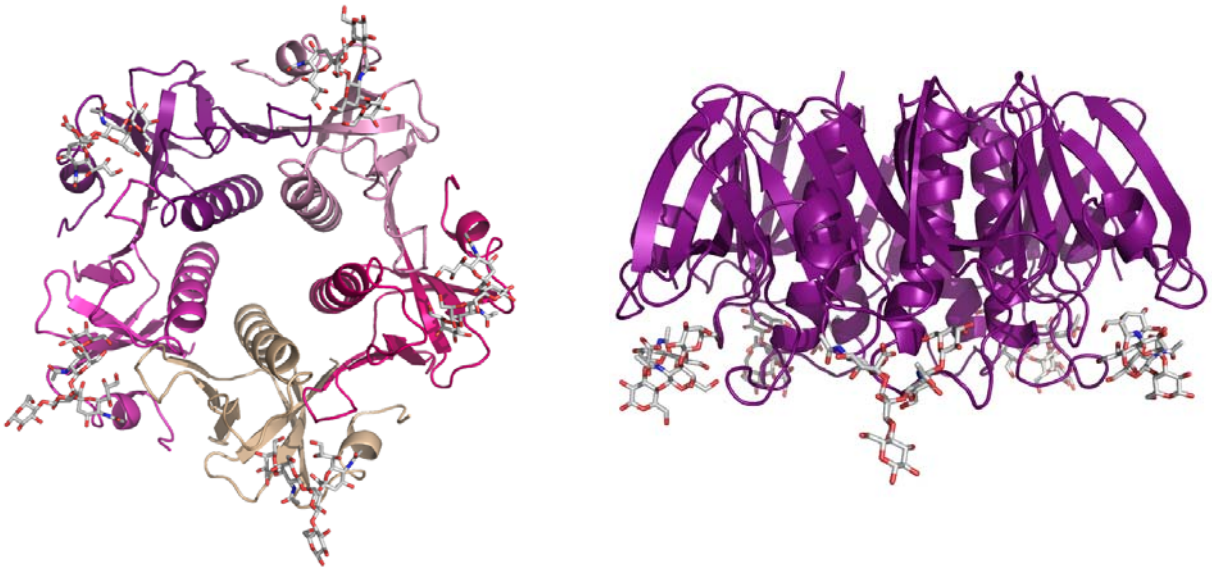


Figure 2.3. Représentation de la sous-unité B de la toxine du cholera en complexe avec le GM1.

A gauche, le pentamère vu de dessous. A droite une vue de profil. Dessin réalisé avec le logiciel Pymol à partir de la structure cristallographique du complexe (code PDB 3chb).

Les deux pathogènes *Clostridium tetani* et *Clostridium botulinum* sont responsables de maladies létales, le tétanos et le botulisme. La neurotoxine du botulisme est considérée comme la substance la plus dangereuse au monde, en particulier dans le cadre du bioterrorisme. Pour ces deux pathogènes, les neurotoxines possèdent des domaines lectines

capables de reconnaître les glycoconjugués à la surface des neurones (*Figure 2.4*). Les structures cristallographiques de ces domaines en complexes avec le pentasaccharide du glycolipide GT1b ont été résolues (Fotinou *et al.* 2001, Stenmark *et al.* 2008).

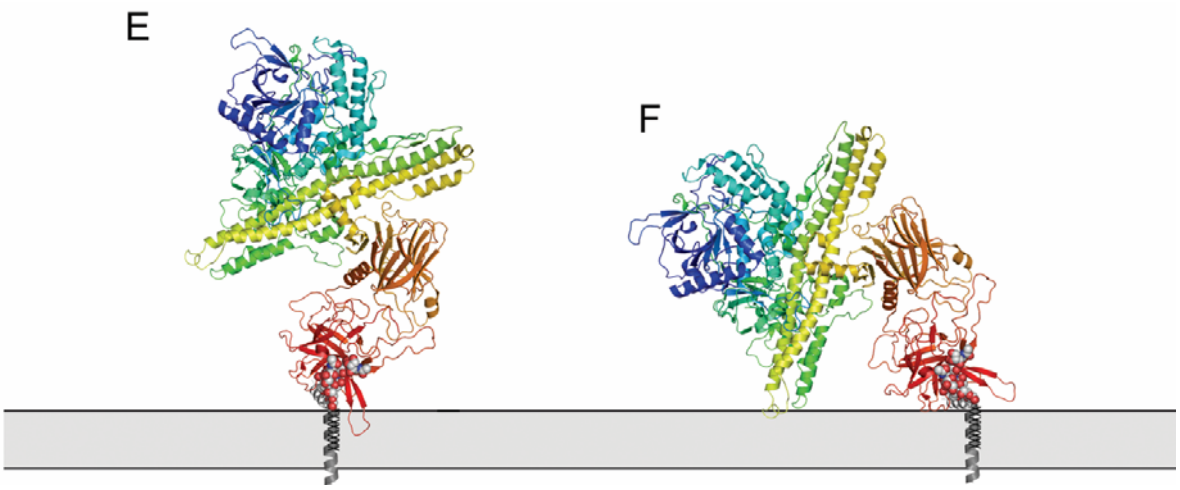


Figure 2.4. Modèle de l'interaction de la toxine de botulisme.

Schéma construit à partir de la structure du complexe BoNT/A avec GT1b et de la structure de BoNT/B. Ancrage de la lectine à la membrane par l'intermédiaire de la lectine et amarrage de la toxine (Stenmark *et al.* 2008)

Récemment, les structures de plusieurs autres domaines lectines associées à des toxines bactériennes ont été résolues. Des domaines répétés de la toxine de *Clostridium difficile,* une bactérie responsables de récentes infections nosocomiales très sérieuses, se lient au cellules épithéliales de l'intestin (Greco *et al.* 2006). Le staphylocoque doré présente également des domaines lectines sur les toxines SSL5 et SSL11 qui jouent un rôle dans l'adhésion aux cellules (Baker *et al.* 2007, Chung *et al.* 2007).

2.2. Les lectines de surface

Les lectines localisées au niveau des organelles de surface (pili et flagelles) sont qualifiées d'adhésines. Ces protéines jouent un rôle important dans la reconnaissance des glycannes de l'hôte et l'adhésion sur les surfaces de cellules cibles. Elles font partie d'une architecture multiprotéique complexe et sont de ce fait difficiles à exprimer sous forme soluble et à cristalliser. Seul un nombre limité de structures cristallographique de ces protéines ont été obtenues (De Greve *et al.* 2007).

Il existe trois types de pili : les types IV, les types I et les types P. Les plus répandues et les mieux caractérisées, que ce soit sur le plan structural, fonctionnel, ou de leur biosynthèse, sont les adhésines d'*E. coli*. Chez les souches uropathogénique et entéropathogénique d' *E. coli*, l'attachement à l'épithélium de l'appareil urinaire et des intestins respectivement se fait par l'intermédiaire de ces protéines. Les adhésines FimH de pili de type 1 et PapG, de pili de type P de la souche uropathogénique d'*E. coli*, ont été cristallisés respectivement en complexe avec le mannose et le galabiose (Galα1-4Gal) (Dodson *et al.* 2001, Hung *et al.* 2002), alors que la structure de l'adhésine GafD des pili de type F17 d'*E. coli* enterpathogénique a été résolue en complexe avec le N-acetylglucosamine (Buts *et al.* 2003, Merckel *et al.* 2003). Il est intéressant de noter que ces domaines adhésines présentent le même type de repliement, bien qu'il n'y ait pas d'identité de séquence significative et que les sites de liaison au sucre et donc la spécificité soient différents (Figure 2.5). Les lectines de cette famille sur une structure allongée avec un repliement de type β-sandwich.

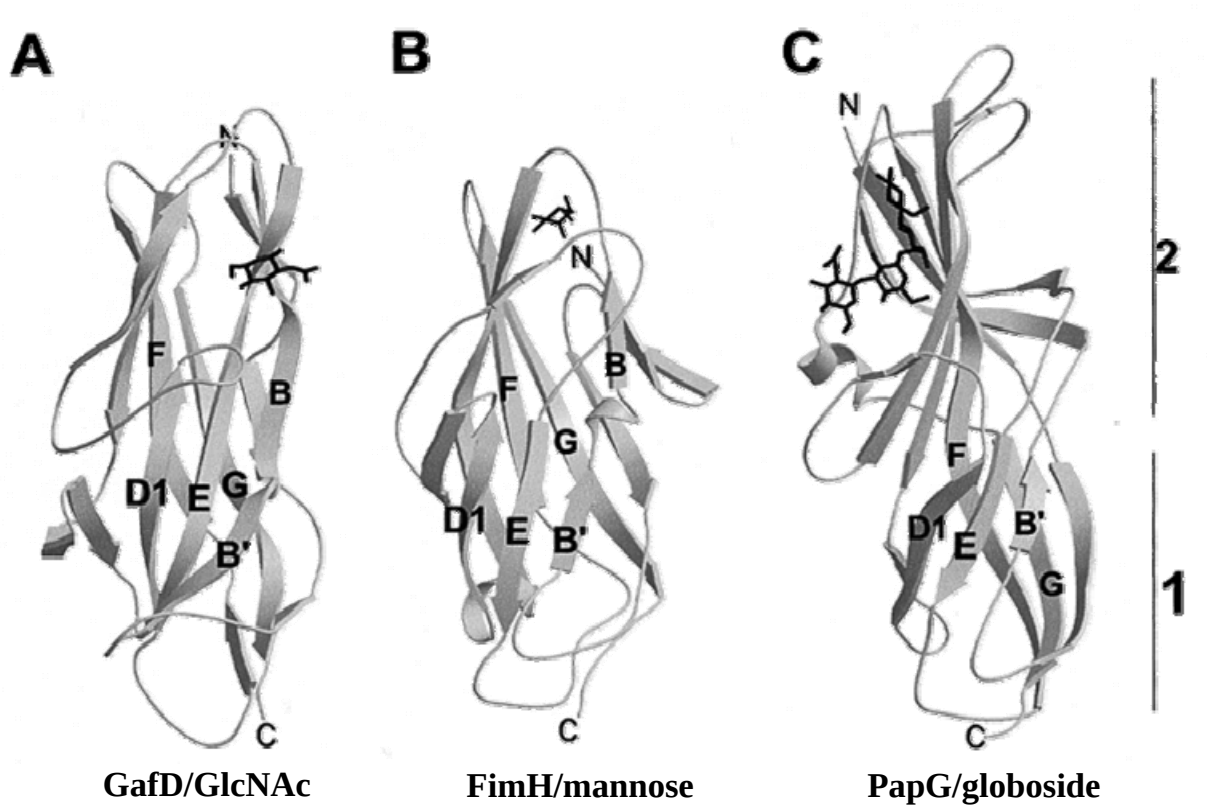


Figure 2.5. Structures cristallographiques d'adhésines fimbriales

2.3. Les lectines solubles

Seul un petit nombre de lectines bactériennes solubles ont été identifiées (*Tableau 2.1*). Les structures tridimensionnelles de ces protéines en complexe avec leur ligand ont été étudiées par cristallographie aux rayons X et trois familles différentes ont été identifiées. Ces lectines sont présentes en grande quantité dans le cytoplasme des bactéries et leur rôle n'est pas encore clairement élucidé.

Organisme	Nom	Classification	Spécificité
Pseudomonas aeruginosa	PA-IL	β-sandwich à 1 calcium	Galactose
	PA-IIL	β-sandwich à 2 calcium	L-Fucose > mannose
Ralstonia solanacearum	RSL	β-propeller	L-Fucose
	RS-IIL	β-sandwich à 2 calcium	mannose > L-Fucose
Chromobacterium violaceum	CV-IIL	β-sandwich à 2 calcium	L-Fucose > mannose

Tableau 2.1. Lectines bactériennes solubles

2.3.1. Lectine en β-propeller de Ralstonia solanacearum.

Une lectine soluble, spécifique du fucose, a été isolée de la bactérie phytopathogène *Ralstonia solanacearum* (Sudakevitz *et al.* 2002). Le trimère de RSL forme un β-propeller à 6 lames, chaque monomère consistant en 2 lames (Kostlanová *et al.* 2005). La lectine montre une forte affinité pour le disaccharide αFuc1-2Gal, qui est présent dans les groupes sanguins humain mais également dans les xyloglucanes, un polysaccharide présent dans les parois primaires des végétaux. La structure cristallographique de la lectine complexée avec un fragment de xyloglucane a été décrite pour et trois résidus sont bien visibles dans le site de reconnaissance (*Figure 2.6*)

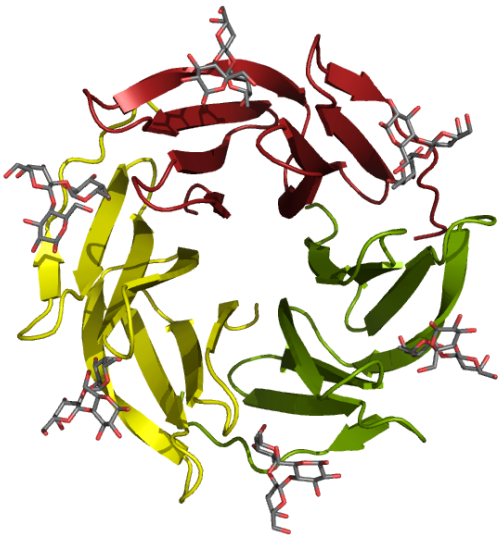


Figure 2.6 Trimère de RSL avec un trisaccharide de xyloglucane dans chacun des six sites de reconnaissances du trimère.

2.3.2 β-sandwich à 1 calcium

Parmi les facteurs de virulence sécrétés par *P. aeruginosa*, la lectine soluble PA-IL (gène *LecA*) a été découverte en 1972 par des expériences d'hémagglutination sur des érythrocytes traités à la papaïne en présence d'extraits de *P. aeruginosa.* La protéine a ensuite pu être isolée par chromatographie d'affinité sur une colonne de sépharose à partir d'une culture de *P. aeruginosa* (Gilboa-Garber 1982). Cette lectine est une protéine de 51kDa composée de quatre sous-unités identiques de 121 acides aminés, liant le galactose et ses dérivés.

PA-IL ne présente pas de similarité de séquences avec des protéines connues. La structure consiste en un tétramère formé de quatre monomères de type β-sandwich (*Figure 2.7*), chacun présentant un site de liaison contenant un ion calcium et un résidu galactose (Cioci *et al.* 2003).

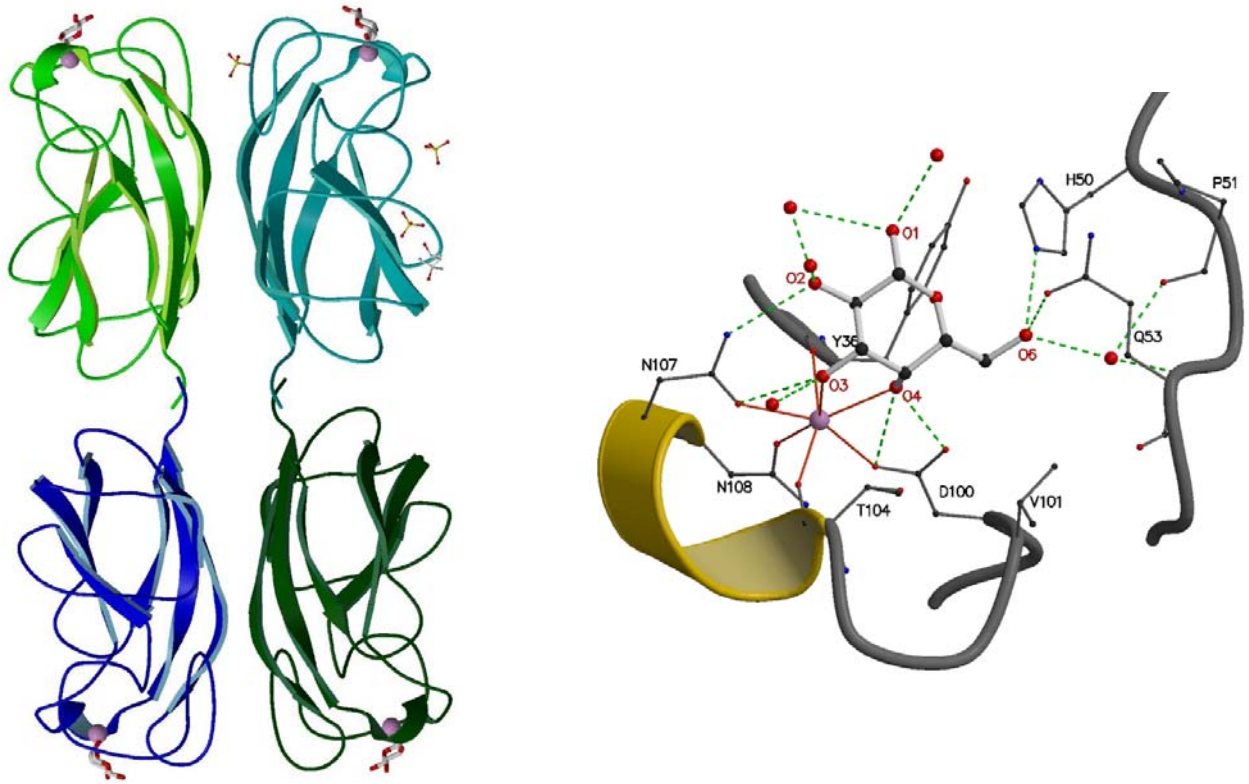


Figure 2.7. Tétramère de PA-IL et site de reconnaissance du complexe avec le galactose

Plus récemment une forte affinité de 68 et 77 µM a été mesurée pour l'interaction de PA-IL avec le Galα1-3galβ1-4Glc et le Galα1-4galβ1-4Glc respectivement. La structure

cristallographique de la protéine en complexe avec le Galα1-3galβ1-4Glc a été résolue (Blanchard *et al.* 2008).

Des études ont montré que l'expression de cette lectine est sous le contrôle du quorum sensing (Winzer *et al.* 2000). Son rôle exact dans la virulence de la bactérie n'est pas claire. Elle pourrait être impliquée dans la formation et la stabilisation du biofilm. En effet, des analyses statistiques d'épaisseur de biofilm sur un mutant sauvage et délété du gène codant pour la protéine ont montré que la délétion du gène *lecA* engendre une diminution de la quantité de bifilm formé. De même, la surexpression de la lectine PA-IL chez la bactérie conduit à une augmentation de la quantité de biofilm (Diggle *et al.* 2006).

2.3.3 β-sandwich à 2 calcium

La lectine PA-IIL a également été purifiée à partir de *P. aeruginosa* par le groupe de N. Gilboa-Garber (Gilboa-Garber *et al.* 2000). Cette protéine (gène l*ecB*) d'une taille de 11.7 kDa, est un homotétramère formé de quatre sous unités de 114 acides aminés. PA-IIL lie le fucose avec une très forte affinité (de l'ordre de 1.5 10^5 M^{-1}) mais reconnaît aussi le mannose avec une plus faible affinité. La structure cristallographique du complexe PA-IIL/fucose a été obtenue à haute résolution (Mitchell *et al.* 2002). Le mode de reconnaissance du sucre est nouveau et implique la participation de deux ions calcium (*Figure 2.8*).

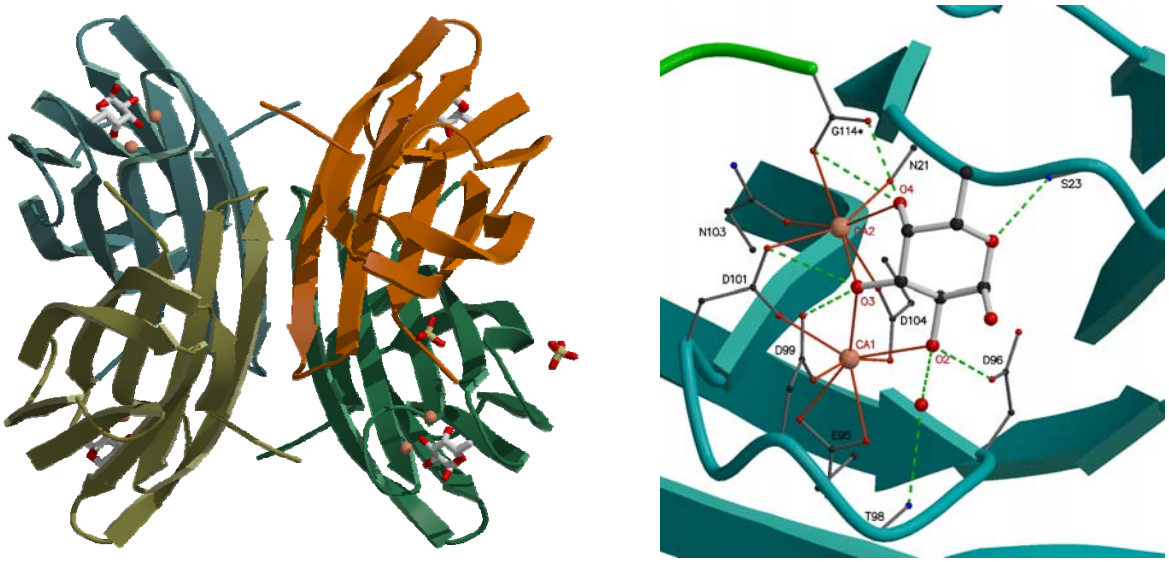


Figure 2.8. Tétramère de PA-IIL et site de reconnaissance du complexe avec le fucose

Une séquence présentant des similarités avec PA-IIL a été identifiée dans d'autres bactéries : *Ralstonia solanacearum* (phytopathogène) et *Chromobacterium violaceum* (intéressante en biotechnologie mais pouvant être dangereuse pour l'homme). Ces lectines ont été clonées et caractérisées et montrent des spécificités différentes pour les monosaccharides (Pokorná *et al.* 2006, Sudakevitz *et al.* 2004). Les structures des complexes RS-IIL/mannose et CV-IIL/fucose, CV-IIL/mannose ont été résolues. Les différences d'affinité ont pu être corrélées à une différence de trois acides aminés dans le site de reconnaissance (Adam *et al.* 2007). Les bases de cette spécificité seront détaillées dans le chapitre 3.

Le rôle des lectines de la famille de PA-IIL dans les processus infectieux, est encore un sujet de discussion. Dans *P. aeruginosa*, la lectine est produite en combinaison avec les facteurs de virulence de la bactérie sous le contrôle du quorum sensing (Schuster *et al.* 2003, Wagner *et al.* 2003, Winzer *et al.* 2000). Une étude *in vivo* a montré que la lectine PA-IIL est capable d'inhiber le battement cilliaire des cellules pulmonaires (Bajolet-Laudinat *et al.* 1994). Une étude récente a suggéré que la lectine soluble, PA-IIL est abondamment présente dans la membrane externe de la bactérie (Tielker *et al.* 2005). Dans ces travaux, il a été montré qu'une souche mutante de *P. aeruginosa*, délétée du gène *LecB*, est déficiente dans la formation du biofilm en comparaison avec la bactérie sauvage. Ces résultats suggèrent que la lectine est attachée aux sucres de la surface des bactéries. Finalement, la lectine pourrait être impliquée dans la reconnaissance et l'adhésion aux cellules hôtes, dans les interactions entre les bactéries ou encore dans des interactions avec les exopolysaccharides du biofilm.

Chapitre 3

Identification des lectines solubles de *B. cenocepacia* et objectifs de ce travail de thèse

---

Le travail décrit dans cette thèse constitue une extension des résultats obtenus dans le laboratoire pour les lectines solubles de *P. aeruginosa*. La lectine PA-IIL présentant une cible éventuelle pour le développement de composés anti-adhésion, il était intéressant de connaître l'extension de cette famille de protéine dans d'autres bactéries.

3.1. Recherche et alignement des séquences homologues à PA-IIL

3.1.1. Recherche des séquences homologues à PA-IIL par BLAST

Les recherches de séquences homologues à PA-IIL ont été effectuées dans la base de donnée de génomes microbiens disponibles sur le site du NIH (http://www.ncbi.nlm.nih.gov/sutils/genom_table.cgi) qui contient actuellement les génomes de 785 bactéries, 51 archaébactéries et 187 eucaryotes, mais aussi sur le site « Sanger *B. cenocepacia* Blast server » (http://www.sanger.ac.uk/cgi-bin/blast/submitblast/b_cenocepacia) pour la souche J2315 de *B. cenocepacia*. Plusieurs modes de recherche de séquence sont accessibles sur ces sites. L'outil BLAST (Basic Local Alignment Search Tool) (Altschul, *et al.* 1997) désigne un algorithme qui permet de rechercher des alignements locaux, appelés HSP (Hight-scoring Segment Pair). Des filtres, sont conçus pour éliminer les régions répétitives et les segments de "faible complexité" qui conduisent à des résultats statistiquement significatifs, mais sans intérêt biologique. Un score global, résultant de la somme de tous les bons appariements pénalisé par les mésappariements (insertions, délétions et mutations introduites) est ensuite attribué à chaque alignement. Deux types de recherche ont été effectués. Dans les BLASTP, la

séquence protéique de PA-IIL a été comparée avec les séquences protéiques disponibles. Les TBLASTN ont permis de comparer cette même séquence protéique avec les traductions possibles des séquences nucléotidiques qui ne sont pas encore annotées.

3.1.2. Alignement des séquences

Les séquences protéiques sont alignées grâce au programme ClustalX. Les alignements finaux sont ensuite obtenus avec le programme ESPript (http://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi). Les pourcentages d'identité et d'homologie entre les séquences protéiques sont déterminés grâce au programme Genedoc.

3.2. Description des protéines "PA-IIL-like"

Seul un nombre restreint d'organismes possèdent un gène codant pour une protéine de type PA-IIL : "PA-IIL like" ont été identifiés. Plusieurs espèces de *Burkholderia* contiennent des gènes codant pour un domaine PA-IIL-like comme décrit dans le paragraphe 3.3. Cependant, ce domaine a été identifié chez trois autres microorganismes.

Dans la classe des *Gammaproteobacteria* qui contient le genre *Pseudomonas*, des similarités n'ont été observées que dans le génome de *Photorabdus luminescens*, une bactérie symbiotique de nématodes qui infecte les insectes (Duchaud, *et al.* 2003). Cet organisme possède dans son génome trois gènes similaires à celui codant pour la lectine PA-IIL : Q9F4V4, Q9F4V3, Q937N3. L'alignement des séquences protéique de la partie *C*-terminale de ces protéines avec la lectine PA-IIL est représenté dans la *Figure 3.1*. Il faut noter que le domaine PA-IIL-like est situé au milieu de la protéine qui présente des extensions en *N*-terminale et *C*-terminal. La présence d'une glycine *C*-terminale étant nécessaire à la liaison au sucre, il est peu probable que le site 2 Ca^{++}/sucre, caractéristique de PA-IIL soit conservé.

Le gène est également présent dans le génome de plusieurs bactéries appartenant à la classe des *Betaproteobacteria*. La bactérie *Chromobacterium violaceum* (famille des *Neisseriaceae*) est présente en abondance dans les sols humides des régions tropicales et sub-tropicales. Cette bactérie présente des intérêts en biotechnologie en raison de son métabolisme produisant une variété d'antibiotiques et d'antifongiques (Duran and Menck 2001)Cependant,

c'est un pathogène opportuniste qui est responsable de diarrhées, mais aussi de septicémies pouvant être mortelles chez les enfants et les personnes immunodéprimés. Elle est maintenant considérée comme un agent pathogène émergent en Australie ou en Floride (Brito, *et al.* 2004). La protéine CV-IIL, produite par *C. violaceum* est une lectine de 114 acides aminés présentant 69% d'homologie et 91% d'identité avec PA-IIL (*Figure 3.1*)

La bactérie *Ralstonia solanacearum* (famille des *Burkholderiaceae*) est un redoutable agent pathogène des plantes, responsable de dommages importants dans plus de 50 famille différentes de plantes dont la pomme de terre et l'aubergine (Salanoubat, *et al.* 2002). Elle peut occasionnellement infecter des personnes immunodéprimées. La lectine RS-IIL de *R. solanacearum* est une protéine de 114 résidus présentant 79% d'homologie et 97% d'identité avec Pa-IIL. (*Figure 3.1)*.

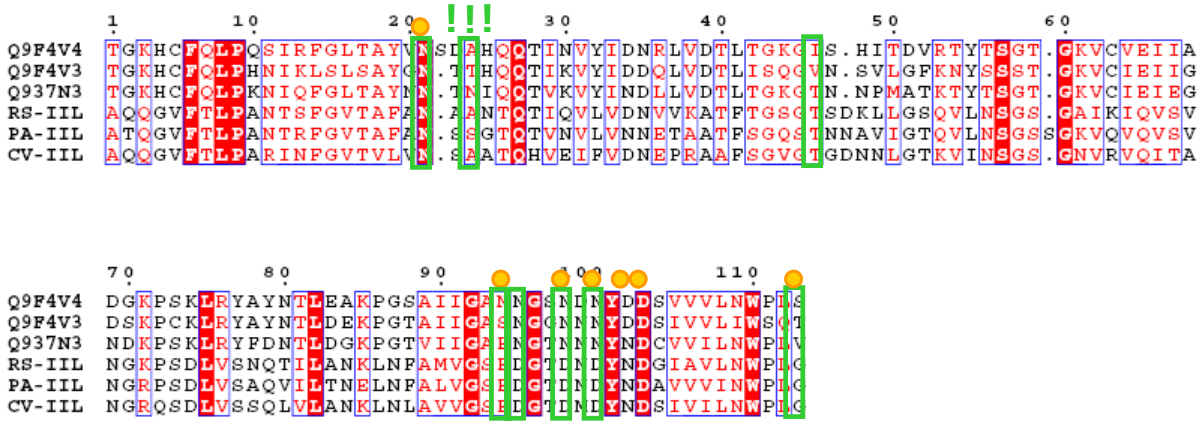


Figure 3.1. Alignement des séquences peptidiques homologues Pa-IIL.

Sont représentée, les séquences des domaines *C*-terminaux des protéines Q9F4V4, Q9F4V3 et Q937N3 de *P. luminescens*, RS-IIL, lectine de *R. solanacearum GMI1000 ;* PA-IIL, lectine de *P. aeruginosa* PAO1 et enfin, CV-IIL, lectine de *C. violaceum* ATCC12472. Sont représentés en blanc sur fond rouge, les acides aminés conservés dans les six séquences ; en rouge sur fond blanc ceux qui diffèrent mais ont des propriétés similaires. Les résidus impliqués dans la liaison du calcium sont signalés par une sphère jaune (●), ceux impliqués dans la liaison au sucre sont encadrés en vert (□). Les acides aminés responsables de la spécificité du sucre (boucle de spécificité) sont indiqués par des points d'exclamation (!!!). Les alignements sont obtenus avec les programmes ClustalX, et Esprit.

Ces deux lectines RS-IIL CV-IIL ont été étudiée dans notre laboratoire. Alors que PA-IIL reconnaît préférentiellement le L-Fucose, la lectine RS-IIL possède une affinité élevée pour le D-mannose (Sudakevitz, *et al.* 2004). La lectine CV-IIL présente quand à elle un comportement intermédiaire puisqu'elle reconnaît le L-Fucose et le D-mannose indifféremment (Pokorná, *et al.* 2006). L'alignement des séquences protéiques montre que les

acides aminés impliqués dans la reconnaissance au calcium sont conservés. En revanche, les résidus impliqués dans la reconnaissance du ligand saccharidique sont différents. Tout comme Pa-IIL, les deux lectines sont organisées en tétramère et reconnaissent leur ligand saccharidique par l'intermédiaire de deux ions calcium.

L'étude structurale comparative de leur site de liaison a montré que la différence de spécificité entre ces protéines est principalement due à la nature des acides aminés 22 à 24 qui constituent ce qu'on appelle la "boucle de spécificité" (*Figure 3.1*).

La séquence Ser22-Ser23-Gly24 chez PA-IIL est remplacée par les résidus Ala22-Ala23-Asn24 chez RS-IIL et Ser22-Ala23-Ala24 chez CV-IIL. La comparaison des structures cristallines de PA-IIL (Loris, *et al.* 2003) et de RS-IIL (Sudakevitz, *et al.* 2004) en complexe avec le D-mannose a permis de mettre en évidence l'importance du rôle que joue l'acide aminé 22 dans la spécificité. Le groupement hydroxyle de Ser22 chez PA-IIL établit une liaison hydrogène avec l'Asp96. L'oxygène O6 du D-mannose n'a pas la place nécessaire pour se positionner au fond du site et pointe par vers l'extérieur du site. Dans le cas de la lectine RS-IIL, l'Asp96 est disponible pour effectuer une liaison d'hydrogène avec l'oxygène O6 du D-mannose. Cette interaction supplémentaire a pour conséquence d'augmenter l'affinité pour le D-mannose. Finalement, les acides aminés Ser22, Ser23 chez Pa-IIL et Ala22, Ala23 chez RS-IIL sont responsable de la spécificité de ces lectine pour le L-fucose et le D-mannose respectivement. Enfin les acides aminésSer22, Ala23 de CV-IIL intermédiaires en terme de similarité de séquence expliquent le comportement intermédiaire de la lectine (Pokorná, *et al.* 2006).

Les structures des complexes RS-IIL/mannose et CV-IIL/fucose, CV-IIL/mannose ont été résolues. Les différences d'affinité ont pu être corrélées à une différence de trois acides aminés dans le site de reconnaissance (Adam 2007) (*Figure 3.2*).

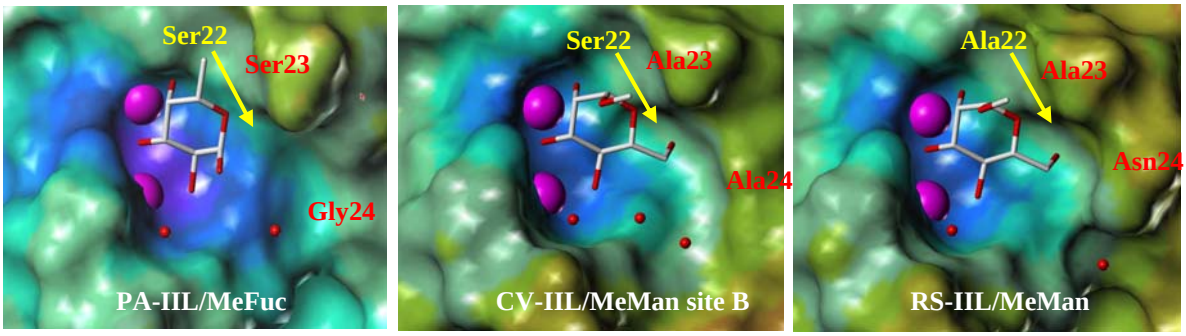


Figure 3.2. Comparaison des sites de reconnaissances de trois lectines PA-IIL, CV-IIL et RS-IIL

3.3. Lectines PA-IIL-like de B. cenocepacia

Certaines bactéries du complexe *B. cepacia* possèdent dans leur génome trois gènes codant pour des protéines ayant un domaine homologue à PA-IIL. Nous les avons apellées BclA, BclB et BclC (Bcl pour *Burkholderia cepacia* lectine) par ordre croissant de taille (*Figure 3.3*). Un quatrième gène, moins fréquent, est appelé BclD. Trois des protéines codées par ces gènes possèderaient un domaine additionnel *N*-terminal de taille non négligeable (de 116 à 168 résidus) et de fonction inconnue (*Figure 3.3.*).

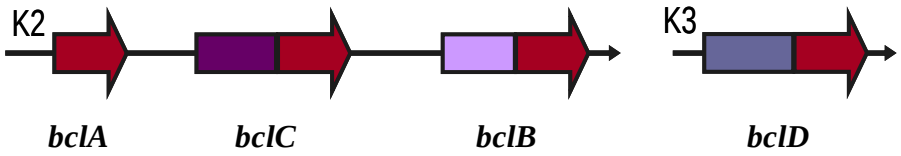


Figure 3.3. Représentation schématique de l'organisation des gènes contenant un domaine codant pour PA-IIL dans le génome de *B. cenocepacia*.

Chaque gène est représenté par une flèche. Les lectines «*B. cepacia* lectin» A, B, C et D sont représentées par des flèches. K2 et K3 correspondent aux chromosomes deux et trois de *B. cenocepacia*. Les régions similaires à PA-IIL sont représentées en bordeaux.

Toutes les souches de *B. cenocepacia* dont le génome séquencé est complet possèdent les gènes codant pour les lectines BclA, BclB et BclC. Les séquences sont également présentes dans le génome de deux *B. ambifaria* et une *B. cepacia*. Les lectines solubles n'ont à l'heure actuelle pas été trouvées dans le génome de la bactérie pathogène *B. multivorans*. Les séquences identifiées dans les différents génomes sont regroupées dans le *tableau3.1*.

Organisme	Taille génome (Mb)	Code d'accession aux protéines dans NCBI				Centre de séquençage
		BclA	**BclB**	**BclC**	**BclD**	
B. cenocepacia AU 1054	7.28	ABF79857.1	ABF79859.1	ABF79858.1		DOE JGI
B. cenocepacia HI2424	7.70	ABK09924.1	ABK09922.1	ABK09923.1		DOE JGI
B. cenocepacia J2315	8.06	BCAM0186	BCAM0184	BCAM0185	C-terminal BCAS0293	Sanger Institute
B. cenocepacia MC0-3	7.97	ACA93447.1	ACA93449.1	ACA93448.1		DOE JGI
B. sp. 383 (B. cepacia R18194)	8.68	ABB13044.1	ABB13046.1	ABB13045.1	ABB06257.1	DOE JGI
B. ambifaria AMMD	7.53		ABI90657.1			DOE JGI
B. ambifaria MC40-6	7.64	ACB65731.1	ACB65730.1	ABI90657.1		DOE JGI

Tableau 3.1. Séquences homologues au gène codant pour Pa-IIL dans le génome des bactéries du complexe BCC.

Les gènes codant pour BclA, BclB et BclC sont sur le chromosme K2. Le gène codant pour BclD est sur le chromosome K3 mais n'est présent que dans le génome *B. cenocepacia* J2315 et de *B. cepacia* R18194. De plus, dans le génome *B. cenocepacia* J2315, auquel appartient le clone épidémique et virulent ET12, la lectine BclD ne peut pas être traduite en raison d'un décalage du cadre de lecture. La comparaison des séquences de chaque lectine entre les différentes souches du Tableau 3.1, ne montre que très peu de différence (95 à 100% d'identité). Seule les lectines de la souche J2315 seront donc étudiées par la suite. De plus, la lectine BclD n'étant certainement pas exprimée, nous nous focaliserons sur les trois lectines BclA, BclB et BclC. L'alignement des séquences protéiques de ces trois protéines avec la lectine Pa-IIL est représenté dans la (*Figure 3.4*).

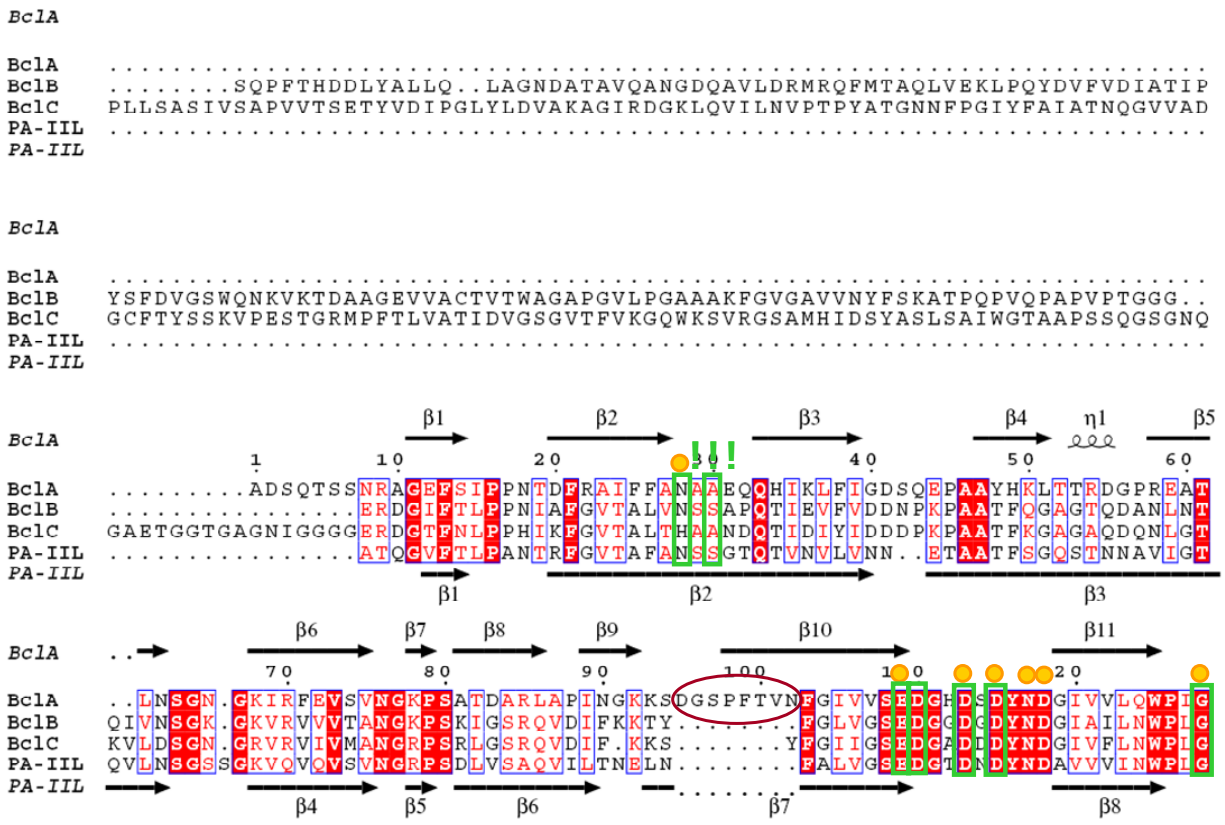


Figure 3.4. Alignement des séquences peptidiques homologues Pa-IIL.

Sont représentée, les séquences de BclA, BclB et BclC, lectines de *B. cenocepacia J2315* et enfin PA-IIL, lectine de *P. aeruginosa* PAO1. Au dessus de la séquence peptique de BclA, sa structure secondaire organisée en onze brins beta et une helice alpha ; déterminée grâce à la résolution de sa structure cristallographique (2NV) au cours de cette étude. En dessous de la séquence de Pa-IIL, sa structure secondaire organisée en huit brins beta (1GZT). Sont représentés en blanc sur fond rouge, les acides aminés conservés dans les quatre séquences ; en rouge sur fond blanc ceux qui diffèrent mais ont des propriétés similaires. Les résidus impliqués dans la liaison du calcium sont signalés par une sphère jaune (●), ceux impliqués dans la liaison au sucre sont encadrés en vert (□). Les acides aminés responsables de la spécificité du sucre (boucle de spécificité) sont indiqués par des points d'exclamation (!!!). Enfin une boucle supplémentaire présente chez BclA est représentée par une ellipse bordeaux (⬭). Les alignements sont obtenus avec les programmes ClustalX, et Espript.

La lectine BclA présente 65% d'homologie et 41 % d'identité avec Pa-IIL. Le domaine *C*-terminal de BclB possède quand à lui un pourcentage d'homologie de 78% et d'identité de 59%. De même, la partie lectine de BclC est homologue à 78% et identique à 51% à Pa-IIL, (*Figure 3.3*). Les acides aminés impliqués dans la liaison au sucre sont presque totalement conservés. Seul l'asparagine 21 de Pa-IIL est remplacée par une histidine chez BclC. L'analyse de la séquence protéique de ces lectines nous permet donc de dire qu'il s'agit très certainement de protéines liant le calcium. De plus, les acides aminés impliqués dans la liaison au sucre sont eux aussi très conservés avec toutefois des différence au niveau de la boucle de spécificité. La séquence des trois acides aminés Ala29-Ala30-Glu31 de BclA indique qu'elle a les caractéristiques d'une lectine à mannose. Les séquences Ser150-Ser150-Ala151pour BclB et Ala178-Ala179-Asn180 pour BclC laissent penser que ces protéines reconnaissent le fucose et le mannose respectivement. Les domaines *N*-terminaux n'ont pas d'identité de séquence entre eux et aucune similarité n'a pu être détectée avec d'autres séquences protéiques.

3.4. Objectif de ce travail de thèse

Les lectines solubles de *B. cenocepacia* n'ont été identifiées que récemment. En raison du pouvoir pathogène de cette bactérie et des enjeux médicaux importants, ce travail de thèse a été mis en place pour comprendre le fonctionnement et le rôle de ces protéines dans l'infection. En effet, ces lectines pourraient être directement ou indirectement impliquée dans l'adhésion de la bactérie à la surface des poumons ou encore dans la formation ou la stabilisation du biofilm, lors des infections survenant chez les patients atteints de mucoviscidose. En raison de la résistance naturelle de cette bactérie aux traitements antibiotiques, la mise en place de nouveaux traitements qui permettraient d'inhiber l'étape initiale d'adhésion de la bactérie à la surface des cellules est un enjeu important. Connaître la cible des lectines est la première étape vers la compréhension du rôle biologique de ces protéines. Associée à la résolution de la structure tridimensionnelle par cristallographie aux rayonsX ces informations pourraient permettre à long terme, la conception de composés thérapeutiques inhibant l'adhésion des bactéries aux cellules épithéliales pulmonaires et donc l'étape initiale de l'infection.

Les deux lectines les plus longues, BclC et BclD font actuellement l'objet des travaux de thèse de Ondrej Sulak et Lenka Malinovska de l'Université Masaryk à Brno en République Tchèque, avec qui nous collaborons.

Mes travaux de thèse portent quant à eux sur l'étude des deux autres lectines, BclA et BclB. Après avoir déterminé la spécificité des lectines recombinantes, notamment grâce à l'utilisation de puces à sucres. L'affinité de celles-ci pour leurs meilleurs ligands saccharidiques a été étudiée par la technique de microcalorimétrie isotherme de titration. En parallèle des essais de cristallogenèse ont été effectués dans le but de résoudre la structure cristallographique des protéines en complexe avec leur ligand. L'obtention de telles structures nous permet de mieux comprendre l'interaction à l'échelle atomique. Enfin, la localisation ainsi la fonction de ces lectines chez la bactérie *B. cenocepacia* ont pu être étudiées.au "Department of Microbiology and Immunology" de l'université de Western Ontario

Chapitre 4

Outils pour l'étude des interactions protéine - sucre

Afin d'élucider le rôle des lectines chez la bactérie *Bulkholderia cenocepacia* et de comprendre leur fonctionnement moléculaire, plusieurs techniques ont été utilisées de façon complémentaire. Dans un premier temps, leur spécificité est étudiée grâce à des tests d'interaction sur puces à sucre, ce sont les *Glycan-* et *Printed- arrays* ; mais aussi grâce aux tests d'*inhibition de l'hémagglutination* (HIA, Hemagglutination Inhibition Assay) et à la *résonnance plasmonique de surface* (SPR, Surface Plasmon Resonance). Puis la *microcalorimétrie par titrage isotherme* (ITC, Isothermal Titration Microcalorimetry) nous permet d'accéder directement à l'affinité ainsi qu'aux grandeurs thermodynamiques de l'interaction entre la protéine et ses meilleurs ligands. La structure tri-dimensionnelle des lectines en complexe avec des ligands saccharidiques est résolue par *cristallographie aux rayons X*. Elle fournit des informations sur la topologie de surface, les liaisons impliquées, et les conflits stériques, permettant ainsi de mieux comprendre la spécificité et la force des interactions mesurées. Enfin une nouvelle approche de quantification de biofilm a été testée.

4.1. Les tests d'hémagglutination

La mesure d'activité hémagglutinante est le test d'interaction le plus simple et le plus largement utilisé pour la détection des lectines et leur caractérisation (Goldstein *et al.* 1980, Rüdiger 1993). Cette technique repose sur la capacité des lectines à former un réseau avec des érythrocytes par interaction avec leurs glycoconjugués de surface. Les hématies de mammifères sont parmi les plus utilisées (homme, lapin, mouton, porc, etc…). Les tests sont réalisés sur des microplaques de titration et l'agglutination (ou agrégation) des globules rouges par les lectines, forme une phase gélatineuse visible à l'œil nu.

Dans un premier temps, la concentration minimale en lectine permettant une agglutination évidente est déterminée ; c'est ce qu'on appelle le *point d'équivalence* (*Figure 4.1*).

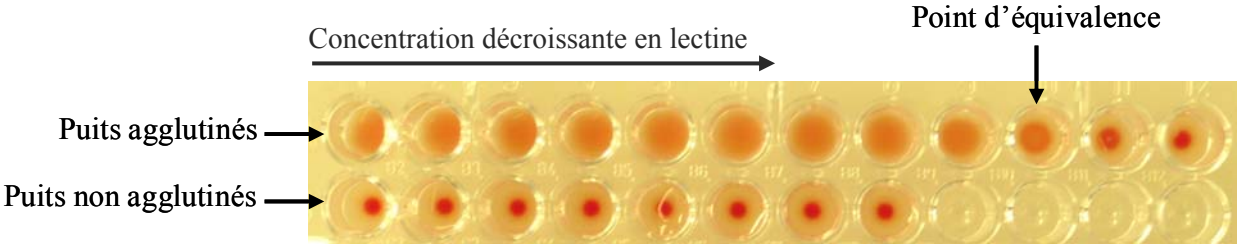


Figure 4.1. Le test d'hémagglutination sur microplaque de titration

Dans les puits agglutinés (en haut), la lectine forme un réseau avec les hématies, empêchant le dépôt de celles-ci au fond des puits (en bas). Le test d'une gamme de concentrations permet de déterminer le point d'équivalence, correspondant à la plus petite concentration en lectine permettant d'agglutiner les globules rouges.

Puis, des tests d'inhibition d'hémagglutination (HIA), permettent de cribler les ligands potentiels de la lectine. Lorsque l'hémagglutination peut être entièrement inhibée par le ligand testé, la concentration minimale en ligand empêchant la formation du réseau est mesurée ; c'est ce qu'on appelle le *potentiel d'inhibition* du ligand. Pour pouvoir comparer les potentiels d'inhibition de plusieurs ligands, il est préférable, de faire les différents tests sur un même lot de globules rouges. En effet, la valeur du potentiel d'inhibition peut dépendre de la source ainsi que de l'âge des érythrocytes. La mesure de l'inhibition d'hémagglutination est un test rapide et simple qui permet d'évaluer les ligands solubles de façon comparative, donnant un ordre approximatif d'avidité.

4.2. Les puces à sucres

Un des défis importants de la glycobiologie est le développement d'outils permettant l'analyse à grande échelle et à haut débit des interactions protéine-sucre. Différentes approches permettant de tester la spécificité des protéines pour leur ligand saccharidique ont été mises en place. C'est le cas des tests ELLA (Enzyme Linked Lectin Assay) qui permettent d'évaluer la faculté que possède une molécule donnée à inhiber l'interaction entre une protéine et une surface recouverte d'un ligand de référence. Mais la grande complexité des structures glucidiques, l'absence de larges bibliothèques de ligands saccharidiques lors des tests effectués ainsi que l'hétérogénéité des méthodes, ne permettaient pas une bonne évaluation de la sélectivité des protéines. Pour remédier à ce problème, le « Consortium for

Functional Glycomics » (CFG) aux Etats-Unis a mis au point de nouveaux outils de routine permettant de mieux comprendre le rôle des interactions protéine-sucre dans la communication entre les cellules (http://www.functionalglycomics.org). Ces outils permettant d'étudier les interactions entre les sucres et les protéines (lectines, enzymes, anticorps), les ARN, mais aussi les cellules entières ou encore des particules virales. L'accès aux biopuces du CFG et la mise à disposition pour la communauté scientifique des résultats générés, ont également permis de créer des bases de données spécialisées sur les « Glycan-Binding Proteins » (GBP), les structures glycanniques, mais aussi les glycosyltransferases.

La spécificité des deux lectines BclA et BclB a ainsi pu être étudiée par « Glycan array » et « Printed array », le terme « array » signifiant réseau ordonné en français. Ces deux types de tests se différencient principalement par la méthode de fixation des sucres. Dans les deux cas, la détection de l'interaction est permise par le marquage préalable des lectines avec un composé fluorescent comme l'Alexa Fluor® 488–TFP (carboxylic acid, 2,3,5,6 tetrafluorophenyl), ou l'utilisation d'anticorps spécifique (Blixt *et al.* 2004).

4.2.1. Les glycan arrays

Les premiers tests à avoir été proposés par le CFG sont basés sur le principe du test ELISA (Enzyme-linked immunosorbent assay). Une librairie de sucres biotinylés sont fixés à l'aide de bras espaceurs sur des plaques de 96 à 384 puits recouverts de streptavidine. (*Figure 3.2*). Cette version miniature du test ELLA permet de tester de nombreux ligands en utilisant peu de matériel biologique et de manière plus reproductible (Bochner *et al.* 2005).

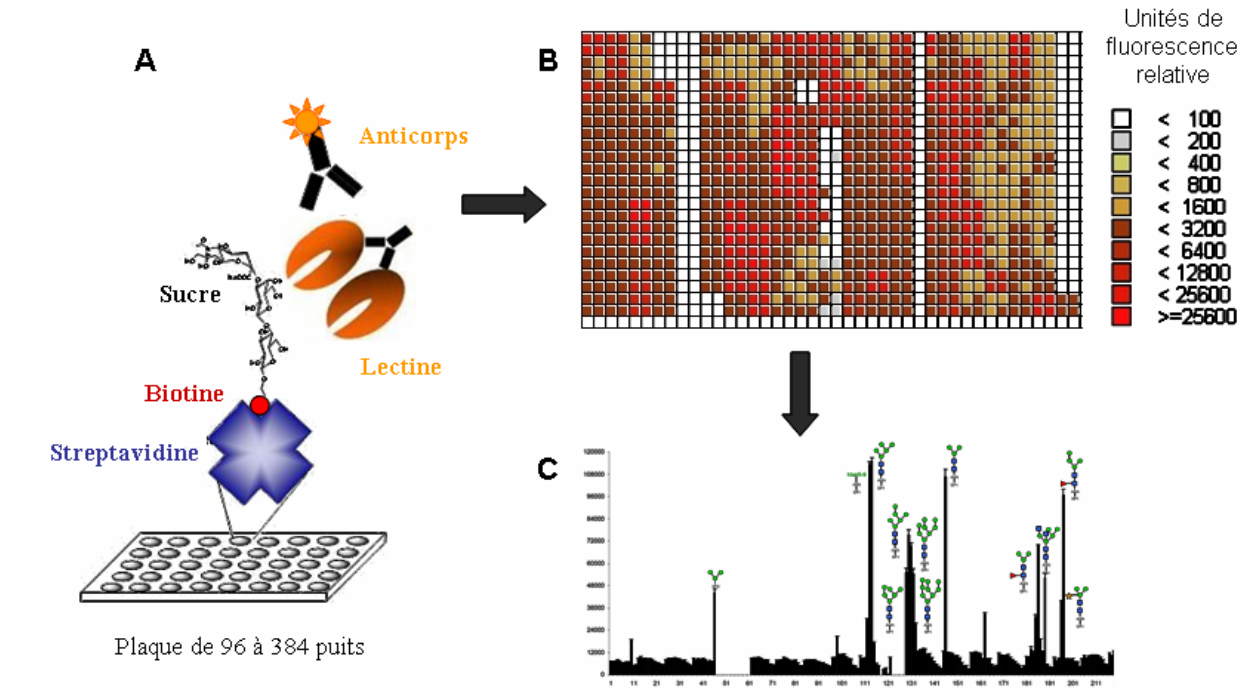


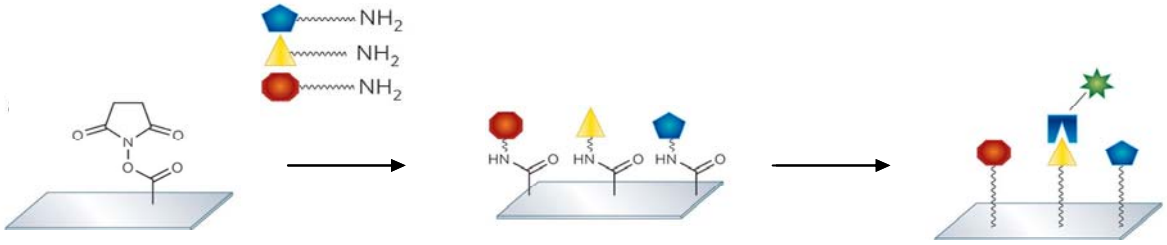
Figure 4.2. Détermination de la spécificité de liaison aux sucres des lectines par Glycan array

A. Principe du Glycan array. **B.** Données brutes. **C.** Représentation de la fluorescence relative mesurée en fonction des sucres testés.

4.2.2. Les printed arrays

Les puces à sucre les plus récentes proposées par le CFG et par d'autres laboratoires dans le monde reposent sur le principe du « Printed covalent glycan array » aussi appelé *printed array*. La fabrication de la puce est faite de manière robotisée grâce à la technologie standard des spotters. Les glycanes ou glycocongugés couplés à des fonctions amine, sont attachés de manière covalente sur des plaques de verre commerciales fonctionnalisées par de la N58 hydroxysuccinimide (NHS) (*Figure 4.3*).

L'intérêt de cette technique réside dans la miniaturisation du système, qui permet de travailler avec de quantités de matériel biologique infimes. En effet, les quantités de ligand saccharidique sont très faibles et la quantité de protéine nécessaire est 100 fois moins importante que pour les glycans arrays. De cette manière, une bibliothèque de plusieurs centaines de sucres naturels (structures glycaniques importantes des glycoprotéines et glycolipides de la membrane cellulaire) mais aussi synthétiques peut être testée.



Plaque de verre commerciale fonctionnalisée par la N-hydroxysuccinimide

Plaque de verre couverte de glucides naturels et synthétiques

Détection de l'interaction grâce à des Anticorps ou à des protéines marquées

Figure 4.3. Printed covalent glycan array

Aujourd'hui, le consortium propose trois types de puces en fonction de la bibliothèque de ligands testée. Le « Mammalian Glycan array » qui remplace les expériences faites sur plaques multi-puits et permet de tester 406 sucres naturels et synthétiques des mammifères. Le « Glycan Array Screening at Core H » permet de tester la spécificité de lectines, anticorps ou de GBPs putatives sur une librairie de glycanes d'origine humaine et pathogène. Et enfin, le « Pathogen Glycan Array », disponible depuis septembre 2008, est la première version de puce à sucres d'origine pathogène. Pour ce test, 96 oligosaccharides provenant de bactéries à gram négatif, comme *Proteus mirabilis* ou *P. aeruginosa,* sont disponibles.

4.3. Le microcalorimétrie isotherme de titration

Selon l'application que l'on en fait, il existe deux types de calorimètres. Le calorimètre différentiel à balayage (Differential Scanning Calorimetry, DSC) permet de mesurer l'énergie conformationnelle de macromolécules en balayant une plage de températures. Il est utilisé pour étudier la stabilité des protéines. L'ITC donne des informations sur la thermodynamique de l'interaction entre deux molécules (protéine-sucre, protéine-protéine, ADN-protéine ou autre) à température constante.

De toutes les méthodes utilisées pour étudier les interactions entre un récepteur et son ligand, la microcalorimétrie est la seule méthode permettant d'obtenir directement et en une seule expérience, les paramètres tels que la constante d'affinité (K_a), la stœchiométrie de l'interaction (n) ainsi que les contributions enthalpique (ΔH) et entropique ($T\Delta S$). De plus, les molécules étudiées sont libres en solution, contrairement à d'autres méthodes comme la résonnance plasmonique de surface, qui nécessite l'immobilisation du ligand ou du récepteur,

gênant ainsi parfois l'interprétation des résultats. L'ITC permet de mesurer une gamme d'affinité s'étalant de 10^3 à $10^8 M^{-1}$.

4.3.1. Notions de thermodynamique

La thermodynamique étudie la physique des échanges d'énergie (thermon: chaleur et dynamis: puissance) permettant ainsi l'étude des aspects énergétiques d'un phénomène biologique. Dans le cas des interactions, une interaction non covalente entre une macromolécule (M) et son ligand (L) conduit à la formation réversible d'un complexe (ML) et la force de l'interaction peut être décrite par la constante d'association, Ka en M^{-1}. Cette constante est définie comme le rapport des concentrations de la forme complexée (ML) sur les formes libres des différentes molécules (M et L). Plus cette constante est grande, plus l'affinité de la protéine pour son ligand est forte. L'affinité est généralement exprimée par la constante de dissociation Kd en M qui n'est autre que l'inverse de la constante d'association.

$$Ka = \frac{[ML]}{[M][L]} = \frac{1}{Kd} \qquad \textit{(Eq. 1)}$$

Comme toute interaction moléculaire, celle-ci donne lieu à des échanges d'énergie qui sont fonction de l'état initial (molécules M et L libres) et final (complexe ML) de la réaction, ainsi que du chemin réactionnel suivi entre les deux états. D'un point de vue thermodynamique, la variation d'énergie libre (ΔG) dépend de la constante d'association Ka, des concentrations initiales en macromolécule et en ligand ainsi que de la température. La variation d'énergie libre ΔG, qu'on appelle aussi énergie de Gibbs, est directement reliée à l'affinité de l'interaction par la relation suivante :

$$\Delta G = -RT \ln Ka \qquad \textit{(Eq. 2)}$$

Où - *T est la température absolue en Kelvins (°K)*
- *R est la constante des gaz parfaits = 8.315 $J.mol^{-1}.°K^{-1}$*
- *Ka est la constante d'affinité en M^{-1}*

La valeur absolue de cette variation d'énergie nous renseigne également sur l'évolution du système. Une réaction ou interaction est dite exergonique (par opposition à endergonique), c'est-à-dire qu'elle est spontanée et susceptible de fournir de l'énergie, lorsque la valeur de

ΔG, est négative. Cette énergie est en fait la résultante de deux contributions énergétiques : l'enthalpie et l'entropie, notées respectivement H et S, selon l'équation suivante.

$$\Delta G = \Delta H - T\Delta S \qquad \textit{(Eq.3)}$$

Où - *ΔG est la variation d'énergie libre ou énergie de Gibbs en (kJ.mol^{-1} ou kcal.mol^{-1})*
- *ΔH est la variation d'enthalpie en (kJ.mol^{-1} ou kcal.mol^{-1})*
- *ΔH est la variation d'entropie en (kJ.mol^{-1} ou kcal.mol^{-1})*

D'un point de vue très général, la variation d''enthalpie est une grandeur thermodynamique rendant compte de la quantité de chaleur (qu'elle soit sous forme thermique ou mécanique), reçue par un système qui évolue à pression constante. Tout système ayant tendance à avoir l'énergie la plus basse possible, la valeur de ΔH tend à être négative. L'entropie quand à elle, reflète le degré de désordre du système. Les mouvements Brownien sont responsables de l'augmentation de l'entropie, d'autant plus que la température augmente. Cette augmentation d'entropie est reflétée par une valeur de T*ΔS* qui tend à être positive.

Associée à des informations structurales, la détermination de ces contributions énergétiques permet une meilleure compréhension de l'interaction à l'échelle atomique. En effet, deux interactions ayant une affinité et une énergie libre identiques peuvent avoir des composantes enthalpique et entropique différentes.

4.3.2. Thermodynamique des interactions non covalentes

Dans le cadre d'interactions non covalentes, l'enthalpie et l'entropie sont impliquées dans des mécanismes différents.

Contribution entropique

L'interaction est favorable sur le plan entropique lorsque la variation ΔS est positive. Cette variation reflète les phénomènes d'hydrophobicité mis en jeu au cours de l'interaction. Deux phénomènes ayant des effets apposés d'un point de vue entropique entrent en jeu. Tout d'abord, au moment de la reconnaissance, les molécules de solvant se réarrangent autour de la protéine. Les molécules d'eau présentes dans les premières sphères de solvatation des deux partenaires de l'interaction sont chassées par le rapprochement des surfaces hydrophobes et sont libérées dans le milieu. Lorsque les surfaces impliquées dans la reconnaissance sont

hydrophobiques, l'entropie de désolvatation est favorable. Dans un deuxième temps, la diminution des degrés de liberté du ligand et/ou de la protéine au moment de l'interaction entrent aussi en jeu (*Figure 4.4*). Cette contribution entropique est souvent défavorable. En solution, le ligand peut adopter plusieurs conformations grâce aux rotations autour des angles de torsion. Au contraire, dans les complexes protéine-sucre, le ligand ainsi que les chaines latérales de certains acides aminés impliquées dans ces interactions sont généralement bloqués dans une seule conformation.

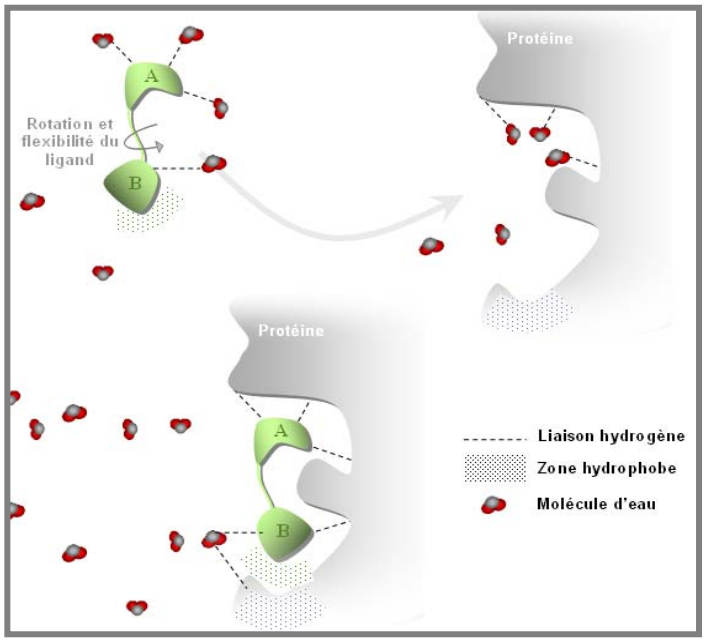


Figure 4.4. Réaménagement du solvant lors d'une interaction non covalente.

En absence de ligand les molécules d'eau établissent des liaisons hydrogènes avec les acides aminés du site de reconnaissance et avec le ligand. Les molécules présentes dans le site de liaison sont chassées lorsque la partie A du ligand se positionne dans le site d'interaction, occasionnant un gain entropique et enthalpique. La capture d'une molécule d'eau et la rigidité occasionnée lorsque la partie B du ligand entre dans le site, représentent une entropie défavorable.

Contribution enthalpique

La variation d'enthalpie reflète quand à elle la force des interactions mises en place entre la macromolécule et son ligand par rapport à celles que ces mêmes molécules entretiennent avec le solvant. Ces liaisons étant principalement des interactions de Van der Waals et surtout des liaisons hydrogènes. La variation d'enthalpie reflète finalement la redistribution du réseau de liaisons hydrogènes entre les espèces, y compris le solvant, au cours de l'interaction (*Figure 4.4*). L'interaction est favorable sur le plan enthalpique quand

ΔH est négatif. Une enthalpie d'interaction favorable nécessite la disposition correcte des groupes donneurs et accepteurs de liaison hydrogène dans le site de fixation. De plus, pour la protéine comme pour le ligand, des groupements polaires, dont l'enthalpie de désolvatation est défavorable, interviennent. Ces deux processus doivent être considérés pour connaître l'enthalpie globale de l'interaction.

4.3.3. Principe de l'ITC

Le microcalorimètre utilisé au cours de ces travaux est le VP-ITC de MicroCal®. Les principales composantes d'un appareil ITC sont illustrées dans la *Figure 4.5.A*. L'appareil est composé d'une cellule de référence, d'une cellule de mesure et d'une seringue d'injection. Les deux cellules sont isolées de l'environnement externe par une enceinte adiabatique, maintenue à température constante pendant le temps de l'expérience. La *cellule de référence* ne contient que du tampon et la *cellule de mesure* contient la protéine préparée dans le même tampon. Des aliquotes de ligand sont injectés à intervalles réguliers dans la cellule de mesure grâce à la seringue. La protéine, est donc progressivement saturée en ligand. Quand celui-ci se lie à la protéine, il se produit un échange thermique qui se traduit par une variation de chaleur dans la cellule de mesure : dégagement ou absorption dans les cas de réactions exothermique ou endothermiques respectivement. Après chaque injection, la température dans la cellule de mesure est ramenée à la même température que celle de la cellule de référence grâce à un signal électrique. La puissance de ce signal en µcal/sec est ainsi enregistrée pour chaque injection. Le thermogramme expérimental est ainsi enregistré au cours du temps. Au début de la titration, des quantités importantes de chaleur sont dégagées, indiquant la formation du complexe protéine-ligand. En fin de réaction, tous les sites de liaison sont déjà occupés et chaque nouvel ajout de ligand a comme seule conséquence sa dilution dans la cellule de travail *Figure 4.5.B.*

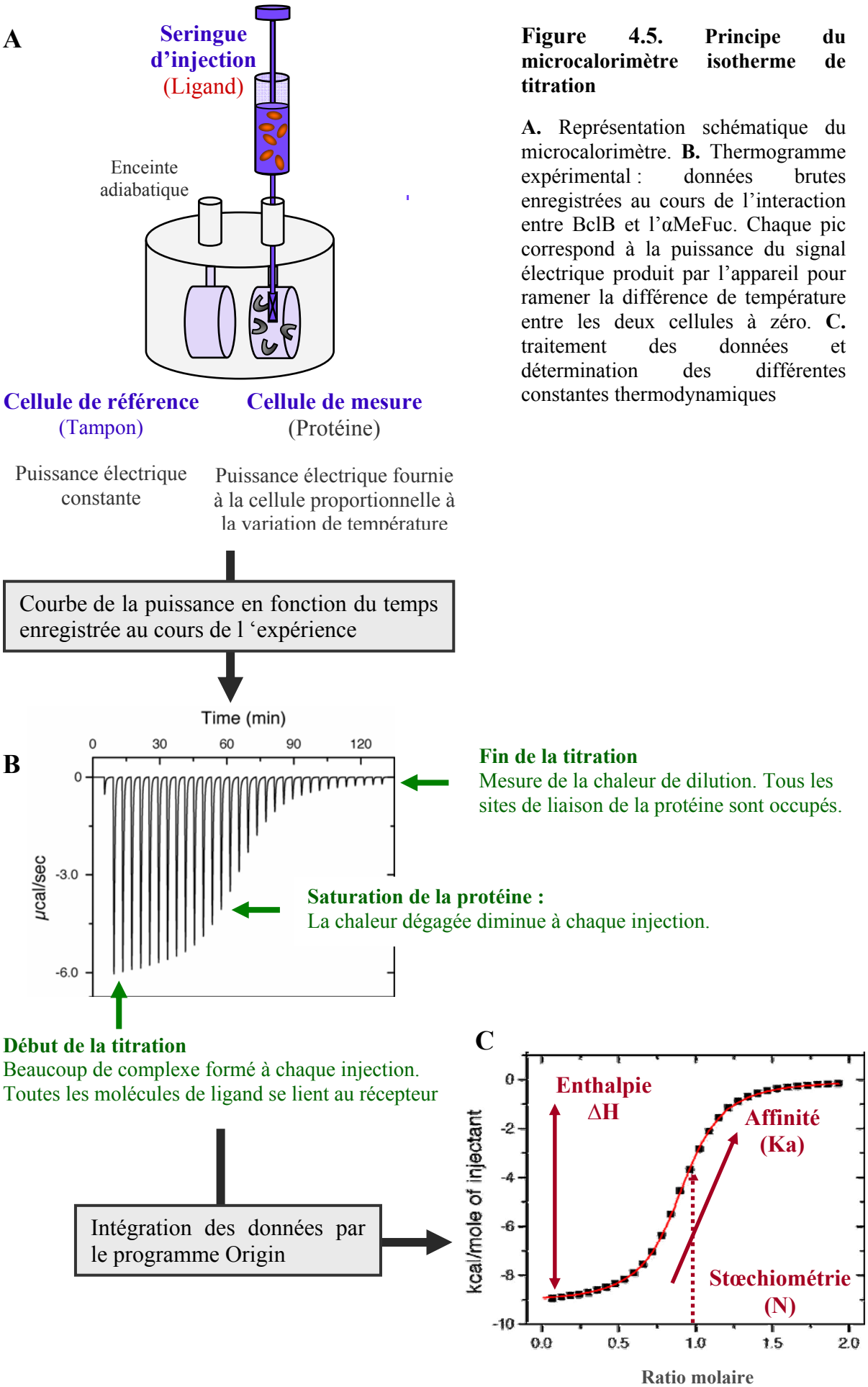


Figure 4.5. Principe du microcalorimètre isotherme de titration

A. Représentation schématique du microcalorimètre. **B.** Thermogramme expérimental : données brutes enregistrées au cours de l'interaction entre BclB et l'αMeFuc. Chaque pic correspond à la puissance du signal électrique produit par l'appareil pour ramener la différence de température entre les deux cellules à zéro. **C.** traitement des données et détermination des différentes constantes thermodynamiques

Ces quantités de chaleur sont ensuite ramenées à la concentration molaire de la protéine, et rapportées en fonction du rapport des concentrations ligand/macromolécule. L'ensemble des points obtenus est appelé *isotherme d'interaction.* L'énergie de dilution peut être évaluée par les dernières injections effectuées lors d'une analyse. Si cette énergie est significative, un « blanc » consistant à titrer le tampon avec le ligand est effectué. L'isotherme ainsi obtenu peut être retranché de l'isotherme d'interaction. L'ajustement d'une courbe théorique adéquate sur la courbe expérimentale permet ensuite au logiciel « Origin » (Microcal) de calculer en une seule analyse la constante d'association, la stœchiométrie ainsi que l'énergie libre du système *Figure 4.5.C.* Des modèles à un site, à deux sites indépendants ou encore à deux sites coopératifs peuvent être utilisés. La variation d'enthalpie influence directement l'amplitude de la courbe, tandis que la position du point d'inflexion donne la stœchiométrie *n*. 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 chacune d'entre elles. Ces paramètres sont donc ajustés de manière à avoir une interaction libérant suffisamment d'énergie pour être mesurée, tout en étant dans une zone qui permette l'application des modèles mathématiques. Pour ce faire, la courbe doit se rapprocher de la forme sigmoïdale. L'étendue de cette zone de validité est déterminée par la valeur sans unité du paramètre de Wiseman, c (Wiseman *et al.* 1989). La valeur de c est calculée par l'équation :

$$c = n[M_o]K_a \qquad \text{(Eq. 4)}$$

Où - c est le paramètre de Wiseman

- $n[M_o]$ représente la concentration initiale de sites de liaison en solution
- Ka représente constante d'association en M^{-1}

En pratique, le paramètre *c* doit être compris entre 10 et 100. Des valeurs supérieures génèrent des courbes dont la pente trop élevée, ne permet pas une mesure fiable de K_a (*Figure 4.6*). Lorsque la constante d'association et la stœchiométrie d'un système sont inconnues, la concentration en macromolécule est choisie de manière à ce qu'une chaleur minimale de 3 à 5 µcal soit dégagée lors de l'injection. La concentration en ligand dans la seringue est généralement de 7 à 20 fois celle de la macromolécule dans la cellule.

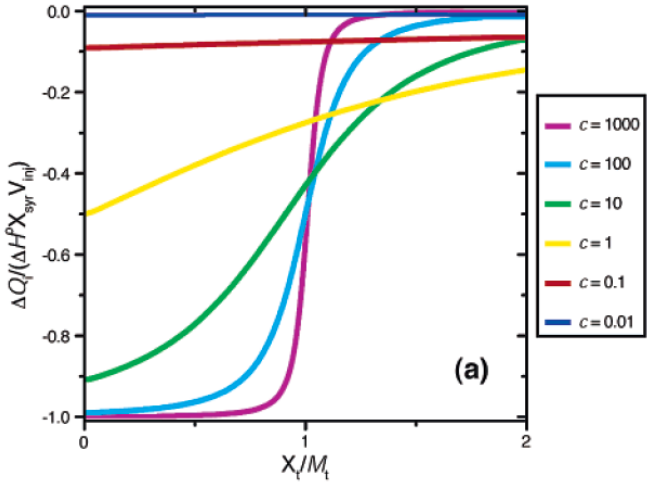


Figure 4.6. Isotherme de Wiseman

La forme de la courbe varie en fonction de la valeur de c qui est le résultat du produit entre la constante d'association et la concentration initiale en macromolécule

4.4. La résonnance plasmonique de surface

Le système Biacore® utilisé au cours de cette étude repose sur le principe de Résonance Plasmonique de Surface (RPS), pour mesurer, en temps réel, toute interaction biologique sans aucun marquage des molécules. Cette technique permet notamment de cribler, de manière qualitative, une bibliothèque de molécules et de déterminer, quantitativement les paramètres cinétiques d'une interaction.

La technique implique la fixation de l'un des partenaires de l'interaction, appelé ligand, sur un biocapteur, la "sensor chip" pendant que l'autre partenaire, l'analyte, passe en solution. Le biocapteur est composé d'une fine couche de verre recouverte d'un film d'or, couplé à différentes fonctions chimiques permettant d'immobiliser le ligand. Lorsqu'un faisceau de lumière monochromatique polarisée illumine, à un angle précis, l'interface entre deux indices de réfraction différents, toute la lumière est réfléchie, aucune réfraction n'est observée. Les électrons délocalisés de la couche périphérique de l'or sont excités par l'onde incidente et entrent en résonance (Figure 4.7.A). Parallèlement, une composante éléctromagnétique de la lumière incidente, l'onde évanescente, se propage perpendiculairement à la surface dans le milieu d'indice de réfraction le plus faible (la solution). La lumière réfléchie présente alors une chute d'intensité à un angle précis, l'angle de résonance, correspondant à l'absorption d'énergie de l'onde évanescente. Les changements de masse induits par l'association ou la dissociation des complexes ligands-analytes modifient la réfringence du milieu et décalent ainsi la position de cet angle.

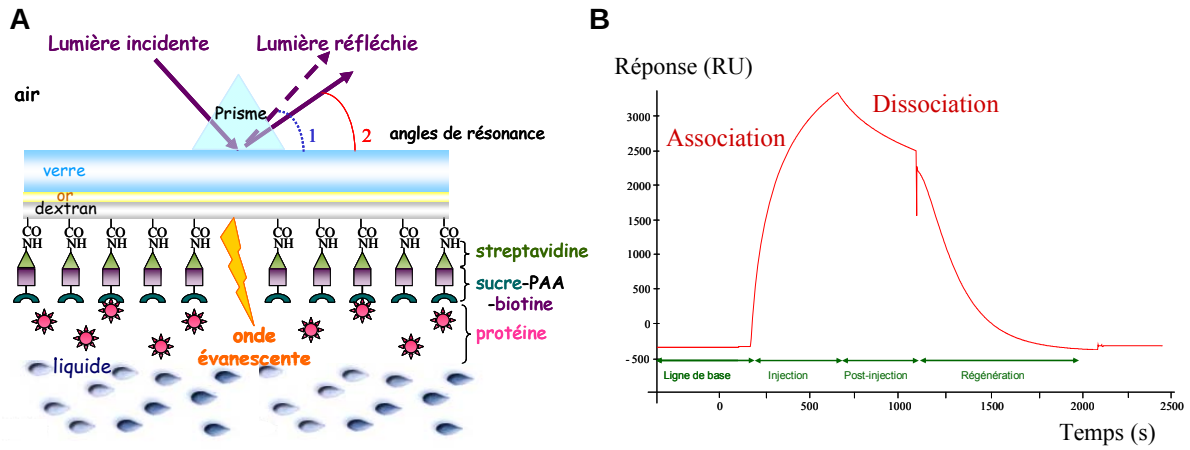


Figure 4.7. Principe du système Biacore®.

A. Schéma du principe de Résonance Plasmonique de Surface. **B.** Sensorgramme.

Ces variations sont enregistrées sur un sensorgramme et quantifiées en unités de résonance (RU), 1000 RU correspondant à un déplacement de l'angle de 0,1° et équivalent à la fixation de 1 ng d'analytes par mm^2. Le sensorgramme est composé d'une phase d'injection (association), d'une phase de post-injection (dissociation) avec l'injection de tampon seul et d'une phase de régénération. Les paramètres cinétiques sont déterminés pendant les phases d'association et de dissociation (*Figure 4.7.B*) et extraits des données expérimentales à partir de modèles mathématiques intégrés avec le logiciel d'analyse.

4.5. La cristallographie aux rayons X

Deux méthodes expérimentales sont aujourd'hui utilisées pour résoudre la structure tridimensionnelle des macromolécules biologiques à l'échelle atomique : la cristallographie aux rayons X et la spectroscopie par résonance magnétique nucléaire (RMN). La limitation la plus sérieuse en RMN est la taille des protéines. Grâce aux progrès réalisés dans ce domaine, il est aujourd'hui possible de déterminer la structure tridimensionnelle en solution de protéines allant jusqu'à 50 kDa, soit moins de 440 acides aminés. A ce jour, plus de 85 % des structures protéiques enregistrées dans la Protein Data Bank, PDB (http://www.rcsb.org/pdb/home/home.do) ont été résolues par diffraction des rayons X (*Tableau 4.1*). La facilité d'accès au rayonnement synchrotron, l'augmentation du pouvoir de calcul des ordinateurs, et surtout les découvertes majeures de la biologie moléculaire ont

contribué à faire de la cristallographie la technique de pointe pour l'étude structurale des biomolécules.

		Type de macromolécule				
		Protéines	*Acides Nucléiques*	*Complexes Protéine/Acide Nucléique*	*Autres*	*Total*
Méthodes expérimentales	*Cristallographie aux rayons X*	42513	1087	1963	24	45587
	RMN	6541	816	138	7	7502
	Microscopie électronique	135	11	49	0	195
	Autres	90	4	4	2	100
	Total	49279	1918	2154	33	53384

Tableau 4.1. Répartition des structures tridimensionnelles des macromolécules recensée dans la PDB en fonction de la technique expérimentale utilisée.

Les méthodes expérimentales rassemblées dans « autres » correspondent à : Spectroscopie infrarouge, Diffraction sur fibres, Diffusion de neutrons. Les macromolécules « autres » rassemblent des molécules comme des antibiotiques ou encore des glucides.

D'après PDB Statistic : http://www.rcsb.org

La résolution des structures cristallographiques de la myoglobine et de l'hémoglobine par John Kendrew et Max Perutz a été récompensée par le prix Nobel de Chimie en 1962. Cette récompense souligne l'importance de la cristallographie aux rayons X et marque l'émergence de la biologie structurale en tant que voie importante pour la recherche en biologie. Contrairement à la RMN, la cristallographie ne souffre pas de limitation en taille de la macromolécule étudiée mais présente une barrière importante : être capable d'obtenir des cristaux de bonne qualité de la macromolécule étudiée. Une fois la protéine cristallisée, se pose encore le *problème des phases*.

4.5.1. La cristallogenèse des protéines

La cristallisation est un transfert de phases dans lequel les macromolécules en solution interagissent et passent progressivement d'un état soluble désordonné à un état solide ordonné. Dans le cristal ainsi formé, la protéine s'organise en un réseau tridimensionnel périodique. Le *diagramme de phases* (*Figure 4.8*) décrit le comportement des

macromolécules en fonction de la concentration en un agent précipitant donné. Dans ce diagramme, la courbe de solubilité définit la limite entre la phase soluble et la phase solide qui sont en équilibre dynamique. Une macromolécule cristallise, lorsqu'elle dépasse cette courbe et entre dans une zone hors équilibre thermodynamique : la sursaturation. Il existe une seconde courbe, qui sépare deux zones sursaturées du diagramme : la zone de nucléation, où la sursaturation élevée conduit à la nucléation du cristal et la zone métastable où la sursaturation plus faible est juste suffisante pour qu'un cristal existe et croisse. Les cristaux sont obtenus en diminuant progressivement la solubilité de la macromolécule, grâce à des agents précipitant qui entrent en compétition avec la macromolécule pour interagir avec l'eau (sel, solvant organique, polymère,…). La protéine peut alors atteindre un état de sursaturation favorable à l'obtention de cristaux, sans toutefois tomber dans la zone de précipitation.

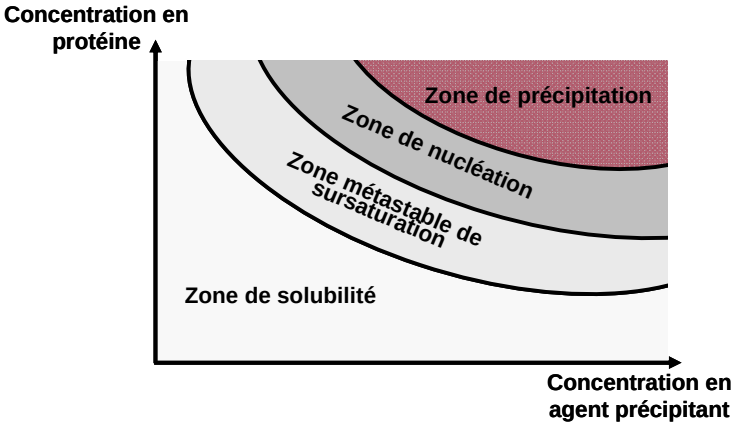


Tableau 4.8. Représentation schématique du diagramme de phase d'une protéine

Alors que la plupart des petites molécules organiques et des substances minérales cristallisent facilement, l'obtention de cristaux de protéines est délicate et constitue en soit un obstacle majeur à franchir pour le bio-cristallographe. Les protéines sont des molécules fragiles et leur stabilité dépend fortement de leur environnement immédiat. En conséquence, ces macromolécules sont plus sensibles aux conditions extérieures et de nombreuses variables doivent être contrôlée pour rendre possible leur cristallisation. Parmi ces paramètres, nous pouvons citer : la pureté des molécules, la solubilité, la température, la concentration, le pH, la force ionique, l'effet de divers additifs, etc… Ainsi, malgré le développement constant de nouvelles méthodes de cristallisation, il est encore parfois difficile voir impossible d'obtenir de bon cristaux de protéine.

Pour être de bonne qualité, les cristaux doivent en général être réguliers avec des faces et formes bien définies (Ducruix and Giégé 1992). Ceci implique de disposer d'une solution initiale de plusieurs milligrammes de protéine pure, soluble et monodisperse. La solution à cristalliser peut être hétérogène en raison d'une purification insuffisante mais aussi d'une protéolyse ou de modifications post-traductionnelles partielles ou hétérogènes. Enfin, des micro-hétérogénéités conformationnelles peuvent apparaître : régions flexibles, équilibre entre différentes formes oligomériques ou agrégation. Mais là encore, il n'existe pas de règle, et dans certains cas, c'est une impureté qui sert de point de départ et permet d'obtenir un cristal ; on parle alors de *nucléation hétérogène*.

La cristallogenèse par diffusion de vapeur

Plusieurs méthodes ont été développées pour cristalliser les protéines**.** Elles ont toutes pour but d'amener la macromolécule en conditions de sursaturation propice à la nucléation, puis à la croissance du cristal. De manière générale, la cristallisation peut être divisée en trois étapes : la nucléation, la croissance cristalline et enfin la cessation de croissance. Une des méthodes la plus fréquemment utilisée est la diffusion de vapeur en goutte suspendue (McPherson 1990) (*Figure 4.9*). Cette technique consiste à déposer sur une lamelle de verre siliconée, une goutte de quelques microlitres de protéine (typiquement 2 à 10µl) mélangée à un volume égal ou similaire de précipitant (tampon, précipitant, sels, additifs ou détergents…). La lamelle est ensuite retournée et scellée hermétiquement sur un réservoir contenant un volume largement supérieur de précipitant (typiquement 0.5 à 1 ml), mais ne contenant pas de protéine. Avec le temps, les composés volatils (eau ou solvants organiques) diffusent par voie gazeuse entre les deux liquides jusqu'à ce que la tension de vapeur soit la même de part et d'autre. L'état de sursaturation s'établit donc progressivement en augmentant simultanément la concentration de la protéine et de l'agent précipitant. Si des conditions appropriées ont été choisies, la protéine peut alors s'agglomérer spontanément et initier la croissance cristalline dans la goutte. Mais, cette méthode reste empirique et le succès aléatoire tant les paramètres physico-chimiques intervenant sont nombreux. La concentration en solutés, le pH et la température, pour ne citer qu'eux, doivent être pris en compte lors des tests successifs tendant à déterminer les conditions de cristallisation. La recherche de ces conditions s'effectue souvent par criblage de nombreuses solutions vendues sous forme de kits. De plus la robotisation permet aujourd'hui de cribler très rapidement un grand nombre de

conditions différentes et ce, avec peu de matériel, augmentant ainsi les chances d'obtention de cristaux.

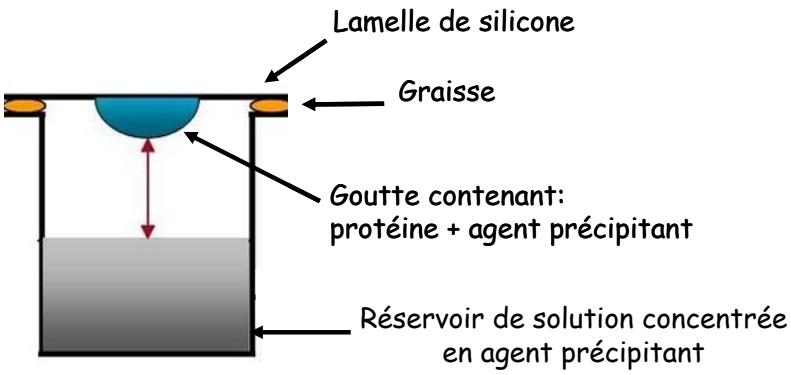


Figure 4.9. Méthode de la goutte pendante.par diffusion de vapeur.

Obtention de complexes

Les cristaux de protéine contiennent en général 30 à 70% d'eau. La majorité de cette eau est désordonnée et occupe les espaces vides entre les protéines. D'autres molécules d'eau, dites ordonnées créent des liaisons hydrogène avec les acides aminés de la surface des macromolécules et occupent le même site sur quasiment toutes les protéines du cristal. Les lectines cristallisées sont donc dans un environnement aqueux quasi naturel. Dans la majorité des cas, leur structure tridimensionnelle native et donc leur activité biologique est conservée.

Caractériser l'interaction entre les lectines et leur sucre nécessite l'obtention de cristaux dans lesquels la protéine est complexée avec son ligand. Pour obtenir de tels cristaux, deux stratégies sont possibles. Tout d'abord, la méthode du « *trempage* », consiste à baigner plus ou moins longtemps un cristal natif (protéine non complexée) dans une solution contenant le ligand afin que celui-ci diffuse dans le cristal. Cette méthode fonctionne à condition que l'accessibilité au site de fixation ne soit pas obstruée par les molécules voisines.

Dans la seconde technique, dite de *co-cristallisation,* la macromolécule et son ligand sont mélangés dans la goutte afin que tous deux interagissent et cristallisent ensemble. Selon les ligands mis en présence, le complexe protéine-ligand peut ne pas cristalliser dans les mêmes conditions que la protéine native et nécessite un nouveau criblage. La co-cristallisation a été choisie au cours de cette étude.

4.5.2. Principe de la diffraction aux rayons X

Définitions du cristal

Un cristal de protéine est un empilement ordonné, infini de motifs identiques (*Figure 4.10*). Dans le cristal, la maille élémentaire est le plus petit volume qui translaté dans les trois directions de l'espace forme un cristal entier. Elle est la plus petite entité contenant toute l'information du cristal et est définie par trois vecteurs qui génèrent ainsi les six *paramètres de la maille* : les trois longueurs des vecteurs a, b, c et trois angles interaxiaux, α (entre c et b), β (entre c et a) et γ (entre b et a).

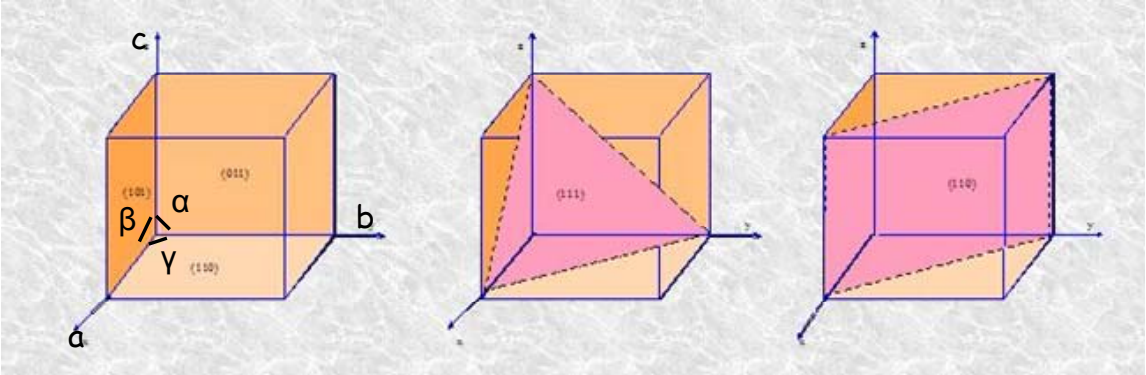


Figure 4.10. Représentation de la maille élémentaire.

Les cubes orange représentent la maille élémentaire définie par les vecteurs a, b et c ainsi que par les angles α, β et γ. Les plans en rose montrent des plans réticulaires

La maille élémentaire est délimitée par un ensemble de points, appelés nœuds du réseau. Il existe 14 types différents de maille élémentaires pour décrire l'arrangement des nœuds en trois dimensions appelés les 14 *réseaux de Bravais*, (*Figure 3.11*). Des opérations de symétries (rotation au tour d'un axe, inversion et réflexion) peuvent être appliquées ensuite dans la maille. Elles appartiennent à un des 32 groupes ponctuels qui combinées aux 14 réseaux de Bravais et à des translations mineures forment les 230 groupes d'espace qui représentent toutes les combinaisons d'opérations de symétrie applicables au motif cristallin. L'unité asymétrique est le plus petit volume qui permet de générer toute la maille à partir des éléments de symétrie de la maille.

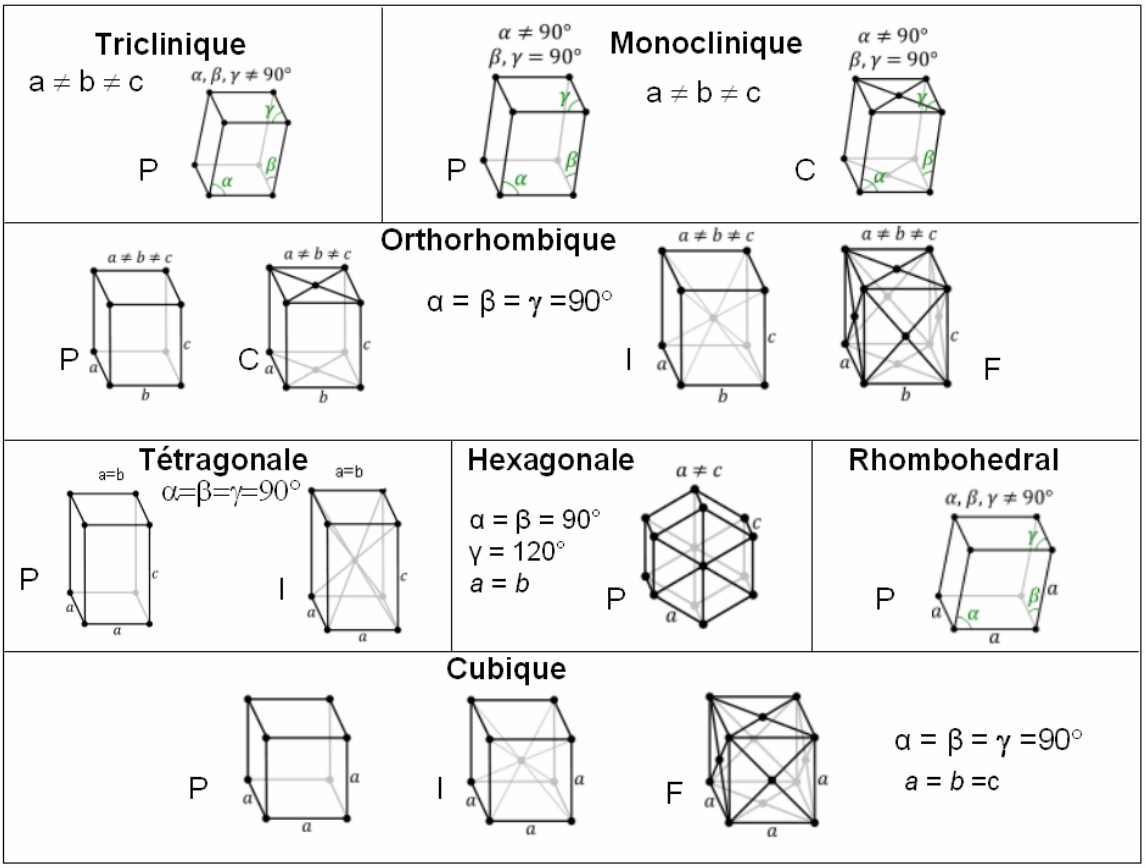


Tableau 4.11. Représentation des 14 réseaux de Bravais

Les protéines étant des molécules chirales, les opérations de symétrie de type réflexion et inversion ne peuvent exister dans leurs cristaux, réduisant ainsi le nombre de groupes d'espace possibles en cristallographie des protéines à 65. L'*unité asymétrique* est la plus petite unité moléculaire (une molécule ou un groupe de molécules) qui suite à l'application d'opérations de symétrie (définis par le groupe d'espace du cristal) permet d'obtenir la maille élémentaire.

Processus de diffraction

Lorsqu'un cristal est placé dans un faisceau de rayons X, la densité électronique de chacun de ses atomes interagit avec les rayons en la diffusant dans toutes les directions de l'espace. C'est la présence du réseau cristallin qui amplifie la radiation dans certaines directions et donnant ainsi naissance au diagramme de diffraction. Les bases mathématiques et géométriques du processus de diffraction ont été formulées par Bragg, Laue et Ewald au début du siècle. Sir Lawrence Bragg a montré que la diffraction par un cristal peut être considérée comme la réflexion du faisceau incident par une série de plans parallèles et équidistants (les *plans réticulaires* du cristal) (*Figure 4.10*), comme le ferait une série de

miroirs. Ces plans également appelés plans de Bragg coupent le cristal dans toutes les directions en passant par les nœuds du réseau. Chaque ensemble de plans est identifié par les indices de Miller (h,k,l) définissant ainsi une famille de plans. Les rayons X réfléchis par des plans adjacents parcourent des distances différentes *Figure 4.12*.

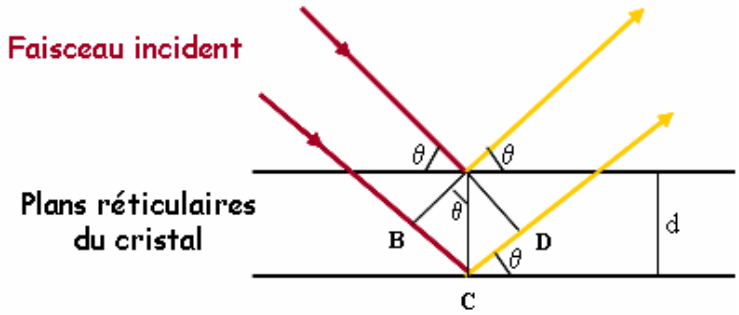


Figure 4.12. Loi de Bragg

La diffraction n'a lieu que si la différence de trajet parcouru par les rayons X correspond à un nombre entier de fois de longueur d'onde.

Lorsque la différence de marche entre les rayons incidents et les rayons diffractés est égale à un nombre entier de fois la longueur d'onde, c'est-à-dire lorsque la loi de Bragg est satisfaite, alors il y interférence constructive.

$$2d_{hkl}\sin\theta = n\lambda \qquad \text{(Eq. 5)}$$

où - *λ est la longeur d'onde du faisceau incident*

- *θ est l'angle sous lequel le rayon traverse le cristal*
- *dhkl est la distance inter-planaire entre les plans cristallins parallèles produisant un rayon diffracté*

Ainsi, lorsque cette condition géométrique est satisfaite, des interférences constructives se produisent et s'additionnent pour être observées sous forme de tâche de diffraction qu'on appelle aussi *réflexions*. En revanche, si les rayons X frappent les plans parallèles avec un angle qui ne satisfait pas cette équation, alors les rayons adjacents seront diffractés avec des phases différents donnants lieux à des interférences destructrices. La position de chaque tache observée sur le cliché de diffraction est donc reliée à une famille de plans réticulaires définie par ses *indices de Miller* (h,k,l). Le cliché de diffraction est appelé *espace réciproque* par opposition à l'*espace réel* du cristal. Cet espace réciproque est défini par les *vecteurs*

réciproques a*, b* et c* perpendiculaires aux vecteurs b et c, a et c et a et b respectivement et de longueur 1/a, 1/b et 1/c. Ils décrivent la position des ondes diffractées alors que l'espace réel décrit la position des atomes dans le cristal.

La construction mathématique de la *sphère d'Ewald* associée au réseau réciproque est une vue de l'esprit qui permet de visualiser la diffraction dans l'espace *Figure 3.13*.
Le faisceau incident est placé sur le diamètre d'une sphère de rayon 1/λ. Le réseau réciproque représente le cristal et son origine est placée à l'endroit ou le faisceau incident émerge de la sphère. Lorsqu'un point du réseau réciproque rencontre la sphère d'Ewald, la loi de Bragg est obéie et il y a diffraction.

Puisque que pour une position fixe donnée, seuls certains plans donneront lieu à des réflexions, il est nécessaire pour explorer tous les atomes, de faire tourner le cristal sur lui-même lors de la collecte de données afin que tour a tour chaque plan puisse engendrer ses réflexions propres.

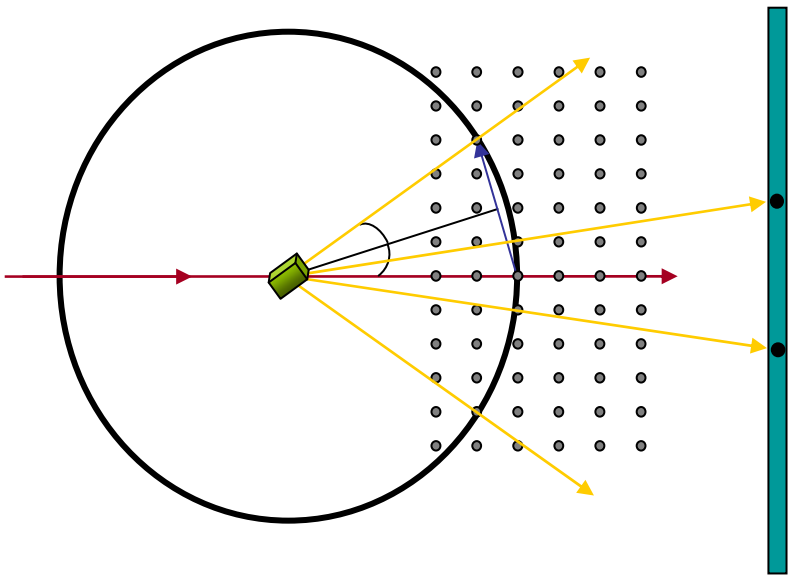


Figure 4.13. Construction d'Ewald

Le cristal placé au centre d'une sphère de rayon 1/λ tourne sur lui-même, ce qui équivaut à tourner le réseau réciproque qui a son origine sur l'intersection de la sphère avec le rayon incident. Quand un nœud du réseau réciproque croise la sphère d'Ewald, le rayon est diffracté d'un angle 2Θ.

4.5.3. Collecte des données des données expérimentales

Préalablement à la collecte, les cristaux sont plongées dans une solution cryoprotectrice afin d'éviter la formation de cristaux de glace qui détruisent la symétrie du réseau. En effet, le faisceau de rayonsX incident endommage la protéine et les molécules de solvant. Il se forme des radicaux libres qui à leur tour endommagent les molécules du cristal. De plus, pendant la collecte de données, l'intensité du faisceau synchrotron, provoque un échauffement local du cristal, pouvant aboutir à sa destruction complète. Les cristaux sont donc congelés dans l'azote liquide (100K) afin de les protéger et de prolonger leur durée de vie. Le choix du cryoprotectant : glycérol, PEG (poly-éthylène glycol), MPD (2-méthyl-2,4-pentanediol) ou huile minéral est donc déterminant. Après un bref trempage dans la solution cryoprotectante (5-15 secondes), les cristaux sont montés sur des boucles de nylons (<1mm) avant d'être plongés dans l'azote liquide.

La collecte de données est ensuite effectuée sous un flux d'azote continu à l'ESRF (European Synchrotron Radiation Facility, Grenoble). Pendant la collecte, le cristal tourne par paliers de 0 à 1° autour d'un axe perpendiculaire au faisceau de rayons X et les rayons diffractés sont collectés sur une caméra CDD capable d'enregistrer un grand nombre de réflexions dans un temps d'acquisition très court. Chaque "image", appelée *cliché de diffraction* correspond à une tranche de l'espace réciproque. La position des taches nous renseigne sur le groupe d'espace alors que l'intensité de ces mêmes taches dépend de la position et de la nature des atomes présents dans l'unité asymétrique.

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

La première étape de traitement des données, consiste à attribuer pour chaque tâche des coordonnées dans l'espace réciproque (indices hkl) ainsi qu'une intensité I associée à une erreur expérimentale σI. Cette étape, dite d'*indexation* est réalisée grâce aux programmes informatiques tel que Mosflm. L'ensemble des images intégrées sont mises à l'échelle, puis les donnés sont réduites (réflexions équivalentes moyennées) pour être ensuite converties en facteurs de structure observés Fobs. La qualité et la résolution du jeu de données est ensuite évaluée lors de l'analyse de plusieurs paramètres tels la *complétude (* pourcentage de réflexions uniques mesurées), la *multiplicité* (rapport entre les réflexions totales et les réflexions uniques) et le rapport entre l'intensité d'une réflexion (I), et son erreur (σI). Dans un bon jeu de données, la complétude doit avoisiner au moins les 95%, la multiplicité doit être

la plus grande possible. La valeur du Rmerge correspond à l'accord entre les réflexions équivalentes reliées entre elles par opérations de symétrie. La résolution maximale d'un jeu de donnée est déterminée par la couche de résolution ayant un rapport (I/σI) de 2 et un Rmerge < 40%.

4.5.5. Facteurs de structure et densité électronique

Chaque tache observée correspond à une onde résultant de la somme des diffractions contributives de tous les atomes appartenant à la famille de plans réticulaires qu'elle représente. La fonction mathématique qui décrit cette onde est appelée le *Facteur de structure* $F(\vec{s})$, que l'on note aussi F_{hkl}. Cette fonction peut être calculée par la somme des contributions des densités électroniques de tous les atomes de la maille cristalline par la formule suivante:

$$F(\vec{s}) = F_{hkl} = \iiint \rho(x,y,z) e^{2\Pi i (hx+ky+lz)} \, dV \qquad \text{(Eq. 6)}$$

Le *facteur de structure* est la transformée de Fourier de la distribution de la densité électronique dans le cristal. En appliquant la transformée de Fourier, on exprime la densité électronique (ρ) 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}:

$$\rho(x,y,z) = \frac{1}{V} \sum_{h,k,l} F_{hkl} \, e^{-2\Pi i (hx+ky+lz)} \qquad \text{(Eq. 7)}$$

La densité électronique peut ainsi se représenter sous la forme d'une somme de termes exponentiels complexes, décrivant chacun une onde plane se propageant dans la direction d'un nœud du réseau réciproque. Chaque onde est caractérisée par son amplitude |Fhkl| qui est proportionnelle à la racine carrée de l'intensité mesurée $(I)^{1/2}$ et par sa phase φ_{hkl}. Malheureusement, les phases indispensables à la construction de la carte de densité électroniques sont perdues au cours de l'expérience et leur obtention représente un problème majeur en cristallographie.

4.5.6. Obtention des phases par la méthode de remplacement moléculaire

Il existe plusieurs méthodes permettant d'obtenir les phases nécessaires à la construction de la carte de densité électronique. Elles utilisent toutes un outil très important en cristallographie appelée la fonction de Patterson P(u,v,w). Celle-ci est indépendante des phases et dépend seulement des intensités déterminées expérimentalement. Elle correspond à la est la transformée de Fourier des intensités. La fonction de Patterson P(u,v,w), est large quand ρ(x,y,z) et ρ(x+u, y+v, z+w) sont tous les deux larges donc quand u,v,w représente un vecteur interatomique.

$$P(u,v,w) = \frac{1}{V}\sum_h \sum_k \sum_l [F_{hkl}]^2 e^{-2\pi i(hu+kv+lw))} \qquad \text{(Eq. 8)}$$

Une carte de Patterson dans laquelle les pics correspondent à tous les vecteurs inter-atomiques qui peuvent être tracés entre les atomes de la maille réelle est calculée. Le nombre de pics est proportionnel au carré du nombre d'atomes N, leur intensité est quand à elle proportionnelle au produit du nombre atomique des deux atomes à chaque extrémité du vecteur.

Différentes méthodes permettent d'obtenir les phases par des méthodes expérimentales. Parmi elles, les méthodes SAD « Single anomalous diffraction » et MAD « Multiwavelength anomalous diffraction » utilisent la diffusion anomale, on trouve également la méthode du replacement isomorphe multiple « MIR multiple isomorphous replacement ». Ces méthodes sont basées sur l'introduction dans le cristal de un ou plusieurs atomes dits lourds (MIR) ou diffuseurs anomaux (SAD/MAD) ayant pour effet des variations d'intensités des ondes collectées. Le calcul de la carte de différences de Patterson permet d'en déduire la position des atomes introduits dans la maille cristalline et de calculer leur contribution aux amplitudes et aux phases des rayons diffractés par les autres atomes et enfin de construire une carte de densité électronique.

Dans le cas de la diffusion anomale, la méthode repose sur la propriété diffusion « anomale » de certains atomes lorsqu'ils sont irradiés par un faisceau de rayons X avec une longueur d'onde proche de l'énergie de leur seuil d'absorption. Parmi les atomes le plus souvent utilisés, on trouve le souffre, le brome, le xénon, le sélénium, qui peut être incorporé à la protéine recombinante en remplaçant les méthionines par des séléniomethionine.

Il est parfois possible d'utiliser les phases d'une autre molécule comme phases de départ. En effet, si par chance, la protéine cristallisée possède un homologue de structure connue, la méthode du remplacement moléculaire est alors la méthode de choix pour résoudre le problème des phases. Pour que cette méthode fonctionne, les deux protéines doivent avoir au moins de 20 % de similarité. Les similarités de séquence en effet impliquent le plus souvent une similarité structurale. En pratique, le remplacement moléculaire procède en plaçant la molécule modèle dans la maille de structure inconnue de façon à reproduire au mieux les facteurs de structures observés et d'utiliser les phases ainsi obtenues comme phases de départ. Pour cela une recherche en six dimensions de toutes les orientations et translations du modèle est nécessaire. Ce procédé est long en raison du nombre de paramètres à considérer. Les programmes de recherche automatique fonctionnent le plus souvent en effectuant deux recherches tridimensionnelles : dans un premier temps une fonction de rotation permet d'orienter la molécule dans la maille puis sa position est déterminée par la fonction de translation. De nos jours des programmes de recherche automatique de plus en plus puissants sont disponibles tels que AMoRe (Navaza 1994, Navaza and Saludjian 1997), PHASER (McCoy 2007) ou encore MOLREP (Vagin and Teplyakov 1997) et permettent d'obtenir une solution en quelques minutes.

4.5.7. Construction, affinement et validation du modèle

Grâce aux phases obtenues par ces différentes techniques, une première carte de densité électronique est construite et visualisée à l'aide de logiciel graphique tel que Coot (Emsley and Cowtan 2004). Cette première carte contient des erreurs et n'est pas toujours de qualité suffisante pour déterminer la position exacte des atomes. Toutefois, la construction d'un premier modèle permet le calcul d'un nouvel ensemble de phases et l'obtention d'une carte de densité électronique améliorée.

Ce premier modèle est ensuite affiné afin d'éliminer les erreurs qu'il contient. Les différentes étapes d'affinement ont pour but de minimiser les différences entre le facteur de structure calculé à partir du modèle (Fc) et le facteur de structure observé expérimentalement (Fo). Entre chaque cycle d'affinement, après inspection des cartes de densités électroniques 2Fo-Fc et Fo-Fc, le modèle est ajusté manuellement. La carte 2Fo-Fc donne la densité électronique pour le modèle observé et la carte Fo-Fc indique les différences de densité entre les cartes observées et calculées. Des logiciels d'affinement comme Refmac (Murshudov *et*

al. 1997) utilisent un algorithme de type statistique nommé l'estimation du maximum de vraisemblance. Cet algorithme de type statistique, nommé le maximum likehood maximise la probabilité que le facteur de structure observé Fo soit en accord avec le modèle, tout en considérant de nombreux paramètres tels que les coordonnées et le type des atomes, le facteur d'agitation thermique B, mais aussi des paramètres structuraux comme les longueurs, les angles de liaison et les torsions. La correspondance entre le modèle et les données expérimentales est mesurée grâce au facteur d'accord cristallographique R qui nous renseigne sur l'accord entre les amplitudes des facteurs de structure calculées et observées et la progression de l'affinement. Il est calculé à partir de la formule suivante :

$$R = \frac{\sum_{hkl} \left\| F_{obs}(hkl) \right| - \left| F_{calc}(hkl) \right\|}{\sum_{hkl} \left| F_{obs}(hkl) \right|} \qquad \text{(Eq. 9)}$$

La valeur de R nous renseigne sur la progression de l'affinement de la structure : plus sa valeur est basse, plus le modèle est en accord avec les données. Un second critère de qualité : le Rfree est calculé de la même manière que le facteur R, mais à partir d'une fraction de réflexions (5% à 10% du jeu de données) qui ne sont pas prises en compte en compte au cours de l'affinement. Si la structure n'est pas suraffinée, les facteurs R et Rfree doivent diminuer en parallèle.

En fin d'affinement, les paramètres structuraux du modèle final sont vérifiés et comparés aux valeurs des structures protéiques ave c les logiciels PROCHECK ou WHATIF (Laskowski *et al.* 1993, Vriend 1990). La structure peut alors être déposée dans la base de données internationale : RCSB Protein Data Bank.

4.6. Le Biofilm Ring Test

Il existe différentes méthodes permettant d'observer, caractériser, mesurer et quantifier les biofilms. Parmi elles, on trouve les techniques de microscopie (épifluorescence, confocale, microscopie électronique à balayage ou à transmission) qui peuvent être utilisée à des fins de dénombrement ou de mesure qualitative des biofilms. Il existe également des techniques de décrochage et dénombrement des bactéries sessiles. Enfin, la méthode de quantification la plus répandue est la coloration au cristal violet. Dans cette technique, les bactéries adhérées

sont révélées par une coloration au crystal violet puis quantifiées par une mesure spectrophotométrique de la densité optique à 570 nm. Ce procédé utilisable en plaque multi-puits permet de tester simultanément plusieurs souches, mutants et paramètres influençant la formation du biofilm. Cette technique présente plusieurs désavantages : des cultures d'au moins 24 ou 48 heures sont nécessaire, les étapes de lavage permettant l'élimination des cellules non adhérées influence très fortement le résultat final.

La méthode du "BioFilm Ring Test®", développée par la société BioFilm Control est un procédé de criblage qui permet d'évaluer l'aptitude des microorganismes à former des biofilms ainsi que la sélection des molécules capables d'inhiber ou retarder leur formation Cette technique ne nécessite ni lavage ni coloration des cellules et permet d'obtenir un résultat en quelques heures. L'analyse est faite dans de microplaque de 96 puits. Une solution bactérienne mélangée à des particules magnétisables est déposée dans chaque puits, sous le fond duquel un aimant est présenté. En l'absence de formation de biofilm, les particules s'amassent au centre sous l'influence du champ magnétisable pour former un spot. Lorsqu'un biofilm se développe dans le puit, les particules sont progressivement immobilisées. Moins nombreuses à réagir au champ magnétique, elles forment alors un spot moins net, qui disparaît, lorsque le biofilm est totalement formé (*Figure 4.14A*).

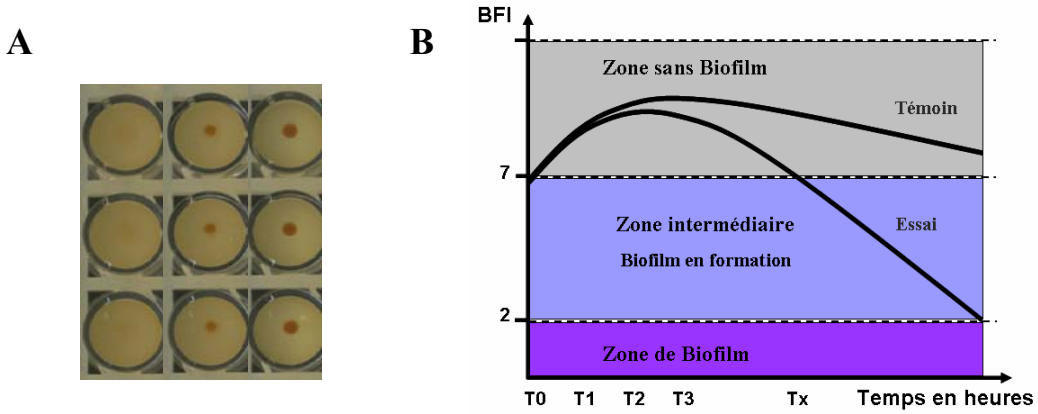


Figure 4.14. Mesure de la formation du biofilm par le Biofilm Ring Test

A. Image acquise après aimantation des particules magnétisables contenues dans chaque puit. En haut, les particules sont mobiles, il n'y a pas de biofilm et l'indice de biofilm, BFI est de 7. En bas, le biofilm couvre toute la surface des puits et les particules sont immobilisées ; le BFI est inférieur à 2. **B.** Cinétique de formation du biofilm

L'analyse des images du fond du puits avant et après aimantation donne un indice biofilm (BFI). Lorsque celui-ci est supérieur à 7, il n'y a pas de biofilm. Lorsque le biofilm se développe, le BFI est alors compris entre 7 et 2, puis se stabilise autour de 2 en présence d'un biofilm totalement formé (*Figure 4.14B*).

Chapitre 5

Etude structurale et fonctionnelle de la lectine BclA

La lectine la plus courte, BclA (*Burkholderia cepacia* lectine A) a été produite de façon recombinante. L'obtention de cristaux nous a permis de résoudre sa structure cristallographique en complexe avec l'α-methyl-mannoside (αMeMan) à une résolution de 1.7Å. La lectine s'organise en un homodimère et la carte de densité électronique montre clairement la présence de deux ions Ca^{++} et un sucre dans chaque site de liaison. L'organisation dimérique de la protéine en solution a été confirmée par une chromatographie d'exclusion. Les expériences de printed array ont permis de mettre en évidence la spécificité stricte de BclA pour les structures glycanniques de type oligo-mannoside. L'affinité de cette lectine pour l'αMaMan a pu être mesurée par microcalorimétrie. Ce travail a fait l'objet d'une publication dans *Biochemical Journal*. Les expériences de résonance plasmonique de surface et de chrommatographie d'exclusion ont été réalisées par Lenka Malinovsk'a (Brno) dans le cadre de sa thèse

Les résultats de Printed array ont montré que la lectine BclA reconnaît les N-Glycannes de type oligo-mannosides. L'extrémité Manα1−3(Manα1−6)Man est particulièrement bien reconnue. Nous avons donc étudié la thermodynamique de l'interaction entre BclA et ce trisaccharide par microcalorimétrie isotherme de titration. Les structures de BclA en complexe avec le Manα1−3(Manα1−6)Man et le Manα1−3Man ont également pu être résolues. Dans le cristal de BclA complexée au trisaccharide, le sucre forme un pont entre deux protéines illustrant ainsi très bien les propriétés d'hémagglutination de la lectine. Ce phénomène étant généré par l'état cristallin, des expériences de microscopie par force atomique ont été entreprise pour confirmer ce phénomène (Collaboration Frédéric Dubreuil, CERMAV-CNRS). Ces résultats sont présentés dans un article en préparation.

(Article n°1)

Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia*

Emilie LAMEIGNERE, Lenka MALINOVSK´A, Margita SL´AVIKOV´A, Eric DUCHAUD, Edward P. MITCHELL,

Biochem. J. (2008) 411, 307–318 (Printed in Great Britain) doi:10.1042/BJ20071276

BJ www.biochemj.org

BJ Structure

Biochem. J. (2008) **411**, 307–318 (Printed in Great Britain) doi:10.1042/BJ20071276

online data

Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia*

Emilie LAMEIGNERE*[1], Lenka MALINOVSKÁ†[1], Margita SLÁVIKOVÁ†, Eric DUCHAUD‡, Edward P. MITCHELL§, Annabelle VARROT*, Ondrej ŠEDO‖, Anne IMBERTY*[2] and Michaela WIMMEROVÁ†[2]

*CERMAV-CNRS (affiliated with Université Joseph Fourier and belonging to ICMG), BP 53, F-38041, Grenoble, Cedex 09, France, †National Centre for Biomolecular Research (NCBR) and Department of Biochemistry, Masaryk University, Kotlarska 2, 61137 Brno, Czech Republic, ‡INRA, Unité Virologie et Immunologie Moléculaires UR892, Jouy-en-Josas, France, §ESRF Experiments Division, BP 220, F-38043, Grenoble Cedex, France, and ‖Department of Functional Genomics and Proteomics, Institute of Experimental Biology, Masaryk University, Kotlarska 2, 61137 Brno, Czech Republic

Chronic colonization of the lungs by opportunist bacteria such as *Pseudomonas aeruginosa* and members of the Bcc (*Burkholderia cepacia* complex) is the major cause of morbidity and mortality among CF (cystic fibrosis) patients. PA-IIL (*lecB* gene), a soluble lectin from *Ps. aeruginosa*, has been the subject of much interest because of its very strong affinity for fucose. Orthologues have been identified in the opportunist bacteria *Ralstonia solanacearum*, *Chromobacterium violaceum* and *Burkholderia* of Bcc. The genome of the J2315 strain of *B. cenocepacia*, responsible for epidemia in CF centres, contains three genes that code for proteins with PA-IIL domains. The shortest gene was cloned in *Escherichia coli* and pure recombinant protein, BclA (*B. cenocepacia* lectin A), was obtained. The presence of native BclA in *B. cenocepacia* extracts was checked using a proteomic approach. The specificity of recombinant BclA was characterized using surface plasmon resonance showing a preference for mannosides and supported with glycan array experiments demonstrating a strict specificity for oligomannose-type N-glycan structures. The interaction thermodynamics of BclA with methyl α-D-mannoside demonstrates a dissociation constant (K_d) of 2.75×10^{-6} M. The X-ray crystal structure of the complex with methyl α-D-mannoside was determined at 1.7 Å (1 Å = 0.1 nm) resolution. The lectin forms homodimers with one binding site per monomer, acting co-operatively with the second dimer site. Each monomer contains two Ca^{2+} ions and one sugar ligand. Despite strong sequence similarity, the differences between BclA and PA-IIL in their specificity, binding site and oligomerization mode indicate that the proteins should have different roles in the bacteria.

Key words: bacterial lectin, *Burkholderia cenocepacia*, *Burkholderia cepacia* complex, mannose-binding lectin, *Pseudomonas aeruginosa*.

INTRODUCTION

The Bcc (*Burkholderia cepacia* complex) is a collection of genetically distinct but phenotypically similar bacteria that are divided into at least nine species [1]. They colonize plant–rhizosphere systems, and aquatic and human environments, and their abilities to degrade organic compounds and to produce antimicrobial compounds give them environmental and biotechnological potential [2]. However, the Bcc bacteria are also opportunistic pathogens. Among them, *Burkholderia cenocepacia* and *Burkholderia multivorans* cause variable lung infections in CF (cystic fibrosis) patients, the most dangerous being 'cepacia syndrome', which is characterized by a rapid decline in lung function [3–5]. Some strains with resistance to antibiotics are also responsible for nosocomial infections.

Bacteria of the Bcc share many characteristics with *Pseudomonas aeruginosa*. Both are opportunistic bacteria that are found in various environments, including soil, water and vegetation. They can both form biofilms, and the occurrence of mixed biofilms has been demonstrated with the use of common signal molecules in the quorum-sensing system [6]. Moreover, a frequent exchange of genetic material has been confirmed between the two organisms [7].

Among the virulence factors deployed by *Ps. aeruginosa* in the infection progress, soluble lectins PA-IL (*Ps. aeruginosa* lectin 1), coded by the *lecA* gene, and PA-IIL (*Ps. aeruginosa* lectin 2), coded by the *lecB* gene, have high affinity for galactose and fucose respectively [8,9]. Their ability to bind to the carbohydrate moiety of mucus or glycoconjugates at epithelial cell surfaces make them good candidates for host recognition and cell-surface adhesion. PA-IIL has also been proposed to play a role in biofilm formation [10]. Although orthologues of the *lecA* gene have not been identified in other pathogenic bacteria at the present time, *lecB*-like sequences are present in the genomes of *Ralstonia solanacearum* and *Chromobacterium violaceum* [11,12], two opportunistic pathogens that can cause rapid infections and septicaemia in immunocompromised humans. Sequence and structural comparisons have shown that the three lectins PA-IIL, RS-IIL (*R. solanacearum* lectin 2) and CV-IIL (*C. violaceum* lectin 2) share the same tetrameric arrangement with the unique characteristic of the presence of two Ca^{2+} ions in the monosaccharide-binding site [13–15]. These lectins have unusually

Abbreviations used: Bcc, *Burkholderia cepacia* complex; BclA, *Burkholderia cenocepacia* lectin A; Biot–PAA, biotinylated polyacrylamide; CF, cystic fibrosis; CV-IIL, *Chromobacterium violaceum* lectin 2; DLS, dynamic light scattering; DSC, differential scanning calorimetry; LB, Luria–Bertani; LC, liquid chromatography; MALDI, matrix-assisted laser-desorption ionization; MS/MS, tandem MS; NCBI, National Center for Biotechnology Information; PA-IIL, *Pseudomonas aeruginosa* lectin 2; PEG, poly(ethylene glycol); RS-IIL, *Ralstonia solanacearum* lectin 2; Req, response units corresponding to steady-state equilibrium; RU, resonance units; SPR, surface plasmon resonance.

[1] These authors contributed equally to the work.

[2] Correspondence may be addressed to either of these authors (email anne.imberty@cermav.cnrs.fr or michaw@chemi.muni.cz).

The structural co-ordinates for BclA have been deposited in the Protein Data Bank under accession code 2VNV.

© The Authors Journal compilation © 2008 Biochemical Society

Biochemical Journal

high affinity for their monosaccharide ligands (in the micromolar range) and they can bind both L-fucose and D-mannose, with their preference depending on a particular stretch of three amino acids in the binding site [9].

Since bacteria of the Bcc are closely related in behaviour and taxonomy to *Ps. aeruginosa*, it is perhaps not surprising to locate *lecB*-like genes in the recently sequenced genome of some strains, particularly of *B. cenocepacia*. Indeed, several genes coding for lectins or lectin domains with similarity to PA-IIL have been identified in each strain. In most cases, an uncharacterized N-terminal domain is attached to the lectin. However, we have demonstrated in the present study that the resulting *Burkholderia* lectins are rather different to PA-IIL and the other members of this family. In the present paper, we describe the cloning, production, specificity and structure of BclA (*B. cenocepacia* lectin A), the product of the smallest *lecB*-like gene from *B. cenocepacia* J2315 which is the ET12 lineage epidemic strain that has caused significant mortality among CF patients in the U.K. [16].

EXPERIMENTAL

Partial purification of native BclA and MS characterization

B. cenocepacia strain J2315 (CCM 4899, LMG 16656) obtained from the Czech Collection of Microorganisms at Brno, Czech Republic, was grown with shaking at 37 °C in LB (Luria–Bertani) medium overnight. Batches of LB medium (250 ml in 1 litre Erlenmeyer flasks) were inoculated with 2.5 ml of the overnight culture, and D-mannose was added to a final concentration of 50 mM. Cells were incubated at 30 °C with shaking for 4 days. Afterwards, cells were removed by centrifugation at 8000 ***g*** for 10 min (4 °C), resuspended in 10 ml of 20 mM Tris/HCl buffer (pH 7.5) and disintegrated by sonication. The cytosolic extract was separated by centrifugation at 8000 ***g*** for 10 min (4 °C), filtered through a 0.22-μm-pore-size sterile filter and applied to a D-mannose–agarose HR5/10 column equilibrated in 20 mM Tris/HCl (pH 7.5), 100 mM NaCl and 100 μM $CaCl_2$. Bound proteins were eluted by 0.1 % SDS, dialysed overnight against distilled water and freeze-dried. Freeze-dried proteins were separated by SDS/PAGE (14 % gel) and stained with Coomassie Blue R250. A band corresponding to the expected molecular mass of the BclA lectin was excised from the gel and analysed by MS. After destaining, the protein was incubated with trypsin (sequencing grade, Promega) at 37 °C for 16 h. Digested peptides were extracted from gels using 50 % acetonitrile solution with 5 % methanoic (formic) acid. Protein digests were analysed by MALDI (matrix-assisted laser-desorption ionization)-MS and subsequently using LC (liquid chromatography)–MS/MS (tandem MS).

MALDI-MS analyses were performed on a Bruker Reflex IV (Bruker Daltonik) operated in reflector mode with detection of positively charged ions using α-cyanohydroxycinnamic acid as matrix. LC–MS/MS experiments used an HPLC system consisting of a gradient pump (Ultimate), autosampler (Famos) and column switching device (Switchos; LC Packings) coupled with an Esquire 2000 ion-trap mass spectrometer (Bruker Daltonik). Tryptic digests were concentrated and desalted using a PepMap C_{18} trapping column (300 μm × 5 mm; LC Packings) with a typical sample volume of 10 μl. After washing with 0.1 % methanoic acid, the peptides were eluted from the trapping column using an acetonitrile/water gradient (4 μl/min) on to a fused silica capillary column (320 μm × 180 mm, 4 μm pore size; Jupiter Proteo, Phenomenex), for peptide separation.

The MASCOT 2.0 (MatrixScience) search engine was used for processing the MS and MS/MS data. Sequence searches were performed against the NCBI (National Center for Biotechnology Information) protein database and against a local database containing corresponding protein sequences.

Gene cloning

The oligonucleotide primers were as follows: 5′-GGAGAAA-CATATGGCTGATTCTCAAACGTC-3′ (30-mer) and 5′-GT-AAGCTTAGCCGATCGGCCACTG-3′ (24-mer). The former was designed for the introduction of NdeI and the latter for HindIII restriction sites (underlined sequences). PCR was performed using Pfu polymerase (Promega) and genomic DNA from *B. cenocepacia* strain J2315 as a template. After digestion with NdeI and HindIII, the amplified fragment was introduced into the multiple cloning site of pRSET vector (Invitrogen), resulting in plasmid pRSETbcla.

Lectin expression, purification and characterization

Escherichia coli BL21 (DE3) cells containing the plasmid pRSETbcla were cultured in LB broth medium at 37 °C. When the culture reached a D_{600} of 0.5–0.6, IPTG (isopropyl β-D-thiogalactoside) was added to a final concentration of 0.5 mM. Cells were harvested after 3 h incubation at 30 °C, washed and resuspended in equilibrating buffer (20 mM Tris/HCl, 100 mM NaCl and 100 μM $CaCl_2$, pH 7.5). The cells were broken by cell disruption (Constant Cell Disruption System). After centrifugation at 50 000 ***g*** for 30 min at 4 °C and filtration, the supernatant was purified further on a D-mannose–agarose column (Sigma–Aldrich). BclA was allowed to bind to the immobilized mannose in equilibrating buffer. After washing with buffer (20 mM Tris/HCl, 1 M NaCl and 100 μM $CaCl_2$, pH 7.5), it was eluted using 20 mM Tris/HCl, 100 mM NaCl and 1 mM EDTA (pH 7.5). The purified protein was extensively dialysed against 5 mM $CaCl_2$ for 2 days and against distilled water for a further 4 days, freeze-dried and stored at −20 °C.

A Superose 12 column (GE Healthcare), equilibrated with 20 mM Tris/HCl and 0.3 mM NaCl (pH 7.5) and run at 0.75 ml/min, was used to estimate the molecular mass and oligomeric state of the purified BclA protein. A calibration curve for molecular size estimation was generated by individually loading cytochrome *c*, myoglobin, ovalbumin and BSA.

Further characterization of the BclA lectin was made using DLS (dynamic light scattering) and DSC (differential scanning calorimetry). DSC experiments were performed on a VP-DSC differential scanning micro-calorimeter (Microcal) at a BclA concentration of 1.1 mM and scan rate of 1.0 K/min in 20 mM Tris/HCl, 100 mM NaCl and 100 μM $CaCl_2$ (pH 7.5). A background scan collected with buffer in both cells was subtracted from each scan. The DSC data were fitted to a non-two-state unfolding model, which allows estimation of the calorimetric heat change (ΔH) and the van't Hoff heat change (ΔHv) of the thermal unfolding transition. DLS measurements were carried out with a Zetasizer Nano-ZS machine (Malvern Instruments) using DTS software for data acquisition and analysis. Five different concentrations of BclA ranging from 0.4 to 2 mg/ml in 20 mM Tris/HCl, 100 mM NaCl and 100 μM $CaCl_2$ (pH 7.5) were prepared, and size distribution of the particles by intensity and volume were measured.

SPR (surface plasmon resonance)

All SPR experiments were performed using a Biacore 3000 instrument at 25 °C using HBS (Hepes-buffered saline: 10 mM Hepes and 150 mM NaCl, pH 7.4) containing 0.005 % (v/v) Tween 20 and a flow rate of 5 μl/min. Two experimental arrangements were used to evaluate the sugar preference of the lectin.

© The Authors Journal compilation © 2008 Biochemical Society

In the first experimental set-up, measurements were carried out simultaneously on three measuring channels using different immobilized sugars: D-galactose, D-mannose and L-fucose. Biot–PAA (biotinylated polyacrylamide) probes bearing sugar moieties (Lectinity Corp) were trapped on a CM5 sensor chip that was coated with streptavidin using the standard procedure. Each Biot–PAA–monosaccharide (50 μl at concentration 200 μg/ml) was injected to the selected channel. Direct binding curves of the lectin to immobilized sugars were measured over the concentration range 0.06–1 mg/ml. For the inhibition studies, BclA was diluted in running buffer to 200 μg/ml and mixed with 10 × concentrated carbohydrate competitor in a 9:1 ratio. This mixture (20 μl) was injected on to the CM5 chip at a flow rate of 5 μl/min. The CM5 chip was regenerated using 100 mM EDTA. Binding was measured as RU (resonance units) over time, and data were evaluated using the BIAevaluation Software, version 4.1. As BclA does not bind to D-galactose, the corresponding channel was used as a blank. Data were fitted using a kinetic model for 1:1 binding. For IC_{50} evaluation, Req (RU corresponding to steady-state equilibrium), calculated from fitted curves, were taken as the amount of BclA bound to the sugar surface in the presence of competing monosaccharide.

In the second experimental set-up, direct measurements were performed with BclA immobilized on the carboxymethyl-dextran surface of the CM5 chip. Similarly to the streptavidin immobilization, the chip was activated with EDC (*N*-ethyl-*N*′-[3-dimethylaminopropyl]carbodi-imide)/NHS (*N*-hydroxy-succinimide) solution and BclA in 10 mM acetate buffer (pH 4.5) was injected into the flow channel. Finally, the sensor surface was blocked with 1 M ethanolamine. The blank channel was treated similarly, except for the lectin injection.

Carbohydrate solutions (20 μl with concentrations between 0.97 and 500 μM) in running buffer were injected into the flow cells using the kinject mode. As the association and dissociation kinetics of monosaccharides was fast, further regeneration of the chip surface was not necessary. The equilibrium response (after subtraction from the response of the reference surface) of each experiment was used to create curves of analyte binding, which were both evaluated using Scatchard plots and fitted to a 1:1 steady-state affinity model using Origin 7.0 software (OriginLab Corp.).

Glycan microarray analysis

BclA was labelled with Alexa Fluor® 488–TFP (tetrafluorophenyl) (Invitrogen) according to the manufacturer's instructions and purified on a D-Salt polyacrylamide desalting column (Pierce). Alexa Fluor® 488-labelled BclA was used to probe the printed glycan arrays [17] following the standard procedure of Core H of the Consortium for Functional Glycomics (http://www.functionalglycomics.org/).

Microcalorimetry

Titration calorimetry experiments were performed using a Microcal VP-ITC microcalorimeter. All titrations were carried out in 0.1 M Tris/HCl buffer (pH 7.5) containing 3 μM $CaCl_2$ at 25 °C. Aliquots of 10 μl of each carbohydrate at various concentrations from 1.5 to 3 mM, dissolved in the same buffer, were added at 5 min intervals to the lectin solution present in the calorimeter cell. In the titrations, the protein concentration in the cell varied from 0.15 to 0.4 mM. The temperature of the cell was controlled to 25 ± 0.1 °C. Control experiments performed by injection of buffer into the protein solution yielded insignificant heats of dilution. Integrated heat effects were analysed by non-linear regression using a single-site binding model (Origin 7.0). Fitted data yielded the association constant (K_a) and the enthalpy of binding (ΔH). Other thermodynamic parameters, i.e. changes in free energy, ΔG, and entropy, ΔS, were calculated from the equation:

$$\Delta G = \Delta \mathrm{H} - T\Delta S = -RT \ln K_a$$

where T is the absolute temperature and $R = 8.314\ \mathrm{J \cdot mol^{-1} \cdot K^{-1}}$. Two to three independent titrations were performed for each ligand tested.

Crystallization and data collection

Freeze-dried protein was dissolved in water (15 mg/ml) in the presence of methyl α-D-mannoside (400 μg/ml). Initial crystallization conditions were screened using the complete Hampton Screens I and II (Hampton Research). Crystals in the form of thin plates appeared after several days in the presence of ammonium sulfate and PEG [poly(ethylene glycol)] 3350, 5000MME (monomethyl ether) or 8000. Crystals were cryo-cooled at 100 K after soaking them for as short a time as possible in 15 % (v/v) glycerol mixed with precipitant solution. One such crystal was used to obtain a first set of diffraction data collected to 1.95 Å (1 Å = 0.1 nm) resolution. It belongs to space group C2 with cell dimensions $a = 153.7$ Å, $b = 49.6$ Å and $c = 114.3$ Å, and $\beta = 125.9°$.

Optimization of the initial crystallization conditions yielded crystals resembling small triangles from 2.5 μl of BclA (14 mg/ml) solution in complex with methyl α-D-mannoside (700 μg/ml) mixed with 2.5 μl of reservoir solution consisting of 30 % PEG 8000 and 0.2 M ammonium sulfate. This crystal form belongs to space group $C222_1$ with cell dimensions $a = 50.0$ Å, $b = 185.7$ Å and $c = 187.5$ Å. After seeding, one thicker crystal was obtained and was cryo-cooled at 100 K after soaking for as short a time as possible in 20 % (v/v) glycerol mixed with precipitant solution. Diffraction data to 1.7 Å resolution were obtained.

Both data sets were collected at ESRF (European Synchrotron Radiation Facility), Grenoble, France, on stations ID14–2 and ID14–3 using an ADSC Q4R CCD (charge-coupled device) detector (Quantum Corp.) Diffraction images were integrated using MOSFLM [18] and scaled and converted into structure factors using the CCP4 program suite [19]. Statistics for the two data collections are shown in Table 1.

Structure solution and refinement

A monomer of PA-IIL (PDB code 1GZT, with removal of ligands and water molecules) was used as a probe for molecular replacement for the monoclinic crystal form of BclA using the program MOLREP [20]. Five monomers were located, and, after an initial refinement cycle, the electron-density maps of this complex showed glycerol instead of methyl α-D-mannoside in the binding site. This structure was not refined further, but was used as the search probe for the orthorhombic crystal, again giving five monomers in the asymmetric unit. After refinement, the electron-density maps of this complex showed clear features corresponding to one methyl α-D-mannoside molecule and two Ca^{2+} ions per monomer.

Crystallographic refinement was carried out with the program REFMAC [21] and manual model building was achieved using Coot [22]. Alternative conformations were constructed where necessary (with occupancies estimated from the refined relative *B*-factors of the conformations). The final model consists of all 630 residues with five sugar molecules, ten Ca^{2+} ions, four sulfate ions, zero glycerol molecules and 1164 water molecules with an R_{crys} of 16.9 % and R_{free} of 20.3 % to 1.7 Å resolution.

© The Authors Journal compilation © 2008 Biochemical Society

Table 1 Data collection and refinement statistics

Values in parentheses refer to the highest resolution shell. $R_{merge} = \Sigma|I - \langle I \rangle|/\Sigma\langle I \rangle$, where I = observed intensity. Geometry analyses were performed using PROCHECK [31] and WHATIF [32]. RMSD, root mean square deviation.

Parameter	Value	
	Monoclinic form	Orthorombic form
Data collection		
Beamline	ID14-2	ID14-3
Wavelength (Å)	0.933	0.934
Resolution (Å)	31.81–1.95 (2.00–1.95)	33.63–1.70 (1.74–1.70)
Space group	C2	$C222_1$
Cell dimensions:		
a, b, c (Å)	153.7, 49.6, 114.3	50.0, 185.7, 187.5
β (°)	125.9	
Measured reflections	180948	702443
Unique reflections	50938 (3737)	96348 (6680)
Average multiplicity	3.6 (3.6)	7.3 (7.1)
Completeness (%)	99.5 (99.3)	99.9 (99.7)
M $I/\sigma(I)$	13.6 (2.8)	19.1 (3.5)
R_{merge} (%)	9.6 (48.1)	7.3 (43.3)
Refinement		
Resolution range (Å)		33.63–1.70
R_{crys} (observation)		0.169 (0.242)
R_{free} (observation)		0.203 (0.294)
Average B_{iso} (Å)		
Chain/protein/sugar/Ca^{2+}		A/15.9/12.7/10.9
		B/15.1/16.0/11.5
		C/15.2/14.1/12.0
		D/14.4/10.5/9.2
		E/15.1/13.4/10.8
Water/sulfate		32.5/34.3
RMSD from ideality		
Bonds (Å)		0.015
Angles (°)		1.58
Ramachandran plot statistics		
Residues in favoured regions (%)		99.6
Residues in disallowed regions (%)		0.4 (Asp^2 of E and D)

Co-ordinates have been deposited at the Protein Data Bank under code 2VNV. Molecular drawings were prepared using Molscript [23].

RESULTS

Occurrence of lecB-like gene in *Burkholderia* genomes

The *Burkholderia* genus comprises more than 40 *Burkholderia* species [24], 12 of which have had their genome sequenced or are in the process of being sequenced. Among them, the Bcc is represented by several strains of *B. cenocepacia* and others. A similarity search based on the *lecB* sequence allowed the identification of genes coding for related proteins in several of these genomes, all of them corresponding to species both belonging to the Bcc and reported in CF isolates (see Supplementary Table S1 at http://www.BiochemJ.org/bj/411/bj4110307add.htm), namely *B. cenocepacia*, *B. cepacia*, *Burkholderia ambifaria* and *Burkholderia dolosa*. However, no *lecB*-like sequence could be found in the completed genomes of *Burkholderia thailandensis*, a soil inhabitant, or of *Burkholderia mallei* or *Burkholderia pseudomallei*, which are primary pathogens that cause glanders in animals and melioidosis in humans respectively. Surprisingly, in all genomes from Bcc species where it was identified, the *lecB*-like gene does not appear as a single copy, as previously observed in *Ps. aeruginosa*, *R. solanacearum* and *C. violaceum*, but at the 3′-terminus of several genes coding for proteins of various sizes (Supplementary Table S1).

The *B. cenocepacia* J2315 genome sequence consists of three chromosomal replicons and a small plasmid. Three *lecB*-related genes, with sizes of 384, 732 and 816 bp, are present in the second chromosome, and the corresponding gene products are referred to as BclA (129 amino acids without the initiating methionine residue), BclB (244 amino acids) and BclC (272 amino acids) (Figure 1, top panel). A fourth gene with sequence similarities to *lecB* was identified on the third chromosome. This last gene is invalidated due to a frameshift, but in the genome of other *Burkholderia* species such as *B. cepacia* R18194-ATCC17660 (DOE) or *Burkholderia* sp. SAR-1, it could code for a protein of 289 or 296 amino acids respectively.

In the present study, we focused on the analysis of BclA, the shortest protein encoded by a *lecB*-like gene in *B. cenocepacia*. This gene was identified in five strains of *B. cenocepacia* and one strain of *B. ambifaria*, as well as an environmental sample from the Sargasso sea that is likely to contain *B.* (*ceno*)*cepacia*. An alignment of the deduced amino acid sequences is shown in Figure 1(B). All *Burkholderia* BclA protein sequences are very similar, having 32 % identity with PA-IIL. The amino acids involved in the binding of the two Ca^{2+} ions are conserved, as well as the C-terminal glycine residue that is involved in the architecture of the binding site through dimer formation. However, the BclA sequences are longer than the PA-IIL one, with seven additional amino acids at the N-terminus and a nine amino acid insertion in the region corresponding to Asn^{85} of PA-IIL (all sequence numbering is given without initial methionine residues). The amino acids involved in sugar binding are partly conserved with a specificity loop consisting of Ala^{29}-Ala^{30}-Glu^{31} in BclA compared with Ser^{22}-Ser^{23}-Gly^{24} in PA-IIL. The higher affinity for mannose over fucose has been previously mainly attributed to the presence of an alanine residue instead of a serine residue at the first position of the specificity loop triplet in RS-IIL from *R. solanacearum* [15], and BclA would therefore be predicted to display some preference for mannose.

Identification of BclA in *B. cenocepacia*

Expression of native BclA lectin in *B. cenocepacia* strain J2315 was tested under different growing conditions, including cultivation on minimal and rich media in the presence and absence of iron and various sugars. Finally, 4 days of cultivation on rich medium in the presence of D-mannose resulted in a reasonable amount of protein for MS analysis after purification. The results of a search using MASCOT based on the MALDI-MS and LC–MS/MS analyses of the prepared tryptic digest against the NCBI database indicated the presence of protein BcenP_01004462 (*B. cenocepacia* PC184) (gi 84353445), which is a BclA orthologue from another *B. cenocepacia* strain. The search found five protein fragments matching the protein sequence with a final score of 77. In addition, a sixth matching fragment with a signal close to noise could be detected in the spectrum. The same result was obtained with a search against a local database containing data from *B. cenocepacia* strain J2315 (data from The Sanger Institute, which are not included in the NCBI sequence database). The search identified the BCAM0186 protein (chromosome 2, residues 216624–217013) that corresponds to BclA. The peptides identified in the BclA sequence are shown in Figure 2.

Cloning of lecB-like gene, production of recombinant BclA protein in *E. coli* and characterization

Using the expression plasmid *pRSETbcla*, BclA was expressed in *E. coli* BL21(DE3) with a typical yield of approx. 40 mg of

© The Authors Journal compilation © 2008 Biochemical Society

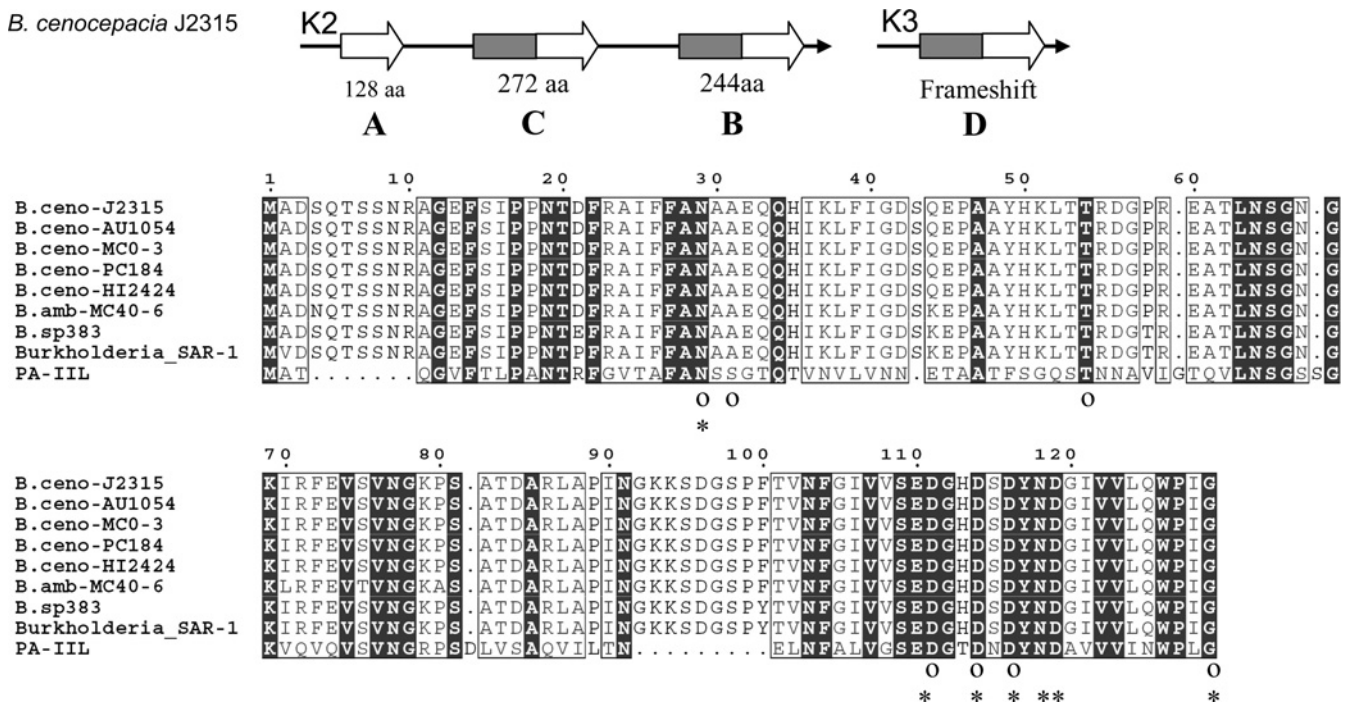


Figure 1 PA-IIL-like lectins in bacteria of the Bcc

Top panel: schematic representation of PA-IIL-like containing proteins in *B. cenocepacia*. aa, amino acids. The grey shaded portions of the arrows indicate the predicted lectin domains. Lower panels: alignment of BclA sequences from different strains with PA-IIL sequences. Strictly conserved amino acids are shaded black. Amino acids of PA-IIL involved in Ca^{2+} binding are labelled with *, and those involved in sugar binding are marked with o. Dots in the sequence indicate gaps inserted to aid alignment. B.amb, *B. ambifaria*; B.ceno, *B. cenocepacia*; B.sp, *Burkholderia* sp.

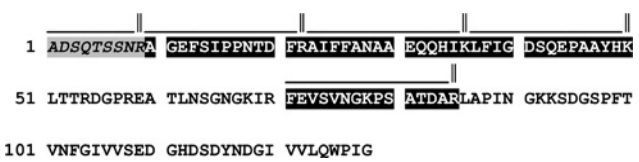


Figure 2 Peptides identified in the tryptic digest of a native protein isolated from *B. cenocepacia* strain J2315, which corresponds to the BclA lectin using MALDI-MS and LC–MS/MS analyses

Five fragments directly identified by the MASCOT search are represented on a black background with a double bar indicating the last amino acid. The sixth fragment close to signal noise is represented with grey background.

the purified protein per litre of culture. The resulting protein was purified by affinity chromatography on D-mannose–agarose and ran as a sharp band of approx. 14 kDa on SDS/PAGE. MS analysis confirmed the molecular mass of 13764 Da, corresponding to the BclA amino acid sequence lacking the initial methionine residue. In contrast with the other PA-IIL-like proteins, size-exclusion chromatography clearly demonstrated that BclA is a dimer in solution (Figure 3).

In order to characterize further the lectin, DLS and DSC experiments were performed (see the Supplementary material at http://www.BiochemJ.org/bj/411/bj4110307add.htm). DLS showed the hydrodynamic size of the lectin particle to be 5 nm. This corresponds to the overall globular shape of the dimeric protein as observed in the crystal structure (see below). The thermal unfolding transition for BclA was measured using DSC to be at 87 ± 0.5 °C with a calorimetric heat change (ΔH) to van't Hoff heat change (ΔHv) ratio of approx. 0.5 (concentration normalization was carried out on a per mol of monomer basis). This also confirms the presence of two unfolding units per mol of protein.

Carbohydrate specificity of BclA

An SPR binding assay was used to determine the affinity of BclA to sugars immobilized on the chip. The results show that, although BclA does not interact with fucosylated and galactosylated surfaces, it does bind strongly to the mannosylated polyacrylamide surface. Figure 4 (top, left-hand panel) shows typical sensorgrams obtained after injection of different concentrations of BclA. The association phase is rapid, but the dissociation is very limited because of the density of mannoside available for binding. The shape of the sensorgrams suggests that the lectin probably undergoes some rebinding phenomenon. The use of EDTA allows a rapid and total removal of the lectin for regeneration of the surface.

The lectin preference for different monosaccharides or monosaccharide derivatives was tested by their inhibition power on the binding of circulating BclA to the mannose surface. Figure 4 (top, right-hand panel) shows the sensorgrams obtained in the presence of increasing concentrations of methyl α-D-mannoside pre-incubated with BclA before injection. A concentration of 30 μM significantly inhibited the binding. The experimental binding curves obtained in the presence of competing sugar were fitted using a kinetic model for 1:1 binding (see the Supplementary material). Req were calculated from fitted curves and taken as the amount of BclA bound to the chip surface in the presence of competing sugar. The IC_{50} values for the different compounds were derived from such competition experiments.

The results show that D-mannose is an excellent inhibitor with an IC_{50} of 39 μM (Table 2), whereas D-galactose does not bind

© The Authors Journal compilation © 2008 Biochemical Society

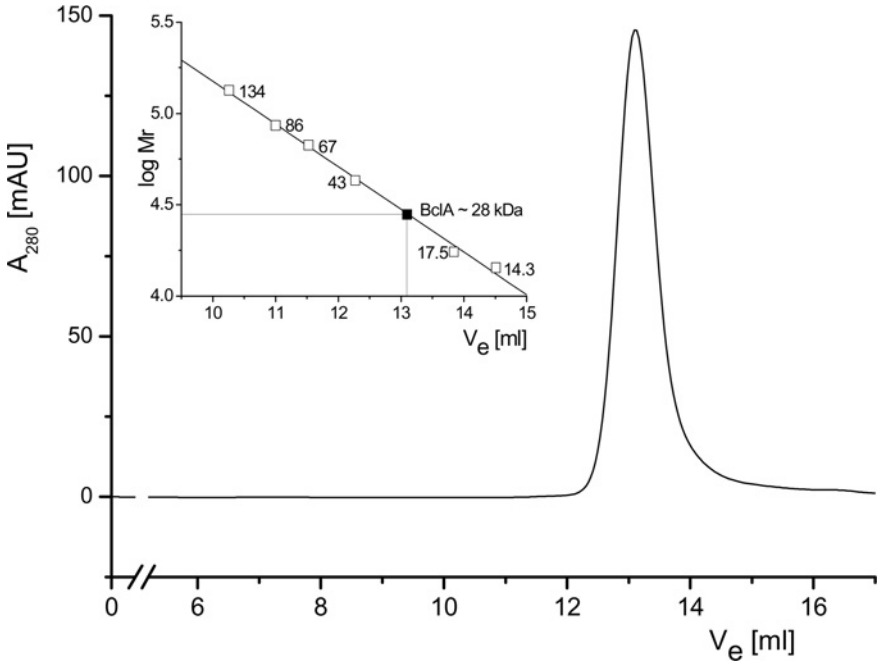


Figure 3 Size-exclusion chromatogram of BclA on Superose 12 column

Using molecular-mass standards under the same running conditions (inset), the molecular mass of the BclA protein was estimated to be 28 kDa, corresponding to a dimer in solution. V_e, elution volume.

significantly to the lectin. Some inhibition effects are shown by the monosaccharides of the L-galacto series, such as L-fucose, D-arabinose and L-galactose, but are two orders of magnitude weaker than D-mannose, indicating that BclA has a high specificity for mannose. As classically observed in lectin–carbohydrate interactions, the BclA lectin has higher affinity towards mannose derivatives with a hydrophobic aglycone group such as methyl (methyl α-D-mannoside) or phenyl (benzyl α-D-mannoside), with IC_{50} values approx. 3-fold smaller.

To take into account the accessibility of the BclA-binding site, which could be modified when binding to the polymeric sugar on the surface, SPR experiments were also conducted by evaluating the direct binding of ten monosaccharides (benzyl α-D-mannoside, methyl α-D-mannoside, D-mannose, L-fucose, D-arabinose, methyl α-L-fucoside, D-galactose, L-galactose, *N*-acetyl-D-glucosamine and *N*-acetyl-D-galactosamine) to BclA immobilized on the surface. The sugars were tested at two different concentrations: 500 μM and 25 mM. Although the mannose and mannosylated derivatives already showed site saturation at 500 μM (see the Supplementary material), for all of the other sugars, only partial saturation was observed (at 25 mM: L-fucose, 86 %; D-arabinose, 73 %; methyl α-L-fucoside, 45 %; L-galactose, 33 %; and *N*-acetyl-D-galactosamine and *N*-acetyl-D-glucosamine, <10 %). Such results indicate that dissociation constants for these other tested sugars are weaker than millimolar values.

Since oligosaccharide specificity could be correlated with the possible function of the bacterial lectin in tissue recognition, binding to a broad spectrum glycan array was performed at the Consortium for Functional Genomics. The screening results show that BclA has a strict specificity for oligomannose-type N-glycans, only binding to oligosaccharides with an α-mannoside-capped branch (Figure 5). The terminal disaccharides that are bound are α-Man1–2Man, αMan1–3Man and αMan1–6Man, indicating that the specificity for the linkage is not strict. The lectin can also tolerate xylose substitution on the non-terminal mannose. Hybrid structures with the presence of *N*-acetylglucosamine on one antenna are also recognized if terminal α-mannoside residues are available on the second antenna.

Thermodynamics of binding

The three best ligands, i.e. mannose and its methyl and benzyl derivatives, were used to characterize further the interaction between BclA and its ligands by titration microcalorimetry, a method that is well suited for the characterization of protein–carbohydrate interactions [25]. A typical titration curve for methyl α-D-mannoside binding to BclA is shown in Figure 6, upper panel. Injection of the ligand yielded strong exothermic peaks in the first phase until reaching a stoichiometry close to 1. Fitting of the integrated peak values to theoretical models yielded the curves shown in Figure 6, lower panel. The fitting is not perfect for a classical one-site model, because of a clear increase in the exothermic reaction during the first ten injections. Such a phenomenon can only be explained by the existence of several binding sites that either have different stoichiometry or have binding co-operativity. Since gel-exclusion experiments demonstrated that BclA is a homodimer in solution, it is possible that binding of mannoside at one site increases the affinity of the other site. Indeed, a fitting procedure using a co-operative two-site model gave excellent agreement with the experimental data (Figure 6, lower panel).

Data obtained by averaging over three experiments for mannose and its derivatives are listed in Table 3. The co-operative binding models indicate the presence of a high-affinity binding site with micromolar affinity that facilitates binding to a second site, with slightly lower affinity (7–12 μM depending on the ligand), but with higher enthalpy. Since both sites do not have very different characteristics, the one-site model gives an 'apparent' thermodynamic behaviour that reflects the overall binding behaviour. For

© The Authors Journal compilation © 2008 Biochemical Society

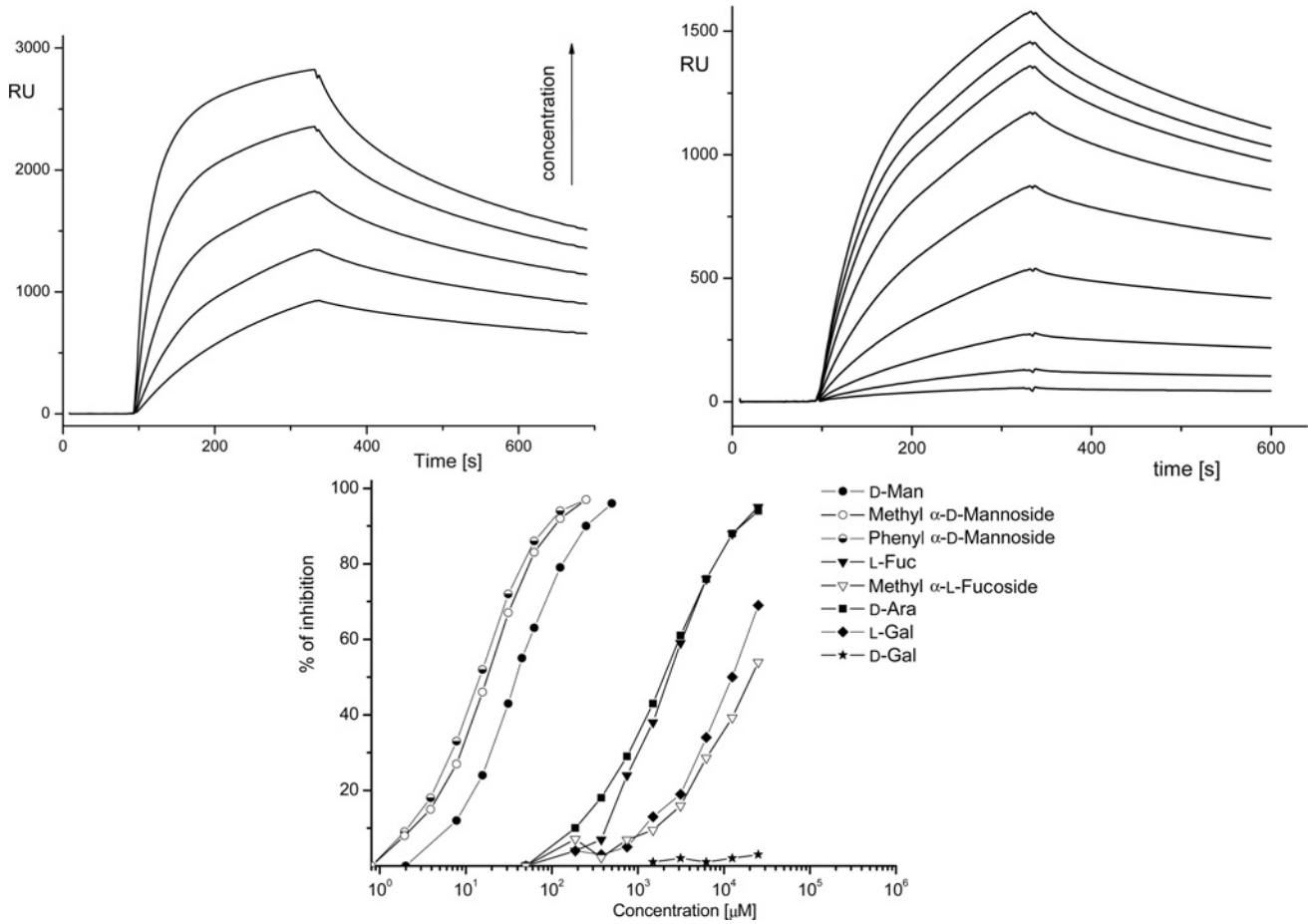


Figure 4 SPR sensorgrams representing binding of BclA on polymeric mannosylated chips

Top left-hand panel: concentration-dependent binding of BclA to the chips from 0.0625 mg/ml (lowest curve) to 0.125, 0.25, 0.5 and 1 mg/ml (highest curve). Top right-hand panel: inhibition curves obtained from different concentration of methyl α-D-mannoside from 0 μM (highest curve) to 1.95, 3.9, 7.8, 15.6, 31.25, 62.6, 125 and 250 μM (lowest curve). Lower panel: percentage of inhibition measured for different monosaccharides and sugar derivatives.

Table 2 IC_{50} values obtained by SPR for the inhibition of binding of BclA to polymeric mannose surface

ND, not determined.

Inhibitor	IC_{50} (μM)	Potency
αMeMan	18.2	2.1
αBenzMan	14.4	2.7
D-Mannose	39.0	1
D-Arabinose	2100	0.018
L-Fucose	2300	0.017
L-Galactose	12 600	0.003
αMeFuc	20 900	0.001
D-Galactose	ND	–

the one-site model, BclA displays a very high affinity for mannose with a dissociation constant of 5 μM, i.e. a much stronger affinity than classically observed for plant or animal lectins. The methyl or benzyl derivatives display twice the affinity, with a K_d of 2.75 μM for the interaction of BclA with methyl α-D-mannoside. This micromolar affinity is due both to a strong enthalpy of binding with ΔH values ranging between -23 and -27 kJ/mol and to a favourable entropy of binding with $T\Delta S$ between 3 and 9 kJ/mol. The more favourable entropy displayed by the mannose derivatives may be due to the limitation of flexibility by blocking the anomeric equilibrium. This favourable entropy of binding is unusual in protein–carbohydrate interactions, but has been reported in this Ca^{2+}-dependent family of bacterial lectins and is attributable to the involvement of the two Ca^{2+} ions in sugar binding [26,27].

Crystal structure of BclA complexed with methyl α-D-mannoside

BclA was crystallized in presence of methyl α-D-mannoside, and crystals were obtained after a few days in the presence of ammonium sulfate and PEG, in both monoclinic C2 and orthorhombic $C222_1$ space groups. The X-ray structure of the C2 form was determined by molecular replacement using the PA-IIL structure as the probe, but was not refined further as no ligand was observed in the binding site. The $C222_1$ form structure, diffracting to 1.7 Å resolution, was in turn determined using the C2 form co-ordinates from the first molecular replacement (Table 1). The asymmetric unit contains five BclA monomers, corresponding to two dimers (on a pseudo 2-fold axis) and a half dimer (forming a dimer by a crystallographic 2-fold axis). The N-terminal extremities exhibit disorder for the first eight amino acids of chain A and the first residue of chain C that could not therefore

© The Authors Journal compilation © 2008 Biochemical Society

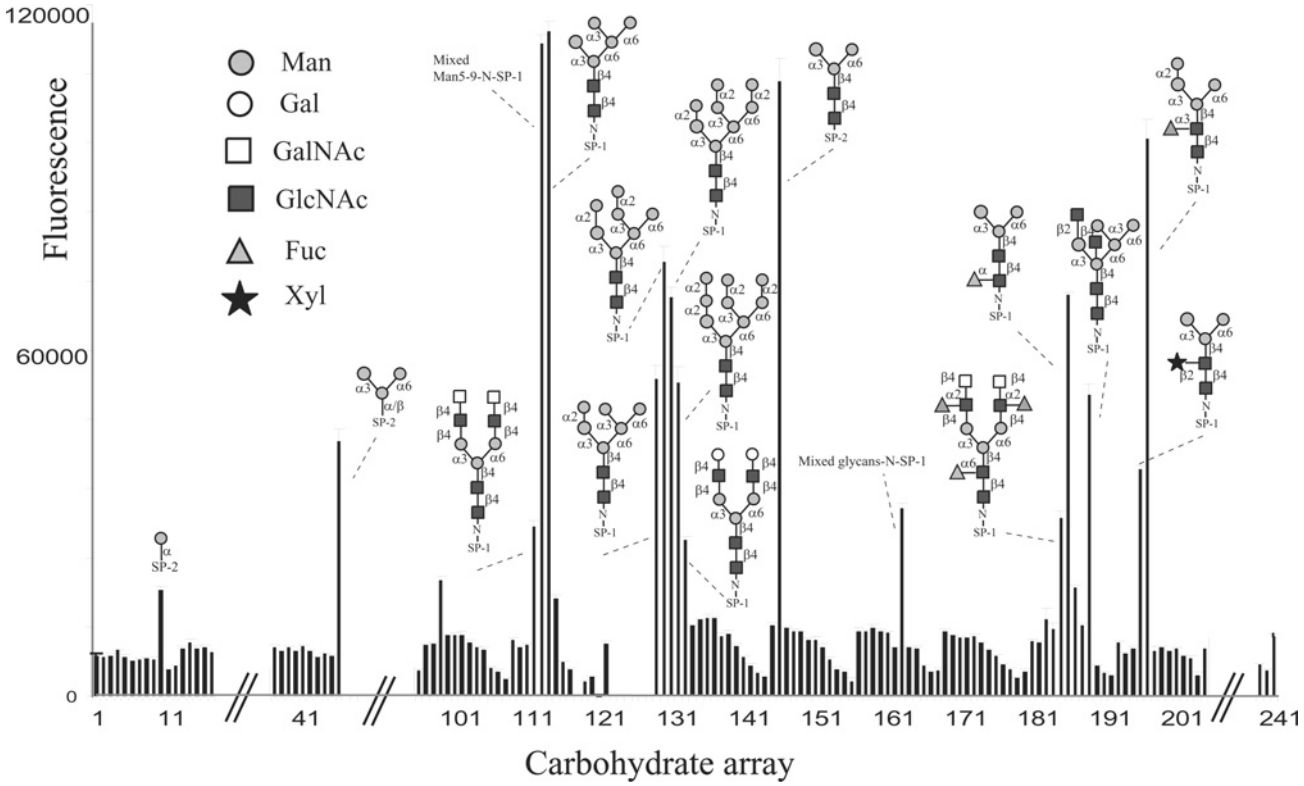


Figure 5 Glycan array analysis of BclA as measured by fluorescence intensity

All glycans yielding to a signal/noise ratio larger than 1.5 are represented. The omitted area corresponds to series of glycans with no significant labelling. Complete results including a comprehensive oligosaccharide list (Plate Array Version 3) are available from the Consortium of Functional Glycomics (http://www.functionalglycomics.org/). Spacers are as follows: SP-1, -O(CH2)2NH(O)(CH2)5NHC(O)(CH2)5NHC(O)-biotin; SP-2, -O(CH2)3NH-CO(CH2)5NH-biotin.

Table 3 Microcalorimetry data

All values are averaged over three experiments (two for the mannose complex). Errors due to the fitting procedure are less than 10% for the one-site model and less than 15% for the co-operative two-site model. S.D.s when averaging are below these values.

(a) Co-operative two-site model (per dimer)

	D-Mannose	αMeMan	αBzMan
K_{a1}/K_{a2} (10^3 M^{-1})	300/80.6	525/118	618/141
K_{d1}/K_{d2} (10^{-6} M)	3.3/12.4	2.7/7.8	1.6/7.1
$-\Delta G_1/-\Delta G_2$ (kJ/mol)	−31.3/−28.0	−32.7/−29.0	−33.0/−29.4
$-\Delta H_1/-\Delta H_2$ (kJ/mol)	−26.0/−31.8	−20.8/−28.4	−22.9/−28.5
$T\Delta S_1/T\Delta S_2$ (kJ/mol)	5.3/−3.8	11.9/0.6	10.1/0.9

(b) One-site model (per monomer)

	D-Mannose	αMeMan	αBzMan
K_a (10^3 M^{-1})	194	364	361
K_d (10^{-6} M)	5.15	2.75	2.80
n (stoichiometry)	0.91	0.83	0.86
$-\Delta G$ (kJ/mol)	30.2	31.8	31.7
$-\Delta H$ (kJ/mol)	27.0	23.0	24.2
$T\Delta S$ (kJ/mol)	3.2	8.8	7.5

be included in the final model. The overall topologies of BclA and PA-IIL are very similar, with a nine-stranded antiparallel β-sandwich consisting of two sheets of four and five strands respectively. Dimerization is created by strong complementary packing of the hydrophobic curved β-sheet to its counterpart and via sheet extension (Figure 7A). This is the first description of a dimeric form for a bacterial lectin that is dependent on two Ca^{2+} ions, with all structures reported previously adopting a tetrameric organization.

Each BclA monomer is associated with two Ca^{2+} ions and one mannose residue, the electron density of which is shown in Figure 7(B). The two ions are located close together, 3.73 ($\pm$0.2) Å apart, with a classical seven-ligand co-ordination, mainly involving one loop via the side-chains of Glu^{109}, Asp^{113}, Asp^{115}, Asn^{117} and Asp^{118} (Figure 7C and see Supplementary Table S2 at http://www.BiochemJ.org/bj/411/bj4110307add.htm). A second loop is involved through the side-chain of Asn^{28}. In addition, the acidic group of the C-terminal Gly^{128} of the other monomer, which is involved in the head-to-tail interaction, also participates in the co-ordination of the Ca^{2+} ions.

The methyl α-D-mannoside molecule participates directly in the co-ordination of the Ca^{2+} ions with hydroxy groups O2, O3 and O4, in a similar fashion to the PA-IIL–D-mannose complex [28]. The O2, O3, O4 and O6 hydroxy groups are also involved in hydrogen bonds with amino acids Asn^{28}, Ala^{30}, Glu^{31}, Asp^{110}, Asp^{113}, Asp^{115} and Asp^{118}, and with the terminal carboxy group of Gly^{128} of the neighbouring monomer (Figure 7C, and see Supplementary Table S2). Two types of hydrophobic contact can be observed. First, the methyl group at the anomeric position is close to the methyl group of Ala^{30} and His^{112}. This histidine residue has no equivalent in the other related bacterial lectins and seems to play a special role. Analysis of the electron density indicates that His^{112} adopts a double conformation with approx. 50/50 occupancy in all of the five binding sites. In one orientation,

© The Authors Journal compilation © 2008 Biochemical Society

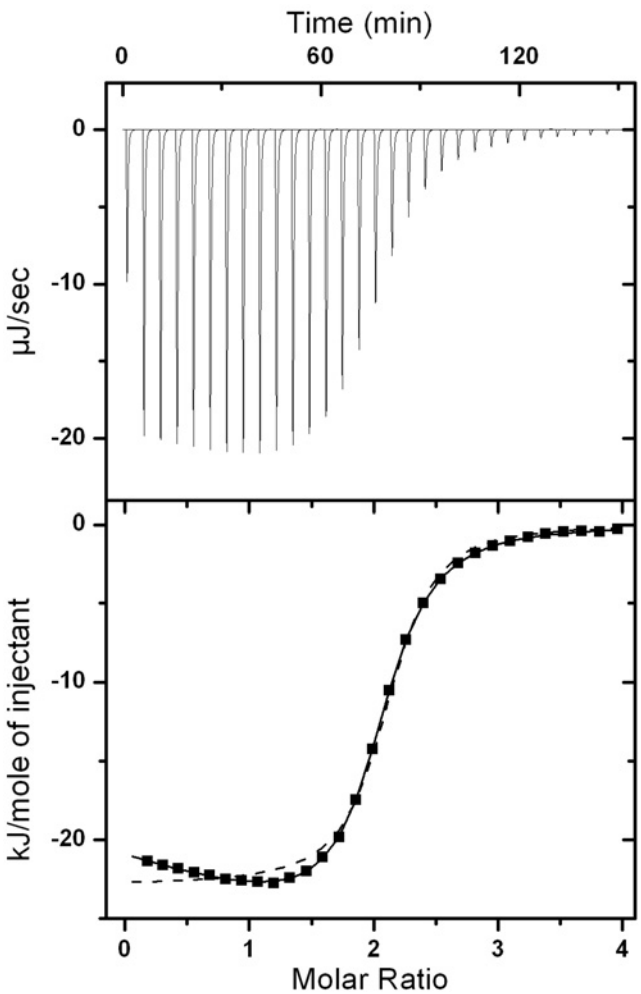


Figure 6 Titration microcalorimetry of methyl α-D-mannoside (3 mM) binding to BclA (0.39 mM) in 100 mM Tris/HCl buffer (pH 7.5) with 3 μM $CaCl_2$ at 25 °C

Upper panel: data for 30 automatic injections of 10 μl of methyl α-D-mannoside each into the BclA-containing cell. Lower panel: plot of the total heat released as a function of molar ratio for dimeric BclA (■). The broken line represents the best fit for a one-site model and the solid line is the best fit for a co-operative two-site model.

it creates hydrophobic contacts with the C5 and C6 atom of the D-mannose hydroxymethyl group. In this conformation, the ND1 nitrogen of His^{112} can give a hydrogen bond to the side-chain of Asp^{110}, while its CE1 creates hydrophobic contact with CG of Glu^{31}, therefore creating a rather 'closed' conformation of the binding site (Figure 7D). In the other conformation, the same histidine residue adopts a different side-chain orientation and makes hydrophobic contact with the hydrophobic faces of mannose (C3 and C5) and with the methyl group at C1. This is the first evidence of flexibility in the binding site of a Ca^{2+}-dependent bacterial lectin.

DISCUSSION

Novel quaternary arrangement in two Ca^{2+}-dependent bacterial lectins

The first crystal structure of a soluble lectin dependent upon two Ca^{2+} ions has been reported for PA-IIL from *Ps. aeruginosa* [13]. Later, two related structures have been described from similar proteins present in two opportunistic bacteria: RS-IIL from *R. solanacearum* [15] and CV-IIL from *C. violaceum* [14]. The three proteins have very similar sequences (60–70 % identity), with identical fold and tetrameric arrangements. The BclA protein belongs to the same family of proteins, but displays clear differences from the lectins mentioned above. The central β-sandwich monomer fold is conserved, but variations in loop conformation can be observed (Figure 8A). The most visible difference is the extension of the N-terminus, with seven additional amino acids for BclA for which electron density is clearly defined, except in monomer A. The loop between β-strands 7 and 8 is also much longer in BclA thanks to an insertion of nine amino acids.

The differences in peptide sequence also influence the oligomeric arrangement. BclA is a dimer in solution as well as in the crystal structure. Superposition of dimeric BclA on tetrameric PA-IIL (Figure 8B) shows that the additional N-terminal amino acids (particularly Arg^{9}) and the longer loop (particularly Pro^{98}) would not allow the creation of the dimer–dimer interface that is observed in the other structures of the family. A thorough examination of the contact region also reveals that bulky amino acids such as Arg^{85} create steric conflicts when trying to build a tetramer by homology modelling.

Structural basis for mannose specificity and possible biological functions

The three lectins characterized previously from opportunistic bacteria have a rather large range of specificity, each binding to several monosaccharides. Their main stereochemical requirement is the presence of one axial hydroxy group located next to two equatorial ones, a sequence that is present in L-galacto sugars (O4 is axial, and O3 and O2 are equatorial) such as L-galactose, L-fucose, D-arabinose and D-fructopyranose, but also in D-manno sugars (O2 is axial, and O3 and O4 are equatorial) such as D-mannose [12,15,29]. However, the specificity is correlated with the nature of a stretch of three amino acids located in the so-called 'specificity loop'. It has been established that the sequence Ser^{22}-Ser^{23}-Gly^{24} (in PA-IIL) does not favour the binding of mannose whereas the Ala-Ala-Asn series of RS-IIL does. A comparison showed that, in PA-IIL, Ser^{22} makes a hydrogen bond to the side-chain of Asp^{96}, preventing it from binding to O6 of mannose. We have demonstrated recently that, by mutating the specificity loop residues, the sugar specificity can be engineered, with the S22A PA-IIL mutant binding more strongly to mannose than to fucose [30].

The situation is very different in BclA since the protein has a strict preference for mannoside. The competition experiments performed using SPR indicate some inhibition power for fucose, but microcalorimetry experiments with the same monosaccharide and SPR experiments using direct binding of BclA to the fucosylated surface did not show any signal, indicating that the binding is at least 1000-fold weaker than for mannose. Indeed, the triplet in the specificity loop is Ala^{29}-Ala^{30}-Glu^{31} which therefore allows O6 of mannose to form a strong hydrogen bond to the side-chain of Asp^{110} (Figure 8D). Furthermore, it seems that fucose binding is excluded because His^{112}, only present in BclA, stacks favourably against C6 or C1 of mannose (depending upon side-chain orientation), but would collide with the O1 of α-fucose (Figure 8C).

BclA also has further differences with the other bacterial lectins when considering the interaction thermodynamics. All of the Ca^{2+}-dependent soluble bacterial lectins studied to date display micromolar affinity for monosaccharides [26], with a strong enthalpy of binding and a small favourable entropy, an unusual

© The Authors Journal compilation © 2008 Biochemical Society

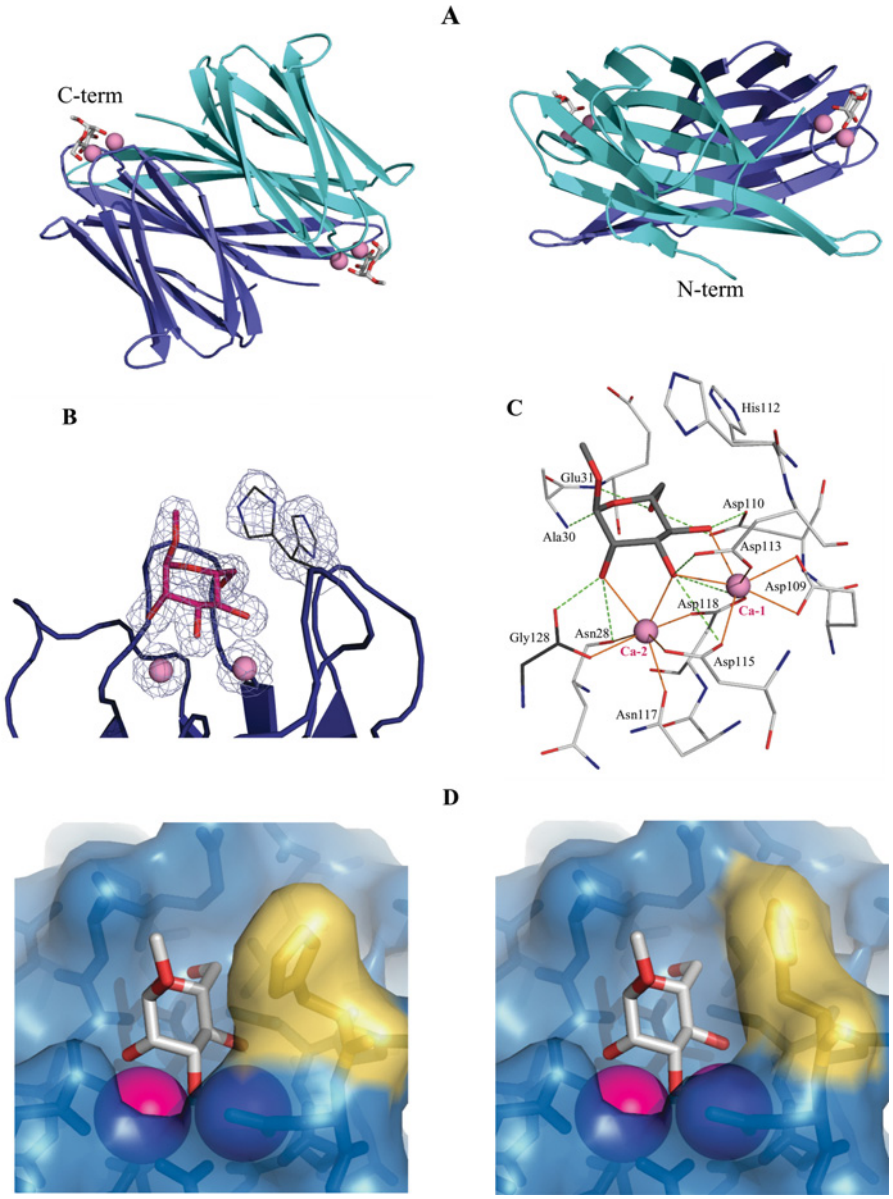


Figure 7 Crystal structure of BclA–methyl α-D-mannoside complex

(**A**) Orthographic representation of ribbon diagram of BclA dimer complexed with methyl α-D-mannoside (sticks) and Ca^{2+} (spheres). C-term, C-terminus; N-term, N-terminus. (**B**) Electron density (1σ = electron per $Å^2$) around the ligand and the two alternative conformations of His^{112}. (**C**) Details of the binding site with co-ordination bonds represented by solid pink lines and hydrogen bonds represented by broken green lines. Ca-1 and Ca-2 indicate the two Ca^{2+} ions. (**D**) Protein surface of the binding site generated by the two possible orientations of His^{112}.

behaviour for protein–carbohydrate interactions that has been correlated with the presence of Ca^{2+} [27]. All of the experiments could be fitted to a one-site model, which is in agreement with the crystal structures. BclA also displays strong affinity for mannoside, but it was not possible to fit the experimental data using a one-site model. In the case of BclA, only once co-operativity between the two binding sites of the monomers was included in the model was a good agreement between fitted and experimental curves obtained. This behaviour can be correlated with the unusual flexibility of His^{112} in the binding site and/or to the unusually long loop in the region of residues 90–100 that only exists in BclA. Indeed, this loop is located very close to the His^{112} loop of the other monomer and it can be postulated that binding of mannose in one site alters the dynamics of one loop which in turn affects the second binding site. A molecular dynamics study will be performed for a better understanding of this unusual co-operativity mechanism.

Conclusions

Strong interest has been shown in bacterial lectins recently because of their novel structure and carbohydrate-binding modes, their high affinity for carbohydrate ligands and their possible role in host recognition, in adhesion and in biofilm formation. A series of such lectins, all based on the sequence of the PA-IIL lectin, has been identified and subsequently studied from several bacteria. They all have a common fold and tetrameric

© The Authors Journal compilation © 2008 Biochemical Society

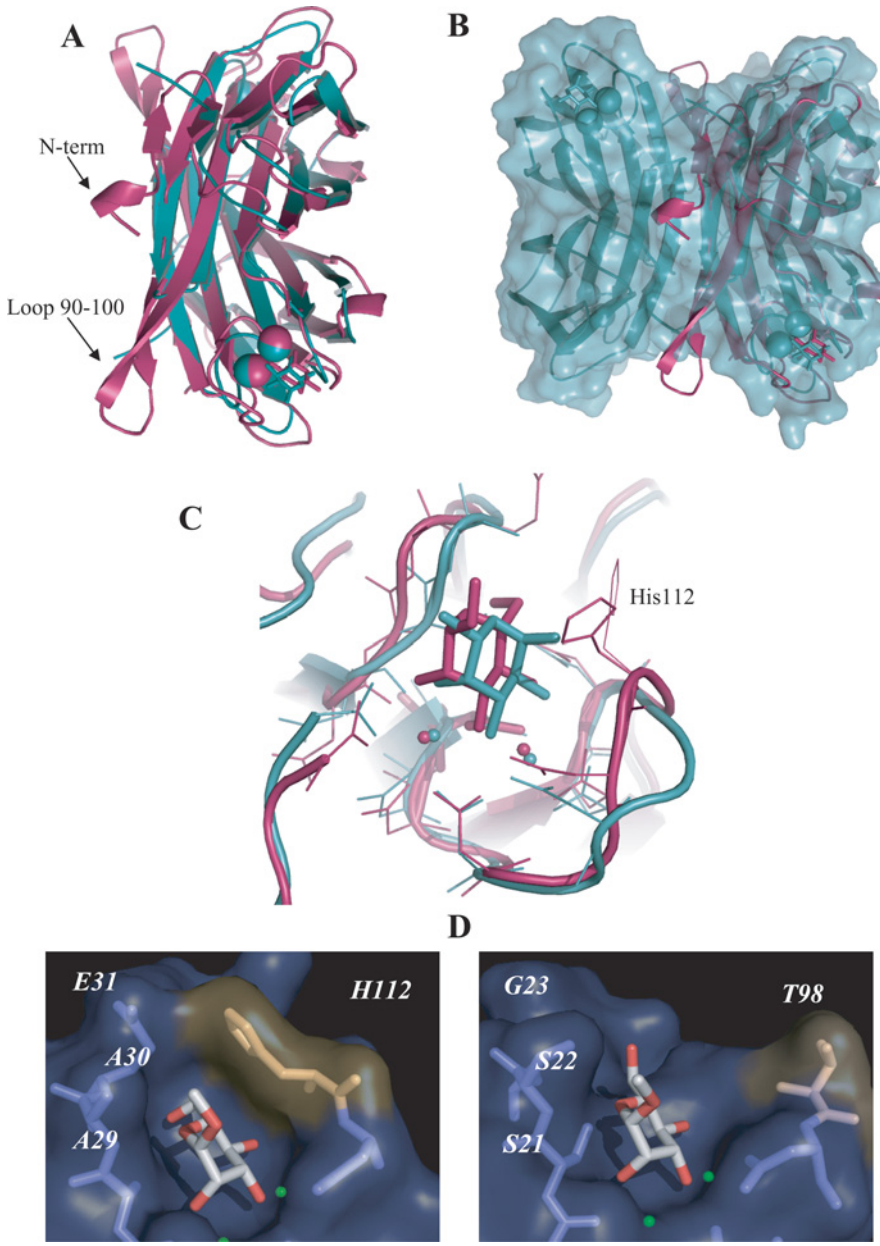


Figure 8 Comparison of crystal structures of PA-IIL–fucose from *Ps. aeruginosa* (PDB code 1GZT) and BclA–methyl α-D-mannoside from *B. cenocepacia*

(**A**) Superposition of the BclA dimer (magenta) and PA-IIL equivalent monomers (blue). N-term, N-terminus. (**B**) Superposition of the BclA dimer on the PA-IIL tetramer and its protein surface. (**C**) Superposition of the binding sites. (**D**) Comparison of the binding-site surface with His^{112} of BclA in the 'closed' conformation. Single-letter amino acid codes are used.

structure, but different sugar specificities while being dependent upon a special two-Ca^{2+} co-ordination. The surprising discovery of several such lectins also in the genome of *B. cenocepacia* led to the characterization (structurally and thermodynamically) of the smallest gene product, BclA. The role of the lectin in bacterial infection has yet to be elucidated, but the present work has yielded significant progress in the elucidation of structure, affinity and specificity of this new lectin.

The present study shows this lectin to have a core fold identical with that of PA-IIL, but tetramer formation is hindered because of the presence of additional loops. The BclA lectin also differs with an apparent novel co-operative binding across the dimer, compared with the purely single-site binding of PA-IIL and its family. BclA is indeed a member of the PA-IIL lectin family, dependent on two Ca^{2+} ions, with a core β-sheet structure, but is the first member to associate only as a dimer and to present site co-operativity. Its specificity is for mannoside sugars and the glycan array experiment shows specificity for oligomannose-type oligosaccharides that are present on human glycoproteins.

This research is part of the Ph.D. theses of E.L. and L.M., who are supported by Conseil Régional Rhône-Alpes (Cluster 10), and by the Ministry of Education (MSM0021622413, MSM0021622415, LC06030) and Grant Agency (GA303/06/570) of the Czech Republic. The glycan array resources were provided by the Consortium for Functional Glycomics grant number GM62116. Travel and visits between NCBR and CERMAV (Centre de Recherche sur les Macromolécules Végétales) are supported by a French–Czech bilateral

© The Authors Journal compilation © 2008 Biochemical Society

exchange programme BARRANDE (M.W. and A.I.) and financial help is gratefully acknowledged from the French cystic fibrosis association 'Vaincre la Mucoviscidose'. We thank Ms Catherine Gautier for excellent technical help. We thank the ESRF, Grenoble, for access to synchrotron data collection facilities.

REFERENCES

1 Mahenthiralingam, E. and Vandamme, P. (2005) Taxonomy and pathogenesis of the *Burkholderia cepacia* complex. Chron. Respir. Dis. **2**, 209–217
2 Chiarini, L., Bevivino, A., Dalmastri, C., Tabacchioni, S. and Visca, P. (2006) *Burkholderia cepacia* complex species: health hazards and biotechnological potential. Trends Microbiol. **14**, 277–286
3 Govan, J. R. and Deretic, V. (1996) Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. Microbiol. Rev. **60**, 539–574
4 Lyczak, J. B., Cannon, C. L. and Pier, G. B. (2002) Lung infections associated with cystic fibrosis. Clin. Microbiol. Rev. **15**, 194–222
5 Saiman, L. and Siegel, J. (2004) Infection control in cystic fibrosis. Clin. Microbiol. Rev. **17**, 57–71
6 Riedel, K., Hentzer, M., Geisenberger, O., Huber, B., Steidle, A., Wu, H., Hoiby, N., Givskov, M., Molin, S. and Eberl, L. (2001) *N*-acylhomoserine-lactone-mediated communication between *Pseudomonas aeruginosa* and *Burkholderia cepacia* in mixed biofilms. Microbiology **147**, 3249–3262
7 Eberl, L. and Tummler, B. (2004) *Pseudomonas aeruginosa* and *Burkholderia cepacia* in cystic fibrosis: genome evolution, interactions and adaptation. Int. J. Med. Microbiol. **294**, 123–131
8 Gilboa-Garber, N. (1982) *Pseudomonas aeruginosa* lectins. Methods Enzymol. **83**, 378–385
9 Imberty, A., Wimmerová, M., Sabin, C. and Mitchell, E. P. (2006) Structures and roles of *Pseudomonas aeruginosa* lectins. In Protein-Carbohydrate Interactions in Infectious Disease (Bewley, C., ed.), pp. 30–48, Royal Society of Chemistry, Cambridge
10 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 **151**, 1313–1323
11 Sudakevitz, D., Imberty, A. and Gilboa-Garber, N. (2002) Production, properties and specificity of a new bacterial L-fucose- and D-arabinose-binding lectin of the plant aggressive pathogen *Ralstonia solanacearum* and its comparison to related plant and microbial lectins. J. Biochem. (Tokyo) **132**, 353–358
12 Zinger-Yosovich, K., Sudakevitz, D., Imberty, A., Garber, N. C. and Gilboa-Garber, N. (2006) Production and properties of the native *Chromobacterium violaceum* fucose-binding lectin (CV-IIL) compared to homologous lectins of *Pseudomonas aeruginosa* (PA-IIL) and *Ralstonia solanacearum* (RS-IIL). Microbiology **152**, 457–463
13 Mitchell, E., Houles, C., Sudakevitz, D., Wimmerová, 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. Nat. Struct. Biol. **9**, 918–921
14 Pokorná, M., Cioci, G., Perret, S., Rebuffet, E., Kostlánová, N., Adam, J., Gilboa-Garber, N., Mitchell, E. P., Imberty, A. and Wimmerová, M. (2006) Unusual entropy driven affinity of *Chromobacterium violaceum* lectin CV-IIL towards fucose and mannose. Biochemistry **45**, 7501–7510
15 Sudakevitz, D., Kostlánová, N., Blatman-Jan, G., Mitchell, E. P., Lerrer, B., Wimmerová, M., Katcof, D. J., Imberty, A. and Gilboa-Garber, N. (2004) A new *Ralstonia solanacearum* high-affinity mannose-binding lectin RS-IIL structurally resembling the *Pseudomonas aeruginosa* fucose-specific lectin PA-IIL. Mol. Microbiol. **52**, 691–700
16 Govan, J. R., Brown, P. H., Maddison, J., Doherty, C. J., Nelson, J. W., Dodd, M., Greening, A. P. and Webb, A. K. (1993) Evidence for transmission of *Pseudomonas cepacia* by social contact in cystic fibrosis. Lancet **342**, 15–19
17 Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M. E., Alvarez, R., Bryan, M. C., Fazio, F., Calarese, D., Stevens, J. et al. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. Proc. Natl. Acad. Sci. U.S.A. **101**, 17033–17038
18 Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 and ESF-EAMCB Newsletter on Protein Crystallography, no. 26
19 Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr. Sect. D Biol. Crystallogr. **50**, 760–763
20 Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Cryst. **30**, 1022–1025
21 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect. D Biol. Crystallogr. **53**, 240–255
22 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. Sect. D Biol. Crystallogr. **60**, 2126–2132
23 Kraulis, P. (1991) Molscript: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. **24**, 946–950
24 Vandamme, P., Govan, J. and LiPuma, J. (2007) Diversity and role of *Burkholderia* spp. In *Burkholderia*: Molecular Biology and Genomics (Coenye, T. and Vandamme, P., eds.), pp. 1–28, Horizon Bioscience, Wymondham
25 Dam, T. K. and Brewer, C. F. (2002) Thermodynamic studies of lectin–carbohydrate interactions by isothermal titration calorimetry. Chem. Rev. **102**, 387–429
26 Imberty, A., Mitchell, E. P. and Wimmerová, M. (2005) Structural basis for high-affinity glycan recognition by bacterial and fungal lectins. Curr. Opin. Struct. Biol. **15**, 525–534
27 Mitchell, E. P., Sabin, C., Šnajdrová, L., Pokorná, M., Perret, S., Gautier, C., Hofr, C., Gilboa-Garber, N., Koca, J., Wimmerová, M. and Imberty, A. (2005) High-affinity fucose binding of *Pseudomonas aeruginosa* lectin PA-IIL: 1.0 Å resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches. Proteins: Struct. Funct. Bioinform. **58**, 735–748
28 Loris, R., Tielker, D., Jaeger, K.-E. and Wyns, L. (2003) Structural basis of carbohydrate recognition by the lectin LecB from *Pseudomonas aeruginosa*. J. Mol. Biol. **331**, 861–870
29 Sabin, C., Mitchell, E. P., Pokorná, M., Gautier, C., Utille, J.-P., Wimmerová, M. and Imberty, A. (2006) Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa*: thermodynamics data correlated with X-ray structures. FEBS Lett. **580**, 982–987
30 Adam, J., Pokorna, M., Sabin, C., Mitchell, E. P., Imberty, A. and Wimmerová, M. (2007) Engineering of PA-IIL lectin from *Pseudomonas aeruginosa*: unravelling the role of the specificity loop for sugar preference. BMC Struct. Biol. **7**, 36
31 Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291
32 Vriend, G. (1990) WHAT IF: a molecular modeling and drug design program. J. Mol. Graphics **8**, 52–56

Received 18 September 2007/20 December 2007; accepted 16 January 2008
Published as BJ Immediate Publication 16 January 2008, doi:10.1042/BJ20071276

© The Authors Journal compilation © 2008 Biochemical Society

(Article n°2)

Structural basis of the affinity for oligomannosides displayed by BclA, a *Burkholderia cenocepacia* soluble lectin

Emilie Lameignere, Annabelle Varrot, Frédéric Dubreuil, Michaela Wimmerova, Anne Imberty,

Article en préparation

Structural basis of the affinity for oligomannosides displayed by BclA, a *Burkholderia cenocepacia* soluble lectin.

Emilie Lameignere[1], Annabelle Varrot[1], Frédéric Dubreuil[1], Michaela Wimmerova[2], Anne Imberty[1*]

[1]*CERMAV-CNRS (affiliated with Université Joseph Fourier and belonging to ICMG), BP 53, F-38041, Grenoble, Cedex 09, France*

*** Author for correspondence**

Dr. Anne Imberty, CERMAV-CNRS (affiliated with Université Joseph Fourier), BP 53, F-38041, Grenoble, Cedex 09, France; Fax: 33-4-76547203 Email: anne.imberty@cermav.cnrs.fr

Running title: *B. cenocepacia* lectin, BclA, complexed with di- and tri-mannosides

Key words: bacterial lectin / *Burkholderia cenocepacia* / Burkholderia *cepacia* complex / mannose-binding lectin / *Pseudomonas aeruginosa.*

Abstract

The opportunistic pathogen *Burkholderia cenocepacia* contains three soluble carbohydrate-binding proteins, related to the fucose-binding lectin PA-IIL from *Pseudomanas aeruginosa*. All display a PA-IIL-like domain and two of them have an additional N-terminal domain that displays no sequence similarities with known proteins. Printed arrays screening performed on the shortest one, *B. cenocepacia* lectin A (BclA), demonstrated the strict specificity for oligomannose-type N-glycan structures (Lameignere, *et al.* 2008). The disaccharides αMan1-2Man, αMan1-3Man and αMan1-6Man and the trisaccharide, αMan1-3(αMan1-6)Man were tested by titration microcalorimetry in order to evaluate their affinity for BclA in solution and to characterize the thermodynamics of the binding. The crystal structures of BclA complexed with αMan1-3Man disaccharide and αMan1-3(αMan1-6)Man trisaccharide have been solved at 1.75 and 1.1 Å resolution respectively. In the BclA-Trisaccharide complex, the sugar is positioned on a symmetric axis and creates a bridge between two adjacent dimmers. AFM experiments permitted the visualization of the filament formed in solution by this type of interaction. .

Introduction

Because of their ability to specifically bind to complex carbohydrates on cell surfaces, lectins are proteins that are part of the strategies adopted by pathogens for binding and adhering to host tissues (Sharon 1996, Sharon 2006). Apart from well-defined toxins and fimbrial adhesins, several soluble lectins were recently characterized in opportunistic bacteria (Imberty and Varrot 2008). Among them, the PA-IIL family, first identified in *Pseudomonas aeruginosa* (Gilboa-Garber 1982), has been identified in other opportunistic gram-negative bacteria such as *Chromobacterium violaceum*, *Ralstonia solanacearum* and *Burkholderia cenocepacia* (Lameignere, *et al.* 2008, Pokorná, *et al.* 2006, Sudakevitz, *et al.* 2004). These soluble lectins are hypothesized to play a role in host recognition but also in biofilm formation and cohesion (Imberty, *et al.* 2004, Imberty, *et al.* 2006).

The PA-IIL-like lectins demonstrate an unusual high affinity (micromolar range) for their monosaccharide ligands (Mitchell, E.P., *et al.* 2005, Sabin, *et al.* 2006). They form dimers or tetramers that have the characteristic to bind sugar through bridging by two calcium ions (Mitchell, E., *et al.* 2002). The dimmer architecture is at least requested for sugar binding since the C-terminus acidic group of one monomer participates in the calcium and carbohydrate binding of the other monomer. The lectins in this family bind both L-fucose and D-mannose, the preference depending on a stretch of three amino acids in the binding site (Adam, *et al.* 2007).

Three PA-IIL-like proteins have been identified in the genome of J2315 strain of *Burkholderia. cenocepacia*, a strain that is responsible for epidemic infections among cystic fibrosis patients (Govan and Deretic 1996, Lyczak, et al. 2002). BclA, the lectin coded by the shortest gene, has been recently characterized (Lameignere, et al. 2008). This lectin displays a strict specificity for mannosides with micromolar dissociation constant for the monosaccharide. The crystal structure of recombinant BclA complexed with αMeMan was solved at 1.7Å resolution and displays some differences with other PA-IIL like lectins. Whereas all previously characterized lectins in this family oligomerize as tetramer, BclA is a homodimer. Glycan array data run at Consortium for Functional Glycomics on 240 mammalian glycans established a strong specificity of oligomannose type N-glycans, with higher affinity for the ones presenting a terminal branched trimannoside αMan1-3(αMan1-6)βMan such as Man3-GlcNAc2 and Man5-GlcNAc2.

We present here the characterisation of the interaction between BclA and the dimannoside and trimannoside constituting N-glycans. The combined use of titration microcalorimetry and crystallography aimed at establishing the atomic basis of affinity. The use of atomic force microscopy confirmed the unusual oligomerisation mode observed in the presence of the trisaccharide ligand.

Results

Thermodynamics of binding

The specificity profile obtained from Glycan Array Data indicates that BclA binds to oligomannose-type N-glycans (Lameignere, *et al.* 2008). The two highest affinity oligosaccharides present the branched trisaccharide αMan1-3(αMan1-6)Man (TriMan) at their extremity. Glycans with branches capped by the αMan1-2Man disaccharide also appear to bind efficiently to the lectin. In order to assess the role of each mannose in the interaction, the trisaccharide TriMan and the three disaccharides αMan1-2Man (M12M), αMan1-3Man (M13M), and αMan1-6Man (M16M) have been tested by titration microcalorimetry, a method that is well suited for the characterization of protein-carbohydrate interaction (Dam and Brewer 2002).

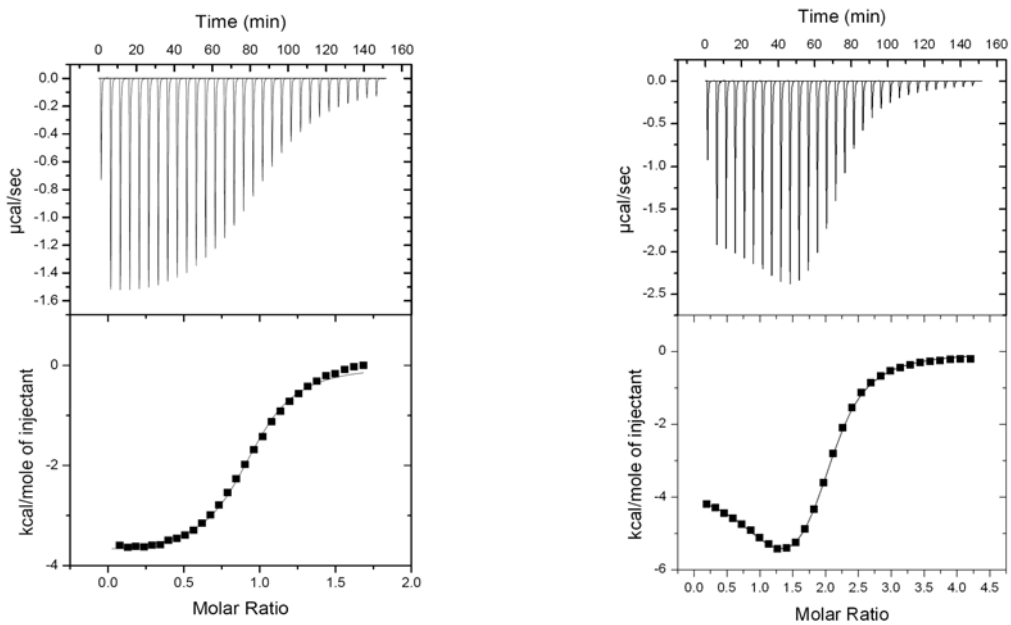


Figure 1. Microcalorimetry data. Left: upper panel: thermogram of BclA binding to M12M. Lower panel: integration of data with curve fitted for "one binding site" model. B: thermogram of BclA binding to M16M. Lower panel: integration of data with curve fitted for "sequential two binding sites" model.

In all cases, the interaction lead to strong exothermic peaks, indicating the enthalpic characteristics of the binding. The titration of BclA by M12M displays the classical sigmoid shape with strong exothermic peaks at the beginning, getting smaller during the titration and finishing by negligible peaks due to dilution effects at the end of the experiments (Figure 1A). Fitting of the data on a "one model" site yielded the expected stoechiometry value close to 2 (n=1.64) per BclA dimmer, with micromolar affinity (Kd= 6.7 μM). The free energy of binding (ΔG = -29.5 kJ/mol) has a strong enthalpy component (ΔG = -18.3 kJ/mol), but the entropy contribution is also favorable to the binding (TΔS = 11.2 kJ/mol) and participate to 30% of the free energy of binding (Table 1). Such a thermodynamical behavior is unusual in protein/carbohydrate interaction that classically presents a strong entropy barrier (Dam and Brewer 2002), but is has already been observed in this family of bacterial lectins where it has been proposed to be correlated to the direct involvement of the two calcium ions in the binding site (Mitchell, E.P., *et al.* 2005).

Table 1. Microcalorimetry data for BclA dimmers. All values are averaged over two experiments. Standard deviations are indicated in parenthesis. The "two sequential binding sites" model has been applied to all complexes except for the BclA/αMan1-2Man complex.

		Ka (10^4 M^{-1})	**Kd (10^{-6} M)**	**ΔG (kJ/mol)**	**ΔH (kJ/mol)**	**-TΔS (kJ/mol**
αMan1-2Man	n=1.64	15 (± 1)	6.7	-29.5	-18.3 (± 0.1)	-11.2
αMeMan [a]	Site 1_b	52.5	2.7	-32.7	-20.8	-11.9
	Site 2	11.8	7.8	-29.0	-28.4	-0.6
αMan1-3Man	Site 1	49 (± 3)	1.3	-32.5	-22.9 (± 1.2)	-9.6
	Site 2	17 (± 2)	5.8	-29.9	-27.7 (±1.9)	-2.2
αMan1-6Man	Site 1	127 (±	0.80	-34.8	-17.7 (± 0.3)	-17.1
	Site 2	15 (± 2)	6.8	-29.5	-29.5 (±0.7)	0
αMan1-3(αMan1-	Site 1	32 (± 9)	3.4	-31.3	-14.8 (± 0.6)	-16.6

[a] From reference (Lameignere, *et al.* 2008)

[b] Calculations based on the dimmer with two sequential binding sites.

The other ligands (M13M, M16M and TriMan) present a different profile of heat release. The first injections give rise to peaks with increasing exothermic signals before the titration occurs. Such behavior has been already observed for the BclA/αMeMan interaction and is correlated to a cooperative effect (Lameignere, *et al.* 2008). The two binding sites of the homodimmer are not independent and the binding of the ligand in one site enhances the enthalpy of binding in the second site. All curves have been therefore fitted using a "sequential two binding sites model". The thermodynamic data are listed in Table 1. In the cooperative mode, the first site appears systematically to be the highest affinity one, with Kd varying from 0.8 μM for M16M to 3.4 μM for the trisaccharide. This binding site presents a moderate enthalpy of binding and the entropy contribution can be very strong, up to 50% of the free energy for M16M and TriMan, the two ligands with stronger cooperative effect. The second binding site has slightly lower affinity with Kd varying from 6 to 72 μM. It is characterized by a higher value of the enthalpy of binding, giving rise to the strong peaks observed after few injections and a negligible entropy contribution.

Crystal structures of BclA in complex with dimannoside and trimannoside

The complex of BclA with M13M crystallized in $P2_1$ space group with two dimmers per asymmetric unit (Table 2). The data were collected to 1.75 Å and the structure was solved by molecular replacement using the BclA/αMeMan structure (Lameignere, *et al.* 2008).

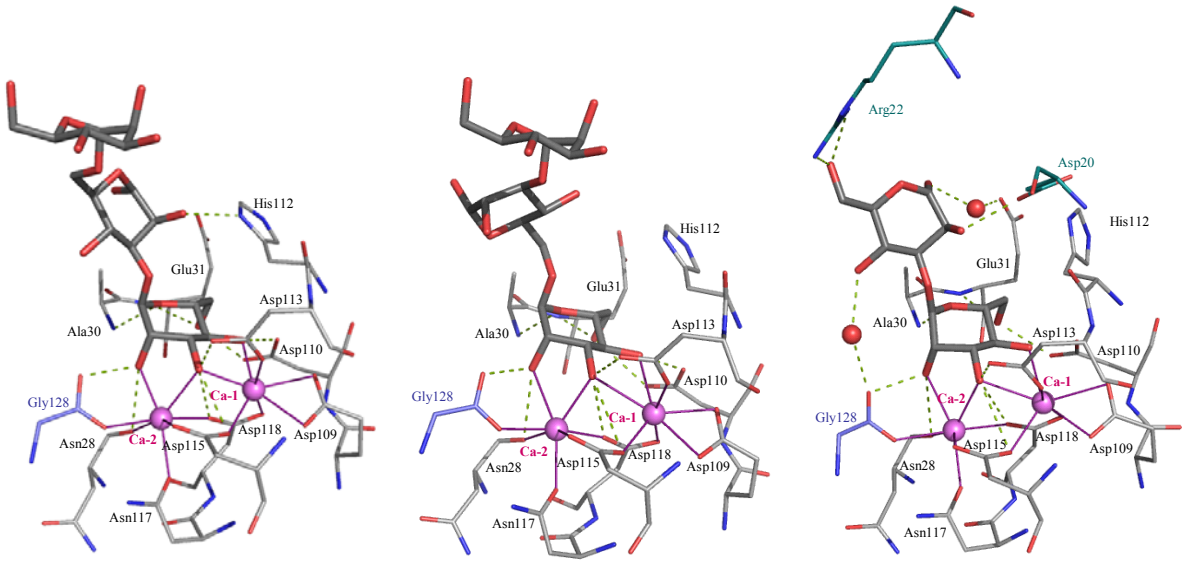


Figure 2: Graphical representation of the binding site of BclA complexed with M13M (left) and TriMan with either the 3-arm (center) or the 6-arm (right) in the binding site.

Table 2: Data collection and refinement statistics for the crystal structures of BclA in complex with di and trisaccharide.

Ligand	Manα1-3Man	Manα1-3(Manα1-6)Man
Data collection		
Beamline	ID29	ID29
Wavelength (Å)	0.979	0.979
Résolution (Å)	75.38-1.75 (1.84-1.75)	57.45-1.10 (1.16-1.10)
Space group	$P2_1$	$P\ 3_1\ 2\ 1$
Cell dimensions		
a, b, c (Å)	57.96 48.97 76.26	66.29 66.29 46.16
α, β, γ (°)	90.00 98.89 90.00	90.00 90.00 120.00
Measured reflections	141954 (13356)	182133 (10765)
Unique reflections	41914 (5322)	45990 (5456)
Average multiplicity	3.4 (2.5)	4.0 (2.0)
Completness (%)	97.8 (86.7)	96.5 (79.8)
R_{merge} (%)	7.9 (35.4)	4.6 (21.7)
$< I/\sigma(I)>$	11.2 (1.9)	16.8 (2.4)
Refinement		
R_{cryst} (%)	18.0	14.2
R_{free} (%)	22.1	17.3
Rms from ideality		
Bonds (Å)	0.016	0.018
Angles (°)	1.681	1.973

The first 4 to 9 N-terminal amino acids depending of the chains are disordered and could not be included in the final model. As previously described, BclA presents a nine-stranded antiparallel β-sandwich fold associated in a dimmer. Inspection of the first electron density maps reveals without ambiguity the disaccharide M13M in chains A and D. Only the first mannose could be placed in molecule C since the second mannose is too disordered to be modelled. No ligand is observed in molecule B since its binding is hindered by a symmetric molecule but a sulfate ion is found in the binding site. The non-reducing mannose participates directly in the coordination of the Ca^{2+} ions through hydroxyl groups O2, O3 and O4 as previously observed for the αMeMan complex and the contacts are listed in Table3. The reducing mannose is stabilized through interactions with a symmetric molecule or mediated by water molecules. The O1 hydroxyl interacts with Glu31 OE2 via a water molecule as well

as the O4 hydroxyl with the main chain oxygen of Gly128 of the other dimmer monomer. The O2 and O6 hydroxyls interact with Asp20 OD2 and Arg22 NH2 of the symmetric monomer of the dimmer respectively. Following steric hindrance with the reducing mannose, the His112 side chain, which adopted several conformations previously, is only observed in the "open" shape.

Table 3: Contacts in the binding site of BclA with di and trisaccharide

Sugar atoms		**Manα1-3(Manα1-6)Man**	**Manα1-6(Manα1-3)Man**	**Manα1-3Man**
O3	Ca-1	2.31	2.75	2.57
O4	Ca-1	2.30	2.78	2.52
O2	Ca-2	2.27	2.66	2.57
O3	Ca-2	2.50	2.49	2.49
Hydrogen bonds between the first mannose residue and the protein				
O2	Asn28 (O)	2.99	3.01	2.96
	Gly128 (O)a	3.05	2.57	2.49
O3	Asp113 (OD2)	2.54	2.70	2.56
	Asp115 (OD1)	2.75	3.00	3.02
	Asp118 (OD1)	2.93	3.10	3.09
O4	Asp110 (OD1)	2.60	2.64	2.53
O5	Ala30 (N)	3.30	2.83	3.06
O6	Glu31 (N)	3.32	2.76	3.03
	Asp110 (OD2)	2.45	2.90	2.61
Hydrophobic interactions		His 112		
Hydrogen bond between the second Man residue and the protein				
O1	Wat – Glu31 (OE2)			2.73-2.77
O2	His112 (ND1)	2.80		
	Asp20 (OD2) sym			2.57
O4	Wat – Gly128 (O)a			2.68-2.66
O6	Arg22 (NH2) sym			3.12
	Arg22 (NH1) sym			3.08

a: from the second monomer of the dimmer

BclA complexed with the trisaccharide formed crystals that diffracted to high resolution (1.1 Å) in space group $P3_121$. The content of the asymmetric unit is a monomer and the functional dimmer is generated by the two fold axis. Examination of the residual electron density revealed that the trisaccharide TriMan is present in the binding site. It is centered on the two fold axis and bridges two neighboring dimmers. This is not compatible with the geometry of the ligand since the trisaccharide αMan1-3(αMan1-6)Man does not present a 2-fold axis. Further analysis demonstrated that 50% of the ligand is in one orientation (the Man of the 3-

arm in one binding site and the Man of the 6-arm in the second one) and 50% in the reverse orientation. Each orientation was therefore refined to 50% occupancy.

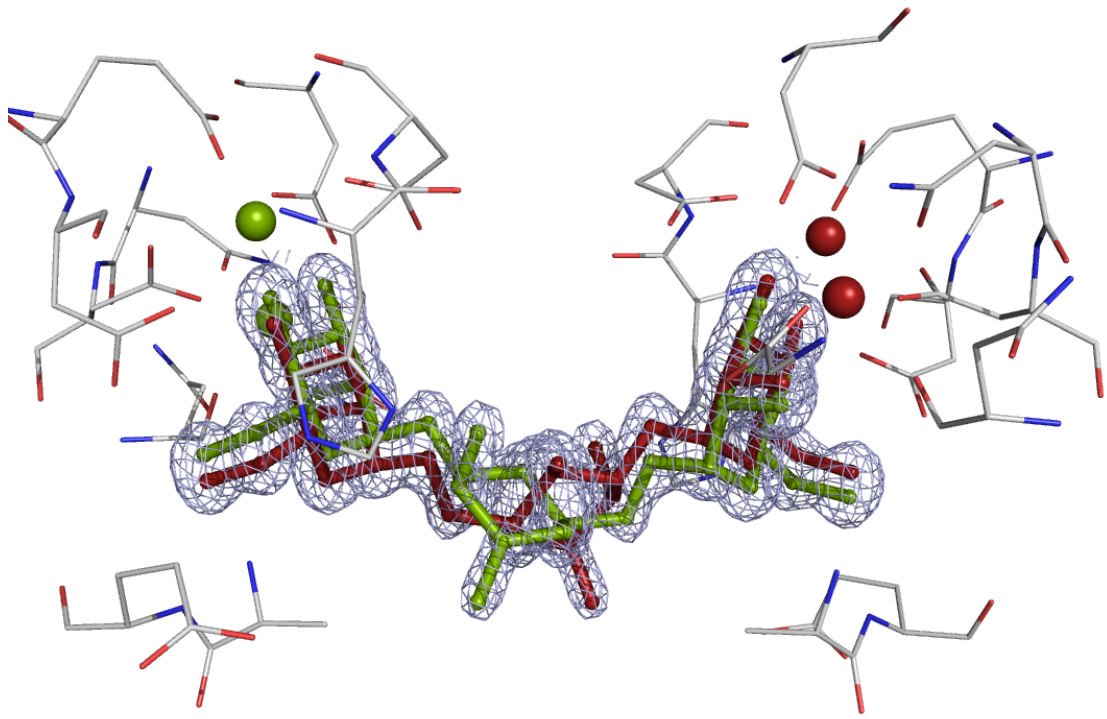


Figure 3. Representation of the 2mFo-DFc electron density for the bridging TriMan in the crystal structure of BclA. The map is contoured at 1 ρ representing 0.48 eÅ^{-3}.

The mannose on the 6-arm does not have so close contact with the calcium ions and the hydroxyl groups present coordination closer to 2.6 Å than to 2.3 Å. On the opposite the mannose on the 3-arm is more deeply buried in the binding site. In addition, the reducing mannose in this case establishes an additional hydrogen bond with His112. The bridging position of the ligand results in a macromolecular architecture inside the crystal structure. Neigbouring BclA dimmers share the TriMan ligand and are therefore assembled in an infinite chain (Figure 4). No branching points can occur because of the dimeric character of the protein and the divalent properties of the ligand. To our knowledge, no such one-dimensional architecture has been observed previously in lectins crystal structures.

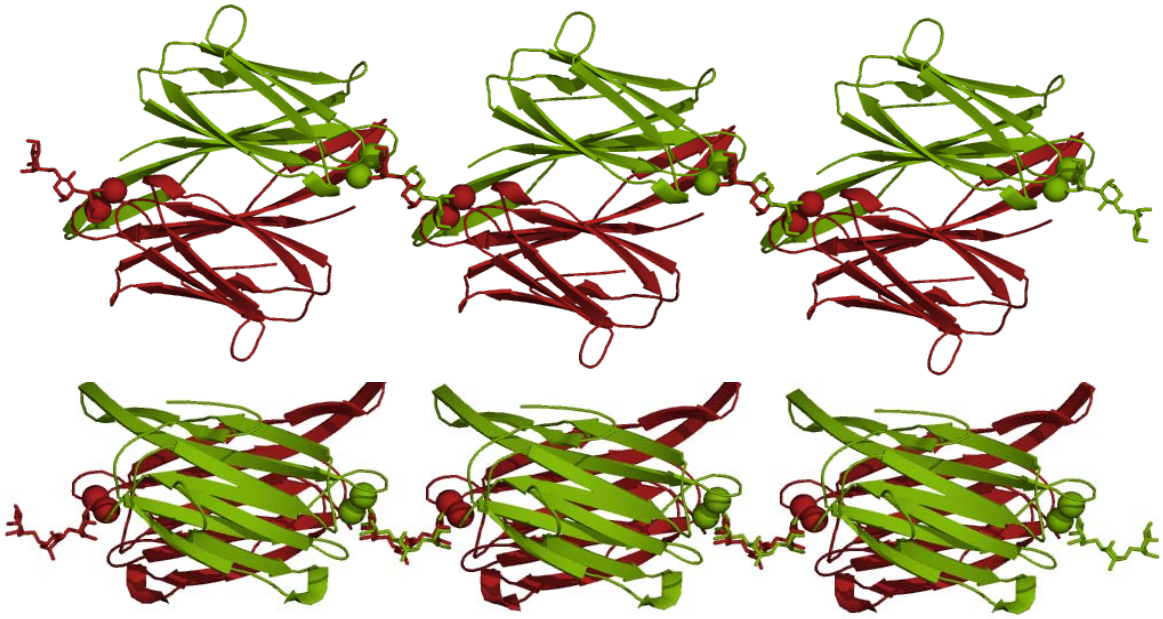


Figure 4. Two orthogonal views of the infinite chains of protein-ligand-protein interactions occurring in BclA/TriMan crystal

Analyse of the AFM samples

The solutions of BclA alone and in complex with BclA were prepared and spread on mica surface to analyze the patterns imaged by atomic force microscopy under the same conditions. In absence of ligand, the images reveal the existence of big and widely spread aggregates (Figure 5A). The height of those aggregates is between 5 and 20 nm and their size of a few hundred nanometers. On the contrary Figure 5B evidences a structured film consisting in small filaments of BclA/TriMan complexes. Height histogram gives variation of 1.5 nm which are relevant with the size of the proteins in a dried state. Moreover the lateral size of the observed filaments (below 20nm) is also consistent with the imaging of couples of molecules with classical non-contact tip (tip radius better than 10 nm).

The filaments are mostly linear with rare branching points. Their lengths vary between 20 to 200 nm, which would correspond to 4 tot 40 trimannoside-bridged BclA dimmers in the arrangement observed in the crystal structure (Figure 4). The kinks observed along the filament generate variations of directions of about 120°. It is expected that the α1-6 linkage in the TriMan ligand can display some flexibility and generates such defects in the linearity.

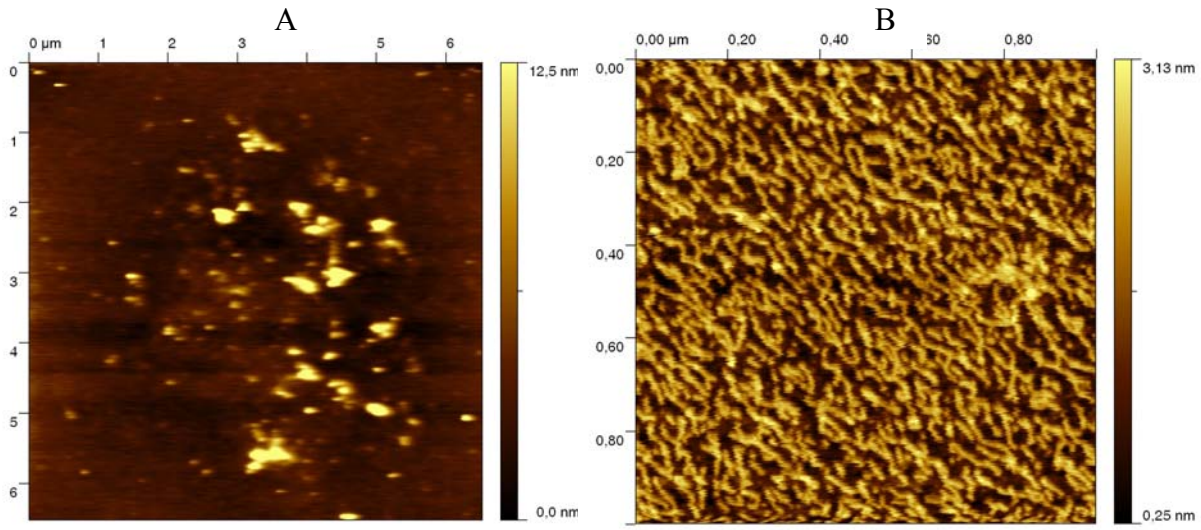


Figure 5: A) 6,5 x 6,5 µm^2 topography picture of proteins onto freshly cleaved mica B) 1x1 µm^2 topography picture of proteins and trisaccharides deposited onto freshly cleaved mica. The lateral color bar gives the vertical scale bar on each sample.

Discussion

The dimeric BclA has the potential to generate linear filament when in the presence of a cross-linking ligand such as the trimannoside ManTri. In the crystal state, this results in perfectly aligned succession of protein-carbohydrate-protein- arrangement. The AFM study demonstrates that such linearity is mostly conserved in solution. The resulting images are drastically different from what has been obtained recently for ConA interacting with synthetic mannosides (Gour, *et al.* 2008). The ConA sample was a mixture of dimmer and tetramers and the interaction with dimeric ligand resulting in a two-dimensional organized lattice similar to a fisherman net. In our study, the absence of tetrameric protein prevents the formation of such network and only linear filament is observed.

Material and methods

Cloning

The recombinant protein BclA was cloned and expressed using the procedure previously described (Lameignere, *et al.* 2008). Briefly, the BCAM0186 gene was amplified by polymerase chain reaction using genomic DNA from *Burkholderia cenocepacia* strain J2315

obtained from the Czech Collection of Microorganisms at Brno, as a template. Fragment corresponding to the protein was then introduced into pRSET-A vector (Invitrogen), resulting in plasmid *pRSETbcla*.

Production and purification of recombinant BclA protein in E. coli

Using the expression plasmid *pRSETbcla*, BclA was expressed in Escherichia *coli* BL21(DE3). Cells were grown in 500ml of Luria broth medium at 37 °C. When the culture reached a D_{600} of 0.6, IPTG (isopropyl β-Dthiogalactoside) was added to a final concentration of 0.05 mM. Cells were harvested after a 4-h incubation at 30 °C, washed, and resuspended in 15 ml of the equilibrating buffer (20mM Tris/HCl pH 7.5, 100mM NaCl and 100µM $CaCl_2$). Then, cells were broken by cell disruption (Constant Cell Disruption System). After centrifugation at 50 000g for 45min at 6°C, the supernatant was further purified by affinity chromatography on a D-mannose–agarose column (Sigma–Aldrich). BclA was allowed to bind to the immobilized mannose in equilibrating buffer. After washing with buffer (20 mM Tris/HCl pH7.5, 1 M NaCl and 100 µM CaCl2), it was eluted using 20 mM Tris/HCl pH7.5, 100 mM NaCl and 1 mM EDTA. The purified protein was intensively dialyzed against distilled water containing $CaCl_2$ (5mM) for 2 days, and against distilled water for a further 4 days, freeze-dried and lyophilized, and kept at −20 °C. BclA is obtained with a typical yield of about 40 mg of the purified protein per litre of culture. The resulting protein moved as a sharp band of about 14 kDa on SDS-PAGE. Mass spectroscopy analysis confirmed the molecular mass of 13764 kDa, corresponding to the BclA amino acid sequence lacking the initial methionine residue

Microcalorimetry

Titration calorimetry experiments were performed using Microcal VP-ITC microcalorimeter (Microcal, Northampton, MA). All titrations were performed in 100mM Tris/HCl pH 7.5 buffer containing 3µM $CaCl_2$, at 25 °C. Aliquots of 10 µl of each carbohydrate at concentrations varying form 1.5 to 3 mM, dissolved in the same buffer, were added at 5 min intervals to the lectin solution present in the calorimeter cell. In the titrations, the protein concentration in the cell varied from 0.15 to 0.4 mM. The temperature of the cell was controlled to 25 ± 0.1 °C. Control experiments performed by injections of buffer in the protein solution yielded to insignificant heats of dilution. Integrated heat effects were analysed by non-linear regression using a single site-binding model (Microcal Origin 7).

Fitted data yielded the association constant (K_a) and the enthalpy of binding (ΔH). Other thermodynamic parameters, i.e. changes in free energy, ΔG, and entropy, ΔS, were calculated from equation

$$\Delta G = \Delta H - T\Delta S = -RT\ln K_a$$

where T is the absolute temperature and R = 8.314 $J.mol^{-1}.K^{-1}$. Two to three independent titrations were performed for each ligand tested.

Atomic Force Microscopy :

Samples were prepared in double distilled water containing $CaCl_2$ to a 0.3 µM final concentration. Carbohydrate solution (10 µl, 100 µM) was added to protein solution (10 µl, 0.5 mg ml^{-1}) and the volume was made up to 40 µl with $CaCl_2$ containing water. These solutions were then incubated for 1 h at room temperature. Each sample was diluted 50-fold in water and a 20 µl aliquot was deposited on the freshly cleaved mica surface. The sample was dried over night in desiccator. AFM experiments were performed on a Pico plus (Molecular Imaging) commercial instrument. Topography pictures were obtained using tapping mode with Mikromash NSC 36 tips. Data treatment (height measurements after base line correction only) and presentation were realized under Gwyddion Software

Crystallization and data collection

Crystals were obtained by the hanging-drop vapor diffusion method using 5µl drops containing a 50:50 (v/v) mix of reservoir solution combined to BclA-sugar solution at 20 °C. Lyophilized protein was dissolved in water (15 $mg.ml^{-1}$) in the presence of Manα1-3Man and Manα1-3(Manα1-6)Man with a concentration of 1.49 and 2.20 $mg.ml^{-1}$ respectively.

The first crystals of BclA / Manα1-3(Manα1-6)Man complex was obtained with the initial condition: 30% (w/v) PEG 3500MME (monomethyl ether) in 0.2M ammonium sulphate. Optimization of the initial crystallization conditions yielded crystals resembling small cubes with 80% of the initial precipitant solution (30% PEG in 0.2M ammonium sulphate). This crystal form belongs to space group $P3_121$ with cell dimensions a=66.3 Å, b=66.3 Å and c=46.2 Å.

Crystals of BclA / Manα1-3Man complex appeared in the form of thin plates after several days in the presence of 30% w/v PEG 5000MME and 0.2M ammonium sulphate. This crystal

form belongs to space group $P2_1$ with cell dimensions a=57.9 Å, b=49.0 Å, c=76.3 Å β= 98.89º.

Data were collected at the European Synchrotron Radiation Facility (Grenoble, France) on ID29 beamline using an ADSC Quantum 210 CCD detector. Data processing were performed using MOSFLM (Leslie 1992) before to be scaled and converted to structure factors using SCALA in the CCP4 suite program (number4 1994). All further computing was performed using the CCP4 suite unless otherwise stated. Statistics for the two data collections are shown in Table 2.

Structure solution and refinement

Molecular replacement technique was used to solve the two structures with MOLREP (Vagin and Teplyakov 1997) using the monomeric coordinates of BclA structure complexed with αMeMan (Protein Data Bank code 2VNV) as the search model. Crystallographic refinement was carried out with the program REFMAC (Murshudov, *et al.* 1997)] interspersed with manual model building using Coot (Emsley and Cowtan 2004). The incorporation of the ligand was performed after inspection of the mFo−DFc weighted maps. Water molecules were introduced automatically using Coot and inspected manually. The stereochemical quality of the model was assessed with the program Procheck (Laskowski, *et al.* 1993). All figures were drawn with PyMOL Molecular Graphics System (2002) DeLano Scientific, Palo Alto, CA.

Acknowledgements

The work was supported by CNRS. This research is part of E.L. Ph.D. thesis who is supported Conseil Régional Rhône-Alpes (Cluster 10). Travel and visits between NCBR and CERMAV were supported by a BARRANDE exchange program (M.W. & A.I.) and financial help is gratefully acknowledged from the French cystic fibrosis association 'Vaincre la Mucoviscidose' and GDR Pseudomonas. The authors wish to thank Ms Catherine Gautier for excellent technical help. We thank the ESRF, Grenoble, for access to synchrotron data collection facilities.

References

Adam, J., Pokorna, M., Sabin, C., Mitchell, E.P., Imberty, A. and Wimmerova, M. (2007) Engineering of PA-IIL lectin from *Pseudomonas aeruginosa* - Unravelling the role of the specificity loop for sugar preference. BMC Struct. Biol., 7 36.

Dam, T.K. and Brewer, C.F. (2002) Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. Chem. Rev., 102, 387-429.

Emsley, P. and Cowtan, K. (2004) Coot: Model-Building Tools for Molecular Graphics. Acta Crystall. D Biol. Crystall., 60, 2126-2132.

Gilboa-Garber, N. (1982) *Pseudomonas aeruginosa* lectins. Methods Enzymol., 83, 378-385.

Gour, N., Purohit, C.S., Verma, S., Puri, R. and Ganesh, S. (2008) Mannosylated self-assembled structures for molecular confinement and gene delivery applications. Biochem. Biophys. Res. Commun.

Govan, J.R. and Deretic, V. (1996) Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. Microbiol. Rev., 60, 539-574.

Imberty, A. and Varrot, A. (2008) Microbial recognition of human cell surface glycoconjugates. Curr. Opin. Struct. Biol., 18, 567-576.

Imberty, A., Wimmerova, M., Mitchell, E.P. and Gilboa-Garber, N. (2004) Structures of the lectins from *Pseudomonas aeruginosa*: Insights into molecular basis for host glycan recognition. Microb. Infect., 6, 222-229.

Imberty, A., Wimmerova, M., Sabin, C. and Mitchell, E.P. (2006) Structures and roles of *Pseudomonas aeruginosa* lectins. In Bewley, C. (ed), Protein-Carbohydrate Interactions in Infectious Disease. The Royal Society of Chemistry, Cambridge, pp. 30-48.

Lameignere, E., Malinovská, L., Sláviková, M., Duchaud, E., Mitchell, E.P., Varrot, A., Šedo, O., Imberty, A. and Wimmerová, M. (2008) Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia*. Biochem. J., 411, 307-318.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr., 26, 283-291.

Leslie, A.G.W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, 26.

Lyczak, J.B., Cannon, C.L. and Pier, G.B. (2002) Lung infections associated with cystic fibrosis. Clin Microbiol Rev, 15, 194-222.

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 Struct. Biol., 9, 918-921.

Mitchell, E.P., et al. (2005) High affinity fucose binding of *Pseudomonas aeruginosa* lectin PA-IIL: 1.0 Å resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches. Proteins: Struct. Funct. Bioinfo., 58, 735-748.

Murshudov, G.N., A.A.Vagin and E.J.Dodson (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr., D53, 240-255.

Collaborative Computational Project Number4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr., 760-763.

Pokorná, M., et al. (2006) Unusual entropy driven affinity of *Chromobacterium violaceum* lectin CV-IIL towards fucose and mannose. Biochemistry, 45, 7501-7510.

Sabin, C., Mitchell, E.P., Pokorná, M., Gautier, C., Utille, J.-P., Wimmerová, M. and Imberty, A. (2006) Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa:* Thermodynamics data correlated with X-ray structures. FEBS Lett., 580, 982-987.

Sharon, N. (1996) Carbohydrate-lectin interactions in infectious disease. Adv. Exp. Med. Biol., 408, 1-8.

Sharon, N. (2006) Carbohydrates as future anti-adhesion drugs for infectious diseases. Biochim. Biophys. Acta, 1760, 527-537.

Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E.P., Lerrer, B., Wimmerova, M., Katcof, f.D.J., Imberty, A. and Gilboa-Garber, N. (2004) A new *Ralstonia solanacearum* high affinity mannose-binding lectin RS-IIL structurally resembling the *Pseudomonas aeruginosa* fucose-specific lectin PA-IIL. Mol. Microbiol., 52, 691-700.

Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. J. Appl. Cryst., 30, 1022-1025.

Chapitre 6

Caractérisation de la lectine BclB

L'étude d'une deuxième lectine soluble de *B. cenocepacia*, BclB, fait l'objet de ce chapitre. L'analyse de la séquence protéique traduite du génome de *B. cenocepacia* J2315 montre que cette protéine de 243 résidus (25.4 kDa) est composée de deux domaines : un domaine *N*-terminal de 128 acides aminés et de fonction inconnue et un domaine lectine *C*-terminal similaire à la lectine PA-IIL. La comparaison entre les séquences du domaine *C*-term de BclB et celle de PA-IIL confirme une très forte conservation des acides aminés impliqués dans la liaison au calcium et au sucre. Le domaine *N*-term ne présente pas de similarité avec des séquences d'autres protéines. L'étude structure-fonction décrite dans ce chapitre a porté sur la lectine BclB recombinante entière, mais également sur les deux domaines produits séparément.

6.1 Matériel et méthodes

6.1.1. Clonage des gènes d'intérêt

Les amorces utilisées pour amplifier les gènes d'intérêt ainsi que les vecteurs de clonage sont récapitulés respectivement dans les Tableaux 6.1 et 6.2.

Protéine BclB entière

Le gène codant pour BclB (*Bcam0184*) a été amplifié par PCR (Polymerase Chain Reaction) avec la Pfu polymerase (Promega) et les amorces BC3-6L et BC3-5U. L'ADN génomique de *B. cenocepacia* souche J2315 disponible à l'Université Masaryk a servi de matrice. Par des méthodes de biologie moléculaire classiques, le gène amplifié a été inséré entre les sites de restriction *NdeI* et *HindIII* (*Tableau 6.1*) du plasmide d'expression procaryotique pRSET-a (Invitrogen) qui comporte un gène de résistance à l'ampicilline (Amp) (*Tableau 6.2*)

Protéine clonée	Nom de l'amorce	Séquence de l'amorce	Sites de restriction
BclB	BC3-6L	R GCAAGCTTCAACCCAGCGGCCAG	*HindIII*
	BC3-5U	F CGAGGACCATATGTCACAACCCTTTAC	*NdeI*
BclBLec	BC3-6L	R GCAAGCTTCAACCCAGCGGCCAG	*HindIII*
	BCLBC-up	F CTTAGCGGCCATATGGAACGTGATGGC	*NdeI*
BclBNter17	3317	R AAAACTCGAGTCAGAAGATGCCGTCGCGTT	*XhoI*
	BclB-5U	F GGACTACCATATGTCACAACCCTTTACCCAC	*NdeI*
BclBNter14	3314	R AAAACTCGAGTCAGCCGCCGCCGGTCGGCAC	*XhoI*
	BclB-5U	F GGACTACCATATGTCACAACCCTTTACCCAC	*NdeI*

Tableau 6.1. Liste des amorces nucléotidiques utilisées lors des clonages

F : Forward, sens, R : Reverse, antisens

Domaine Nterminal de BclB : BclBNter14 et BclBNter17

Deux constructions de longueur différentes, BclBNter14 et BclBNter17, ont été produites pour ce domaine. Les régions codantes pour BclBNter14 et BclBNter17 ont été amplifiées par PCR avec la Taq polymérase (Qiagen) et les amorces 3314 / BclB-5U et 3317 / BclB-5U respectivement (*Tableau 6.1*). L'ADN de *B. cenocepacia* souche K56-2 disponible à l'université de Western Ontario (BCRRC, *B. cepacia* Complex Research and Referral Repository for Canadian CF Clinics) a servi de matrice pour la PCR.

Les fragments amplifiés ont été clonés en phase avec une queue hexahistidine *N*-terminale clivable par la thrombine dans le vecteur d'expression procaryotique pET-28a (Novagen®) qui comporte un gène de résistance à la Kanamycine (Kan) (*Figure 6.1*) entre les sites de restriction NdeI et XhoI (*Tableau 6.2*). Ces constructions comportent donc les 20 résidus MGSSHHHHHHSSGLVPRGSH supplémentaires à leur extrémité *N*-terminale

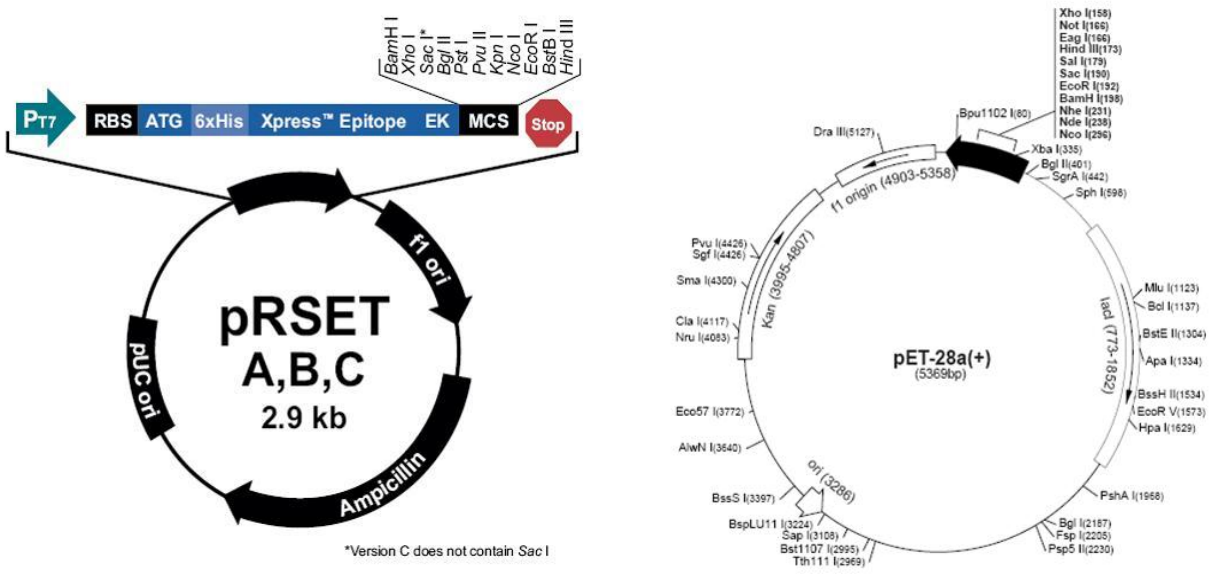


Figure 6.1. Représentation schématique des vecteurs utilisés.

Domaine lectine de BclB : BclBLec

La partie codante pour le domaine lectine de BclB a été amplifiée par PCR en utilisant comme matrice la construction pRSET-a_*BclB*, la HotStar High Fidelity DNA Polymérase (Qiagen®) et les amorces BC3-6L et BCLBC-up (*Tableau 6.1*). Le produit de la PCR est ensuite extrait sur gel d'agarose 1.5% grâce au kit NucleoTrap (Macherey Nagel®). Le plasmide pRSET-a ainsi que l'insert sont digérés par les enzymes de restriction *NdeI* et *Hind III* (Invitrogen®). L'insert est purifié avec le kit SureClean Plus (Bioline®) pour éliminer sels et enzymes. Le vecteur digéré est purifié sur gel d'agarose 1,5% et extrait grâce au kit NucleoTrap. Après une ligation rapide (5min à température ambiante) grâce au kit Mighty Mix (Takara®), le produit de ligation est transformé dans la souche d'amplification XL1-blue d'*E. coli* par choc thermique 45 secondes à 42°C. L'accord de séquence a été vérifié par séquençage.

Protéine clonée	Vecteur de clonage	Fournisseur	Taille (kb)	Gène de résistance	Remarque
BclB	pRSET-a	Invitrogen	2.9	Ampicilline	
BclBLec					
BclBNter17	pET28-a	Novagen	1.2	Kanamycine	Queue $(His)_6$ clivable à la thrombine en *N*-term
BclBNter14					

Tableau 6.2. Vecteurs utilisés pour les clonages de BclB et de ses domaines.

6.1.2. Expression des protéines recombinantes chez E. coli

Les protocoles d'expression de BclB, BclBLec et BclBNter sont très similaire. La souche d'*E. coli* BL21 (DE3) (Novagen®) a été utilisée pour l'expression des deux domaines séparés. La protéine BclB entière a quand à elle été produite dans la souche E. coli BL21pLysS (DE3) (Novagen®).

Une préculture liquide en Luria Broth (LB) contenant les antibiotiques appropriés est réalisée à 37°C sous agitation jusqu'à une densité optique à 600nm (DO600) de 0.6. Un litre de milieu de culture contenant les antibiotiques appropriés est inoculé par 30ml de préculture et placé sous agitation à 37°C jusqu'à une DO_{600} appropriée. L'expression du gène codant pour la protéine est induite par ajout d'Iso-Propyl-β-D-Thiogalactopyranoside ou IPTG. Au terme de 3 heures d'induction à 30°C, les cellules sont centrifugées 20 minutes à 6000 g et 20°C.

Au cours de l'expression, les méthionines du domaine BclBNter ont été substituées par des sélénométhionines afin d'utiliser le signal anomal du sélénium lors de la collecte de données et d'obtenir les phases nécessaires à la résolution de la structure. Pour cela, la méthode de Doublié qui repose sur l'inhibition de la synthèse de la méthionine a été utilisée (Doublié 1997). En effet, en inhibant l'aspartokinase, des concentrations élevées en isoleucine, lysine et thréonine permettent de bloquer la voie de biosynthèse de la méthionine chez *E. coli*. De plus, la phénylalanine et la leucine agissent de façon synergique avec la lysine.

		Volumes pour 1l de culture (en ml)	Stérilisation
Solution M9 concentrée 5X Na_2HPO_4, 34 $g.l^{-1}$ KH_2PO_4, 15 $g.l^{-1}$ NaCl 2.5 $g.l^{-1}$ NH_4C, 5 $g.l^{-1}$ H_2O qsp 1l		200	Autoclavé
$MgSo_4$	1M	2	Autoclavé
Glucose	20%	20	Filtré
$CaCl_2$	1M	0.1	Autoclavé
Thiamine	1mg/ml	1	Filtré
$FeSO_4$, $7H_2O$	125mg/ml	2	Filtré
H_2O qsp 1 litre		778	Autoclavé

Tableau 6.3. Milieu de culture utilisé pour la production de protéine séléniée

qsp : quantité suffisante pour.

Une préculture liquide de 30ml est réalisée sur la nuit à 37°C sous agitation. Après centrifugation 20 min à 6000 g et 20°C, le culot bactérien est lavé trois fois avec du milieu minimal M9 supplémenté (*Tableau 6.3*), afin d'éliminer toute trace de LB. Les bactéries ainsi resuspendues sont utilisées pour inoculer 1litre de milieu minimum. Lorsque la DO_{600} atteint une valeur de 0.6-0.8, le mélange d'acides aminés : Lysine, Phénylalanine and Thréonine à 100 $mg.l^{-1}$; Isoleucine, Leucine and Valine à 50 mg l^{-1} ainsi que la sélénométhionine (SelMet) à 50 mg l^{-1} est ajouté. Après avoir laissé les cellules reposer à température ambiante 15 à 20 min sur la paillasse, l'expression est induite par ajout de 0.05mM d'IPTG. Après une induction à 20°C sur la nuit, les cellules sont centrifugées 20 min à 6000 g et 20°C.

6.1.3. Purification des protéines recombinantes

Toutes les chromatographies effectuées au cours de ces travaux ont été faites sur un système ÄKTA-FPLC (Amersham Biosciences). Dans ce système automatisé, les paramètres tels que la pression, le pH et la température sont contrôlés. L'élution des protéines est suivie par mesure de la densité optique à 280nm au cours du temps. Les tampons de purification récapitulés dans le *Tableau 6.4*, sont systématiquement filtrés et dégazés grâce à un système de pompe à vide. Les différentes étapes de purification et la pureté des protéines ont été analysées par électrophorèse sur gel en conditions dénaturantes ou SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis) 12% pour la protéine entière et 15% pour les domaines séparés. Les quantités de protéines sont dosées soit selon la méthode de Bradford à l'aide du kit Biorad (Quick Start™ Bradford Protein Assay) en suivant les

instructions du fabricant, soit par mesure de l'absorbance à 280 nm en utilisant leur coefficient d'absorption théorique.

Préparation des extraits bruts

Les culots correspondant à 1 litre de culture sont repris dans 20ml du tampon d'équilibration utilisé pour la purification. Afin de faire éclater les membranes cellulaires, les cellules concentrées dans le tampon d'équilibration sont soumises à une forte pression (1.5 kbar) lors de leur passage dans une buse de faible diamètre en utilisant un disrupter de cellules (One shot cell Disrupter, Constant System). Une goutte d'« Antifoam 204 » (Sigma®) est ajouté aux cellules afin d'éviter la formation trop excessive de mousse après passage dans la buse. Les broyats bactériens sont centrifugés 45 minutes à 50 000g et 6°C afin de culotter les débris cellulaires du lysat. Le surnageant contenant les protéines solubles est filtré avec des filtres en microfibres de verre (Glass Microfibre Filters 691, VWR) de 1.6 µm sur un système de pompe à vide.

Purification des lectines : BclB et BclBLec par chromatographie d'affinité

Les protéines BclB et BclBLec ont été purifiées sur une colonne d'affinité D-mannose-agarose (Sigma-Aldrich, USA) en une seule étape. La colonne est équilibrée avec 5 volumes du tampon d'équilibration A1: 20 mM Tris-HCl pH 7.5, 100 mM NaCl, 100 µM $CaCl_2$. L'extrait protéique est chargé sur la colonne avec un débit de 1 ml min^{-1}. La colonne est ensuite lavée avec le tampon d'équilibration puis les contaminants interagissant de façon non spécifique avec la résine sont éliminés par un lavage avec un tampon B1 contenant une forte concentration en sels (20 mM Tris-HCl pH 7.5, 1 M NaCl, 100 µM $CaCl_2$). Pour éluer la lectine, deux tampons ont été testés. Le tampon C1 contenant de l'EDTA (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 100 mM EDTA) et le tampon D1 correspondant au tampon d'équilibration A1, auquel une forte concentration en sucre a été ajoutée (100 mM de L-fucose).

Les fractions correspondant à la lectine sont ensuite rassemblées pour être dialysées contre 5mM de $CaCl_2$ pendant 2 jours puis contre de l'eau pendant 4 jours supplémentaires. Les protéines sont alors concentrées par ultrafiltration sur Vivaspin 6 (Vivasiences) avec un seuil de coupure de 10000 et 5000 daltons pour BclB et BclBLec respectivement. Les lectines sont soit conservées à 4°C soit lyophilisées puis congelées pour une conservation de longue durée.

Purification du domaine N-terminal par chromatographie d'affinité

La présence de la queue histidine en position N-terminale des protéines recombinantes BclBNter14 et BclBNter17, permet leur purification par chromatographie d'affinité de type IMAC (Immobilized Metal ion Affinity Chromatography).

La purification de la protéine se fait à 4°C sur colonne HisTrap™ FF (GE_Healthcare®) de 5 ml et d'une capacité d'environ 40 mg de protéine-$(His)_6$ par ml de résine. La colonne est équilibrée dans 5 volumes de tampon d'équilibration A2: 20 mM Tris-HCl pH 8.5, 500 mM NaCl, 10 mM Imidazole. L'extrait protéique est dilué dans un volume final de 50 ml avant d'être chargé sur la colonne. Après un lavage avec le tampon d'équilibration A2, la protéine est éluée par un gradient d'imidazole grâce au tampon B2 : 20 mM Tris-HCl pH 8.5, 500 mM NaCl, 500 mM Imidazole. Les fractions collectées n'étant pas de pureté égale après examen sur gel SDS-PAGE 15%, elles ne sont pas rassemblées et sont traitées séparément.

L'étiquette histidine est ensuite éliminée par clivage avec la thrombine biotinylée en suivant les instructions du fournisseur « Thrombin cleavage capture kit » (Novagen®). L'ajout de résine Streptavidine Agarose (Novagen®), à raison de 16 µl par unité d'enzyme permet de fixer la thrombine et de l'éliminer après incubation 1 heure à température ambiante sous agitation douce. Le mélange est centrifugé 10 min à 13000 g et 4°C afin de culotter l'agarose-streptavidine lié à la thrombine. La protéine est ensuite redéposée sur colonne de Nickel afin de séparer les protéines clivées des non clivées. Les protéines sans étiquette histidine passent à travers la colonne sans s'y fixer ("flow through"). Les protéines non digérées restent fixés à la colonne grâce à leur étiquette $(His)_6$.

Finalement le tampon de la protéine est échangé pour celui désiré pour les essais de cristallisation grâce à une colonne de dessalage PD-10 (GE Healthcare®) en suivant les instructions du fabricant. La protéine est ensuite concentrée par ultrafiltration comme décrit pour le domaine lectine isolé.

Dans le cas de la protéine séléniée, des agents anti-oxydants ont été rajoutés aux différents tampons lors de la purification afin de prévenir la réduction des sélénométhionines incorporées dans la protéine. 5mM de β-mercaptoéthanol ont été ajoutés aux tampons de purification. Avant concentration et pour les test de cristallogenèse, 0.5 mM de TCEP (tris(2-carboxyethyl)phosphine) , plus stable que le β-mercaptoéthanol ont été ajoutés. Les tampons correspondant sont notés "prime" (*Tableau 6.4*).

Type de chromatographie	Equilibration	Elimination des contaminants	Elution	Tampon final après colonne de dessalage
BclB et BclBLec				
Affinité sur mannose agarose	*Tampon A1* 20mM Tris-HCl pH7.5, 100mM NaCl, 100µM $CaCl_2$	*Tampon B1* 20mM Tris-HCl pH7.5, 1M NaCl, 100µM $CaCl_2$	*Tampon C1* 20mM Tris-HCl pH7.5, 100mM NaCl, 0.1M EDTA	*Tampon A1* 20mM Tris-HCl pH7.5, 100mM NaCl, 100µM $CaCl_2$
	Idem	Idem	*Tampon D1* 20mM Tris-HCl pH7.5, 100mM NaCl, 100µM $CaCl_2$ 0.1M L-Fucose	*Tampon A1* 20mM Tris-HCl pH7.5, 100mM NaCl, 100µM $CaCl_2$
BclBNter				
Affinité sur HisTrap	*Tampon A2* 20mM Tris-HCl pH8.5, 500mM NaCl 10mM, Imidazole		*Tampon B2* 20mM Tris-HCl pH8.5, 500mMNaCl 500mM Imidazole	*Tampon C2* 20mM Tris-HCl pH8.5, 150mM NaCl
	Idem	Idem	*Idem*	*Tampon D2* 20mM HEPES-NaOH pH8.5, 150mM NaCl
BclBNter SelMet				
Affinité sur HisTrap	*Tampon A2'* 20mM Tris-HCl pH8.5, 500mM NaCl 10mM Imidazole 5 mM β-mercaptoethanol		*Tampon B2'* 20mM Tris-HCl pH8.5, 500mMNaCl 500mM Imidazole 5 mM β-mercaptoethanol	*Tampon D2'* 20mM HEPES-NaOH pH8.5, 150mM NaCl 0.5 mM TCEP

Tableau 6.4. Tampons utilisés au cours des différentes chromatographies

Chromatographie d'exclusion

La chromatographie d'exclusion a été utilisée comme technique analytique pour l'étude de l'état d'oligomérisation de BclB et de ses domaines mais aussi préparative, puisqu'elle constitue la dernière étape de purification avant les expériences de cristallogenèse de chaque protéine. La colonne est calibrée grâce au kit de protéines standard pour chromatographie par filtration (Amersham). La protéine est déposée sur une colonne d'exclusion stérique (HiLoad

16/60 Superdex^{TM}75, GE Healthcare) qui possède une gamme de séparation pour les protéines globulaires de 3 à 70 kDa. La colonne est équilibrée avec le tampon d'équilibration A1 pour la protéine entière et le domaine lectine et 20 mM Tris-HCl pH 8.5, 150 mM NaCl, (tampon C2) ou 20 mM HEPES-NaOH pH 8.5, 150 mM NaCl (tampon D2) pour le domaine *N*-terminal. La protéine est déposée dans un volume compris entre 0.5 et 5ml, avec un débit de 0.3 à 0.5 ml.min^{-1}.

Contrôle du poids moléculaire des protéines recombinantes purifiées

La masse moléculaire des protéines est vérifiée par spectrométrie de masse; 1 mg de protéine lyophilisée est dissous dans de l'eau pour être soumise à une analyse de type MALDI (Matrix Assisted Laser Desorption Ionisation).

6.1.4. Diffusion dynamique de la lumière

Les conditions biochimiques de stabilité (tampon, pH…) ainsi que la monodispersité des protéines sont contrôlées par diffusion dynamique de la lumière (Dynamic Light Scattering, DLS). Ces mesures sont réalisées avec un "Zetasizer Nano Series" équipé d'un laser HeNe (Malvern Instruments, UK). Un échantillon de 50µl de protéine pure est placé dans une cuve de mesure thermostatée à 22°C dans laquelle elle est soumise à un faisceau laser. Les molécules en solution présentent un mouvement aléatoire (mouvement Brownien) qui est rapide pour les petites molécules et plus lent pour les grosses particules. La mesure par DLS de la vitesse de diffusion des molécules en solution permet d'en déduire leur rayon hydrodynamique. Les résultats sont présentés sous la forme d'une courbe de l'intensité relative de la lumière diffusée en fonction du rayon hydrodynamique des particules. Cette valeur qui dépend entre autre de la masse et de la forme des particules en solution est directement proportionnelle à leur taille. Dans le cas des protéines, elle permet donc d'obtenir des informations sur la structure quaternaire des macromolécules étudiées.

6.1.5. Résonnance plasmonique de surface

Préparation du biocapteur

Le biocapteur CM5 (GE Healthcare) utilisé est formé de quatre canaux indépendants, contenant chacun une matrice de dextran carboxyméthyl. Après activation par un mélange de 1-éthyl-3-(3-diméthylaminopropyl) carbodiimide (EDC) et *N*-hydroxysuccinimide (NHS), les

groupements carboxyles sont utilisés pour immobiliser, de manière covalente, la streptavidine (Sigma-Aldrich, USA) à 100 µg/ml dans un tampon acétate à 10 mM, pH 5 (Biacore®) via un lien amide sous un flux de 5µl/min pendant 10 min. Une solution d'éthanolamine 1 M est ensuite injectée pour saturer les groupements activés restés libres. En moyenne, 4500 RU de streptavidine sont immobilisées.

Piste	Biocapteur CM5
1	α-D-galactose
2	α-D-fucose
3	α-D-galactose (90%) α-D-fucose (10%)
4	α-D-galactose (97%) α-D-fucose (3%)

Tableau 6.5. Ligands sur les canaux des deux biocapteurs

Des ligands sucre-polyacrylamide-biotine (PAA-sucre-biotine) (Lectinity Holdings, Inc. Moscou) différents sont fixés sur les quatre canaux des biocapteurs grâce à la forte interaction biotine-streptavidine (*Tableau 6.5*). Chaque PAA-sucre-biotine est dilué dans le tampon de course (10mM HEPES- NaOH pH7.5, 0.15 M NaCl, 0.5% Tween 20, 3 µM $CaCl_2$), à 20 µg/ml et injecté sous un flux de 5 µl/min pendant 10 min. Une réponse moyenne de 800 RU est obtenue. L'α-D-galactose, qui n'est pas reconnu par BclB, est utilisé comme blanc sur la piste 1. En mélange avec l'α-D-mannose et l'α-L-fucose, il sert également baisser la densité de récepteur de la surface reconnue par la lectine sur les autres canaux.

Conditions expérimentales

Deux types d'expérience ont été réalisés par SPR. Les paramètres d'interaction lectine-sucre ont été déterminés par des cinétiques d'interaction classiques où seule la concentration en protéines varie. Des cinétiques d'inhibition par compétition ont également été réalisées. Elles permettent de comparer la capacité d'une série de molécules à inhiber la liaison entre la lectine et le sucre immobilisé. La protéine à une concentration fixe est mélangée avec l'inhibiteur, de concentration variable et injectée à la surface du biocapteur. Toutes les conditions expérimentales sont rapportées dans le *Tableau 6.6*.

Conditions	BclB	
	cinétique	**inhibition**
Température	25°C	
Flux	20 µl/min	
Temps de contact	480s	600s
Temps de dissociation	300s	
Concentrations	protéines : de 0,392 à 19,6 µM	protéines : 7,84µM inhibiteurs : de 0,5 à 5000 µM
Solution de régénération	αMeFuc 1mg/ml EDTA 0,5M (1 : 1)	
Temps de régénération	1000s	540s
Temps de stabilisation	300s	

Tableau 6.6. Ligands sur les canaux des deux biocapteurs
Conditions expérimentales pour les cinétiques et les tests d'inhibition entre BclB et le PAA-Fuc. Pour les tests d'inhibition, la protéine est diluée dans leur tampon de course et mélangée avec les solutions d'inhibiteurs concentrées 10 fois selon un rapport 9 : 1 avant injection sur la surface.

Traitement des données

Les sensorgrammes sont traités avec le logiciel BIAevaluation, version AB. Le modèle utilisé correspond à une interaction simple 1 : 1. Les valeurs de RU obtenues sur la piste reconnue par la lectine sont soustraites à celles de la piste 1 avant analyse des résultats.

Le pourcentage d'inhibition de chaque molécule testée dans les tests d'inhibition est calculé à partir de la formule suivante *%I = 100 x (Réponse max – Réponse échantillon) / Réponse max*. Il permet de déterminer l'IC_{50}, concentration en inhibiteur nécessaire à provoquer 50% d'inhibition. Le potentiel d'inhibition de chaque molécule est alors calculé en divisant l'IC_{50} de la molécule inhibitrice par celui de la molécule de référence, le L-fucose.

6.1.6. Marquage des protéines pour les puces à sucre

Les lectines sont marquées avec le tetrafluoro-phenyl (TFP) ester grâce au réactif Alexa Fluor 488-TFP (Invitrogen). Les protéines sont préparées à une concentration de 1mg/ml dans un tampon bicarbonate de sodium à pH9. Le réactif TFP est quand à lui préparé extemporanément dans du DMF (Diméthylformamide) à une concentration de 10 mg/ml. Le mélange est incubé une heure à température ambiante sous agitation. La protéine est finalement séparée du réactif résiduel par colonne de dessalage sur colonne D-salt ™ polyacrylamide (Pierce). Les échantillons marqués sont ensuite envoyés au Consortium for Functional Glycomics, CFG (http://www.functionalglycomics.org/static/index.shtml).

6.1.7. Microcalorimétrie isotherme de titration

Les solutions de protéine et de sucre sont préparées dans le même tampon : 0.1 M Tris-HCl pH 7.5, 0.03 µM $CaCl_2$ de manière à avoir une concentration en sucre 10 fois supérieure à celle de la protéine. Des aliquots de 10µl de sucre sont injectés toutes les 5 min dans la cellule de mesure contenant la protéine thermostatée à 25°C. Chaque expérience consiste en une série de 30 injections.

6.1.8. Tests d'Hémagglutination

Détermination du point d'équivalence

L'activité hémagglutinante du domaine BclBNter17 est mesurée dans des microplaques de titration. La protéine est préparée dans un tampon 100mM HEPES-NaOH pH 7.5, 150 mM NaCl à une concentration de 1 mg/ml. Un volume de 50µl est déposé dans le premier puit, puis une gamme de concentrations est réalisée par dilutions en cascade de deux en deux dans les puits suivants. Les hématies de lapin (BioMérieux) sont diluées dans une solution saline à 150mM de NaCl pour avoir une concentration finale à 2%. Un volume de 25µl de ces érythrocytes est ajouté aux 25µl de protéine dans chaque puit. La lecture de l'activité hémagglutinante est réalisée après 30 minutes d'incubation à 37°C. La plus petite concentration en lectine pour laquelle une hémagglutination est encore visible est appelée point d'équivalence et la concentration en lectine de ce puit correspond à une unité d'hémagglutination (UH).

Inhibition de l'hémagglutination

Une concentration de lectine correspondant à 4 UH est sélectionnée pour les tests d'inhibition. Les différents sucres testés sont préparés à une concentration initiale de 30 $mg.ml^{-1}$ est préparés dans le tampon 100mM HEPES-HCl pH 7.5, 150 mM NaCl. Un volume de 50µl de sucre est déposé dans le premier puit, puis une gamme de concentrations est réalisée par dilutions en cascade de deux en deux dans les puits suivants. Un volume de 25µl de protéine est ajouté à chaque puit contenant 25µl de sucre à tester. Après incubation 30min à 37°C, puis 30min à température ambiante, 50µl d'érythrocytes préparés comme décrit précédemment sont ajoutés à chaque puit. L'inhibition de l'hémagglutination est lue après 30min d'incubation à température ambiante.

6.1.9. Cristallogenèse

La méthode utilisée ici est la diffusion de vapeur en goutte suspendue. Une goutte de solution concentrée en protéine est déposée sur une lamelle de verre siliconée et diluée au demi par une solution de précipitant. La lamelle est ensuite retournée et scellée avec de la graisse sur un puit contenant ce même agent précipitant.

Les expériences de co-cristallisation de BclB et BclBLec avec leur ligand saccharidique ont été effectuées en préparant des solutions protéine-sucre avec un ratio molaire 1 :5 pour la protéine et le sucre respectivement. Le criblage des conditions de cristallisation est réalisé avec des kits commerciaux disponibles au laboratoire, mais aussi grâce à un accès au Robot de cristallisation

6.2. Résultats et discussion

6.2.1. Production des protéines recombinantes

Choix des séquences

Le domaine BclBLec a une longueur de 115 acides aminés et sa limite du côté *N*-term a été choisie par comparaison avec la séquence de PA-IIL. Pour le domaine *N*-terminal, en absence de données structurales sur la limite entre le domaine et le linker, deux constructions de longueur différente, appelées BclBNter14 et BclBNter17, ont été réalisées. La séquence protéique de BclB ainsi que le choix de découpage en domaines lectine (BclBLec) et *N*-terminal (BclBNter14 et BclBNter17) sont présentés dans la *Figure 6.2*.

Figure 6.2. Séquences peptidiques des protéines recombinantes BclB et BclB tronquées.

Les résidus impliqués dans la liaison du calcium sont signalés par des sphères jaunes, ceux responsables de la spécificité de liaison au sucre sont représentés par des points d'exclamation verts.

Le gène codant pour BclB a été cloné avant le début de mes travaux de doctorat par le Dr. Michaela Wimmerova de l'Université Masaryk à Brno en République Tchèque. Le domaine *N*-terminal de BclB a quand à lui été cloné par le Dr. Daniel Aubert du Département de microbiologie et immunologie de l'université de Western Ontario au Canada. Le domaine lectine C-terminal (BclCLec) a été cloné au laboratoire dans le plasmide pRSET-a, comme détaillé ci-dessus

Expression des protéines

L'expression de BclB chez *E. coli* BL21 (DE3) a nécessité une phase d'optimisation. Différentes concentration en IPTG (0.05, 0.5 et 1 mM), plusieurs durées et température d'induction (20, 25 et 30°C de 4h à toute la nuit), ainsi que des DO_{600} d'induction de 0.5. 0.7 et 1 ont été testées mais n'ont pas permis de produire la protéine sous forme soluble. Ce problème a été résolu par l'utilisation de la souche d'*E. coli* BL21 (DE3) contenant le plasmide *pLysS*. Ce plasmide code pour le T7 lysozyme qui est un inhibiteur naturel de l'ARN polymérase du phage T7 limitant ainsi l'expression basale en absence d'induction et possède un gène de résistance au chloramphénicol (Cm).

Toutes les constructions de BclB ont pu être produites sous forme recombinante dans *E. coli* avec de bons rendements. Les protocoles d'expression (température, durée et concentration en IPTG), mis au point pour chacune des protéines sont résumés dans le *tableau 6.7.*

Antibiotique	Préculture	DO_{600} d'induction	IPTG (mM)	Température (°C)	Temps induction
Protéine BclB					
100µg/ml Amp 34µg/ml Cm	30ml à DO_{600} = 0.6	0.9 - 1	0.5	30	3 heures
Domaine BclBLec					
100µg/ml Amp	30ml à DO_{600} = 0.6	0.5 - 0.6	0.5	30	3 heures
Domaines BclBNter					
30µg/ml Km	30ml à DO_{600} = 0.6	0.6 - 0.7	0.05	30	3 heures
Domaines BclBNter Sélénié					
30µg/ml Km	30ml Sur la nuit	0.6 - 0.8	0.05	20	Sur la nuit

Tableau 6.7. Conditions d'expression des protéines BclB, BclBLec, BclBNter et BclBNter séléniée recombinantes.

Purification de BclB et de BclBLec

Les lectines précédemment étudiées dans notre laboratoire ont généralement été purifiées par chromatographie d'affinité sur colonne mannose-agarose et éluée par une solution concentrée de monosaccharide (Mitchell *et al.* 2005). Cette approche est possible pour BclB et son domaine lectine et nous avons vérifié qu'une solution de 100 mM de fucose permet d'obtenir la lectine purifiée. Cependant, cette méthode nécessite des étapes extensives de dialyse pour éliminer le ligand, ce qui peut provoquer une perte de qualité de la protéine.

Une méthode alternative a été utilisée. Après fixation sur colonne mannose-agarose et lavage par une solution contenant 1M de NaCl, les protéines sont éluées par une solution d'EDTA. L'utilisation de cet agent chélateur de cations divalents, permet d'éluer la lectine qui est liée aux sucres de la colonne par l'intermédiaire d'ions Ca^{++}. La concentration nécessaire et suffisante de 1 mM a été préalablement déterminée par un gradient de concentration.
Les profils de purification obtenus pour les deux lectines : BclB et BclBLec sont identiques pour les deux méthodes d'élution (100mM L-Fucose ou 1mM EDTA). Seul le chromatogramme de la purification de BclB avec élution par 1mM d'EDTA est présenté dans ce manuscrit (*Figure 6.3*).

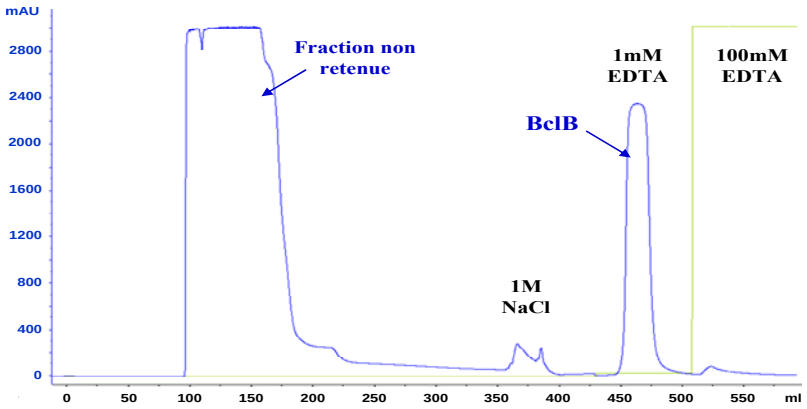


Figure 6.3. Purification de BclB sur colonne mannose-agarose

Chromatogramme de la purification de BclB sur colonne mannose agarose. En bleu, la DO_{280} (en mAU) en fonction du volume d'élution (en ml). En vert, le gradient en tampon d'élution : 1% puis 100% de tampon contenant 100mM d'EDTA. Le premier pic correspond à la fraction protéique non retenue, le second aux protéines interagissant de façon non spécifique et enfin le troisième correspond à l'élution de BclB par 1mM d'EDTA.

Les deux lectines BclB et BclBLec sont produites et purifiées en grande quantité avec un rendement final moyen respectif de 50 et 30mg de protéine par litre de culture. La pureté estimée par électrophorèse sur gel en conditions dénaturantes est proche de 100% (*Figure 6.4*).

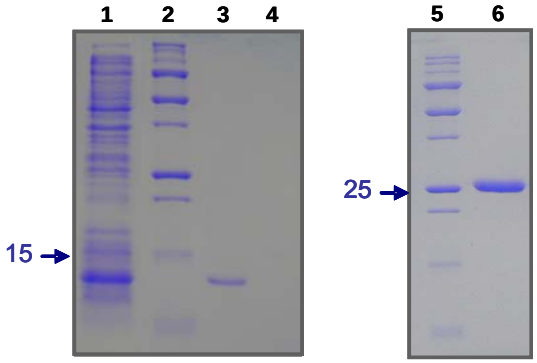


Figure 6.4. Electrophorèse sur gel de BclBLec (à gauche) et BclB (à droite) purifiées sur résine mannoses agarose et éluées par 1mM d'EDTA.

A. SDS-PAGE 15% de BclBLec. (1) extrait brut, (2) marqueur de poids moléculaire, (3) BclBLec purifiée, (4) fraction éluée par 1M NaCl.
B. SDS-PAGE 12% de BclB. (5) marqueur de poids moléculaire, (6) BclB

Production et purification du domaine N-terminal natif et sélénié

Les deux constructions du domaine *BclBNter*, contenant la séquence : MGSSHHHHHHSSGLVPRGSH supplémentaire en position *N*-term ont été exprimées chez *E. coli* BL21 (DE3). Ce domaine a également été produit en incorporant la sélénomethiones à la place de la méthionine afin de faciliter l'étude cristallographique. Les profils de purification obtenus sont très similaires pour ces différentes protéines (*Figure 6.5*).Après coupure à la thrombine, la protéine clivée est de nouveau purifiée sur colonne de Nickel. La protéine possède alors les trois acides aminés GSH supplémentaires en position *N*-term. Après une dernière étape d'échange de tampon, on obtient une protéine pure à 95%, avec un rendement final moyen de 20mg de protéine par litre de culture (*Figure 6.6*).

Après coupure à la thrombine, la protéine clivée est de nouveau purifiée sur colonne de Nickel. La protéine possède alors les trois acides aminés GSH supplémentaires en position *N*-term. Après une dernière étape d'échange de tampon, on obtient une protéine pure à 95%, avec un rendement final moyen de 20mg de protéine par litre de culture (*Figure 6.6*).

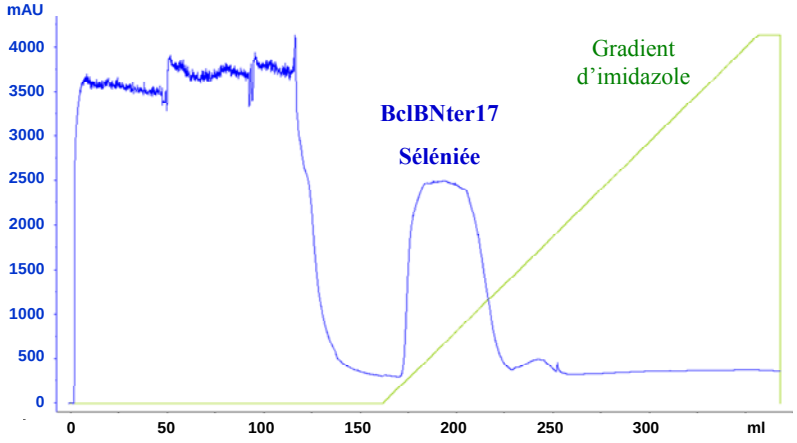


Figure 6.5. Purification de BclBNter17 Séléniée sur colonne de nickel

Chromatogramme de la purification de BclBNter17-1 séléniée sur colonne His-Trap. En bleu, la DO_{280} (en mAU) en fonction du volume d'élution. En vert, le gradient d'imidazole. Le premier pic correspond à la fraction protéique non retenue le second pic rassembles les protéines éluées par le tampon 20mM Tris-HCl / pH 8.5, 20mM, 500 mM NaCl, 10mM Imidazole, 5mM, β-mercaptoethanol.

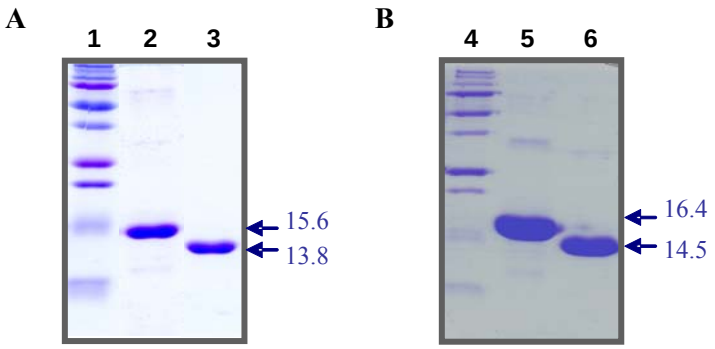


Figure 6.6. Electrophorèse sur gel de BclBNter17 (à gauche) et BclBNter14 (à droite) purifiées sur résine HisTrap.

A. SDS-PAGE 15% de BclBNter17. **B.** SDS-PAGE 15% de BclBNter14.
(1/4) marqueur de poids moléculaire, (2/5) après purification sur colonne de nickel - protéine éluée par 500mM d'imidazole, (3/6) protéine finale après coupure à la thrombine et une deuxième purification sur colonne de nickel - protéine non retenue par la colonne (Flow through).

Caractéristiques théoriques des protéines recombinantes

Les caractéristiques théoriques des différentes protéines produites pour cette étude sont rassemblées dans le *Tableau 6.8*.

	BclB	**BclBLec**	**BclBNter14**		**BclBNter17**		**BclBNter 17 seleniée**
			Avec étiquette His6	Sans étiquette His6	Avec étiquette His6	Sans étiquette His6	Sans étiquette His6
Nombre d'acides aminés	243	115	155	138	148	132	132
Masse moléculaire en Daltons	25429	12051	16408	14526	15559	13808	15698
Point isoélectrique	5.06	5.21	6.81	5.02	7.02	5.15	5.15

Tableau 6.8. Caractéristiques théoriques des protéines recombinantes

6.2.2. Etat d'oligomérisation de la lectine et de ses domaines en solution.

Les lectines bactériennes à deux calciums de la famille de BclB doivent être sous forme oligomérisée pour être actives. En effet la glycine en position *C*-terminal d'un monomère participe au site de liaison au sucre et au calcium de l'autre monomère. Les études cristallographiques ont montré que BclA est un dimère (Lameignere *et al.* 2008) alors que PA-IIL, CV-IIL et RS-IIL sont des tétramères. Toutes ces protéines consistent en un seul domaine lectine, et aucune donnée n'est disponible pour des lectines à deux domaines telles que BclB. La détermination de l'état d'oligomérisation pour la protéine entière et pour ses domaines est donc une étape nécessaire à son étude structure/fonction.

Détermination de l'état d'oligomérisation de BclB

La chromatographie d'exclusion de taille nous a permis d'étudier l'état d'oligomérisation de BclB en solution. La protéine a été chargée sur une colonne Superdex 75 16/60. Le chromatogramme montre qu'il existe trois populations protéiques distinctes dans l'échantillon (*Figure 6.7*). Avec un volume d'élution de 46.34 ml, le premier pic correspond à une masse moléculaire de 95 kDa, soit 3.7 monomère de BclB. La fraction protéique éluée à 49.65ml, correspondant à une masse moléculaire de 78.8 kDa c'est-à-dire 3.1 monomères de BclB. Enfin, le pic majoritaire à 56.42 ml correspondant à la masse moléculaire d'une protéine sphérique de 53.7 kDa, soit 2.1 monomères de BclB.

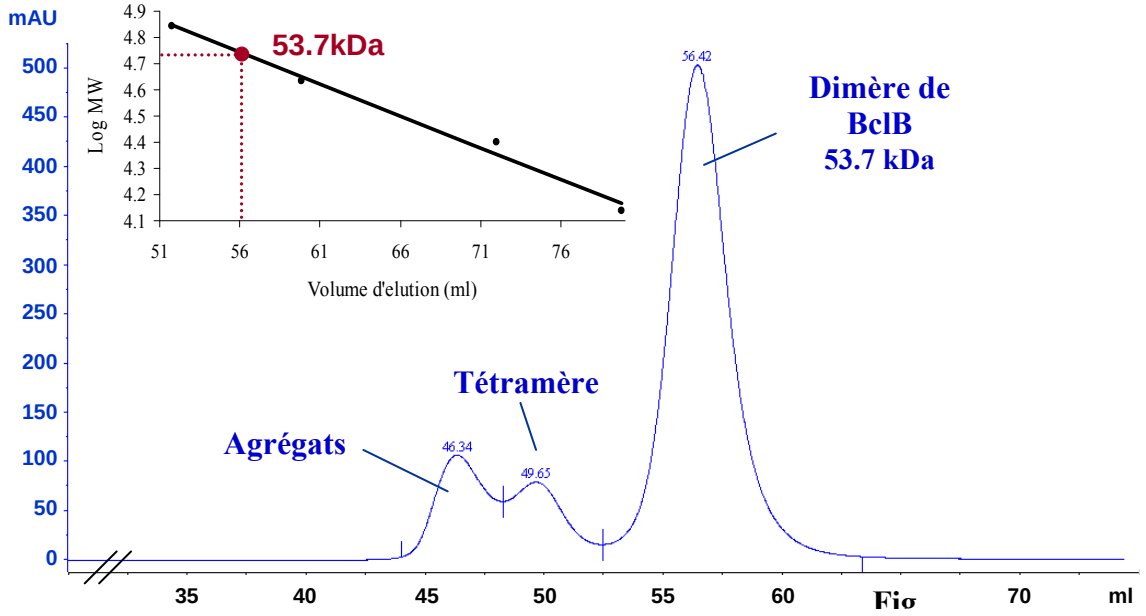


Figure 6.7. Détermination de l'état d'oligomérisation de BclB par chromatographie d'exclusion sur colonne Superdex 75 16/60

Après équilibration de la colonne avec un tampon 20mM Tris-HCl, 150mM NaCl, 100µM $CaCl_2$ / pH 7.5, la protéine BclB est chargée dans un volume de 2ml avec un débit de 0.5ml/min. En bleu, la DO à 280 nm (en mAU) en fonction du volume d'élution (en ml). Le premier pic, à 46.34ml correspond aux agrégats, le second pic à 49.65ml contient une fraction protéique organisée en tétramère et enfin le dernier pic à 56.42ml contient la protéine organisée en dimère.

Après conservation des échantillons 24h à 4°C, les fractions correspondant à chaque pic ont été analysées par diffusion de la lumière (DLS). Les résultats nous permettent d'estimer la taille des protéines en solution et donc d'obtenir des informations sur leur structure quaternaire. (*Figure 6.8*). Avec une taille de 383 kDa, la fraction protéique correspondant au premier pic contient de la protéine agrégée. Dans le deuxième pic, très minoritaire, la masse moléculaire de 87.5 kDa indique qu'une faible partie de la protéine est sous forme tétramèrique. Enfin, le volume hydrodynamique observé pour les protéines contenues dans le troisième pic correspond à la taille d'une protéine de 49kDa. Les résultats confirment donc que la protéine BclB éluée dans le troisième pic majoritaire est organisée sous la forme d'un homodimère. De plus, avec un indice de polydispersité de 0.040, cette fraction protéique stable à 4°C est monodisperse et peut être utilisée pour les essais de cristallogenèse. Ces expériences faites dans un tampon Tris à pH 7.5 ont été renouvelées avec un tampon MES à pH 6 et des résultats identiques ont été obtenus.

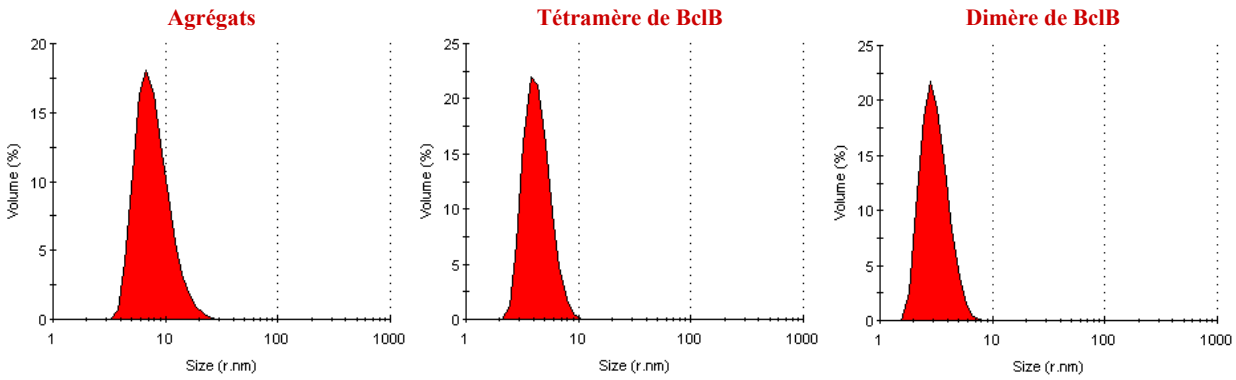


Figure 6.8. Analyse des fractions collectées en gel filtration par DLS

Distribution du rayon hydrodynamique en volume pour les fractions collectées par chromatographie d'exclusion de taille.

Détermination de l'état d'oligomérisation des domaines BclBLec et BclBNter

De la même manière que pour la protéine entière, l'organisation tridimensionnelle des domaines BclBLec et BclBNter a été étudiés par chromatographe d'exclusion. Les expériences ont été conduites sur BclBLec (12kDa), BclBNter14 (16.4 kDa), BclBNter17 (15.6 kDa) (non coupées à la thrombine) ainsi que sur la protéine BclBNter17 séléniée et débarrassée de l'étiquette poly-histidine (15.7kDa). Les résultats montrent que les deux domaines isolés sont probablement eux aussi organisés en dimère (*Figure 6.9*). Le volume d'élution de BclBLec correspond à la taille d'une protéine globulaire de 18.8 kDa, soit 1.6 monomère. L'élution de BclBNter14 et BclBNter17 à 61.49 et 63.09 ml respectivement correspond à des protéines globulaires d'une taille de 40.5 et 37 kDa soit à environ 2.4 monomères. Enfin, le domaine BclBNter17 sélénié sans queue His_6 est élué à un volume de 65.98 ml, correspondant à la taille d'une protéine de 31.5kDa, soit 2 monomères.

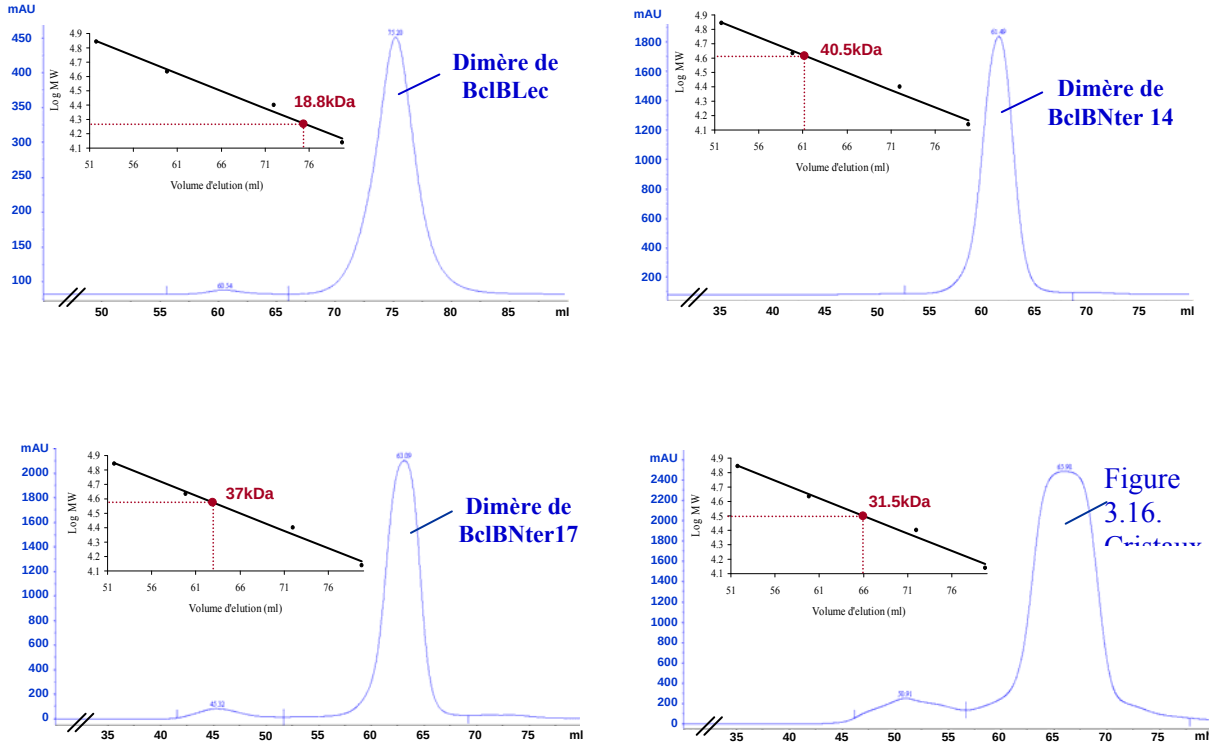


Figure 6.9. Détermination de l'état d'oligomérisation des domaines de BclB par chromatographie d'exclusion de taille sur colonne Superdex 75 1660

A. BclB, **B.** BclBNter14, **C.** BclBNter17, **D.** BclBNter17 SelMet.
En bleu, la DO_{280} (en mAU) en fonction du volume d'élution dans un tampon 20mM Tris-HCl/pH 7.5, 150mM NaCl, 100µM $CaCl_2$.

Les protéines BclBLec et BclBNter17 séléniée utilisées pour la cristallogenèse ont-elles aussi été analysées par DLS. Les résultats montrent que ces deux échantillons, stables à 4°C se trouvent sous forme monodisperse et sans agrégats avec une valeur d'indice de polydispersité (PDI) égale à 0.156 et 0.200 respectivement (*Figure 6.10*). De plus, avec une taille de 16.8 kDa pour BclBLec et 26.8 kDa pour BclBNter17SelMet, l'organisation en dimère de ces deux domaines semble être confirmée.

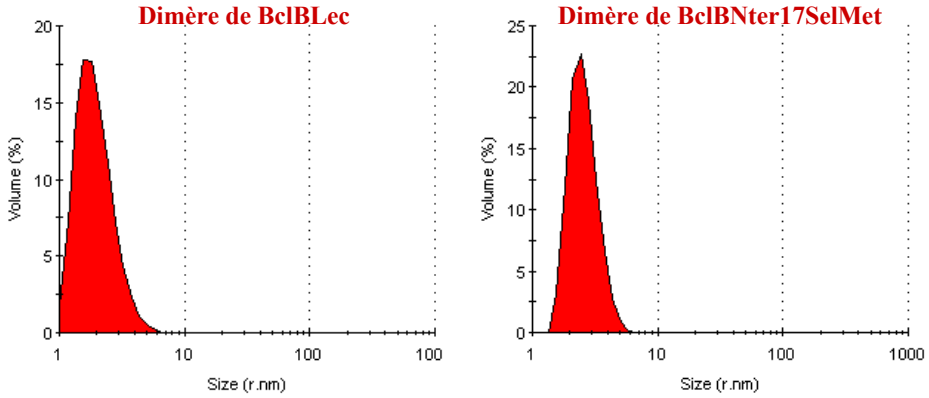


Figure 6.10. Analyse des fractions collectées en gel filtration par DLS

Distribution du rayon hydrodynamique en volume.

6.2.3. Etude de la spécificité de BclB pour les oligosaccharides

Etude de la spécificité de BclB par résonance plasmonique de surface (réalisée dans le cadre du stage de Master2 de Sophie Mathieu, en 2006-2007)

Les analyses de séquences montrent que le domaine lectine de BclB possède la séquence peptidique Ser-Ser-Ala dans le site actif, ce qui indique qu'elle est probablement spécifique pour le Fucose (Adam *et al.* 2007).

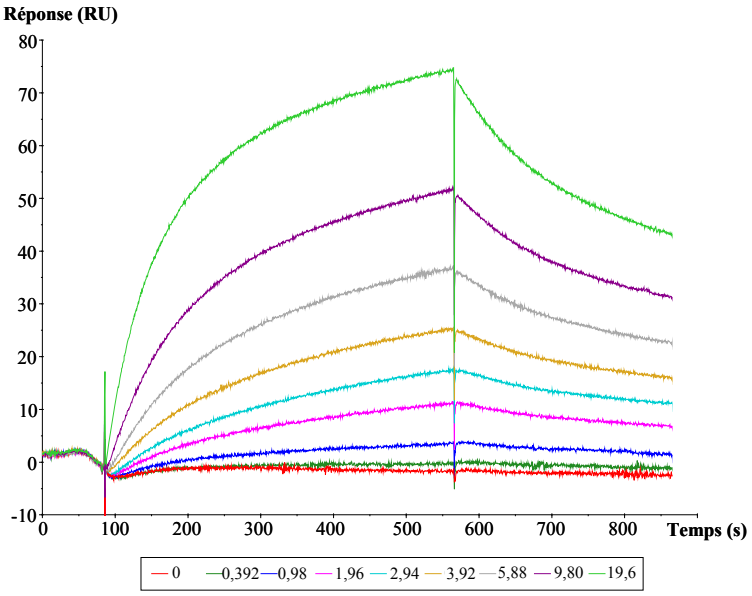


Figure 6.11 Sensorgrammes obtenus à partir de la cinétique réalisée avec BclB.

Ces courbes sont issues de la soustraction de la piste 4 contenant 3% de PAA-Fuc à la piste 1 témoin contenant 100% de PAA-Gal. Chaque couleur représente une concentration en protéine différente exprimée en µM.

Les premières expériences de SPR ont donc été réalisées en testant l'interaction de BclB avec le PAA-α-L-Fuc immobilisé à la surface du biocapteur. Une interaction forte et dépendante de la concentration en lectine a bien été observée. Les mesures de constantes d'affinité ont donné un Kd de 0.53 10^{-5}M (*Figure 6.11*), alors qu'aucune interaction n'est observée sur la surface de PAA-α-D-Gal qui peut donc servir de témoin.

Dans une deuxième étape, la spécificité de BclB pour différentes monosaccharides et oligosaccharides a été déterminée en évaluant la capacité de ces sucres à inhiber l'interaction entre BclB et le PAA-α-L-Fuc.. Le pourcentage d'inhibition permet de déterminer les valeurs d'IC_{50} de chaque inhibiteur. Le meilleur inhibiteur monosaccharidique de BclB est le fucose, mais d'autres monosaccharides peuvent inhiber la fixation à la surface, avec l'ordre d'efficacité suivant : L-fucose > D-fructose > D-arabinose (*Figure 6.12*).

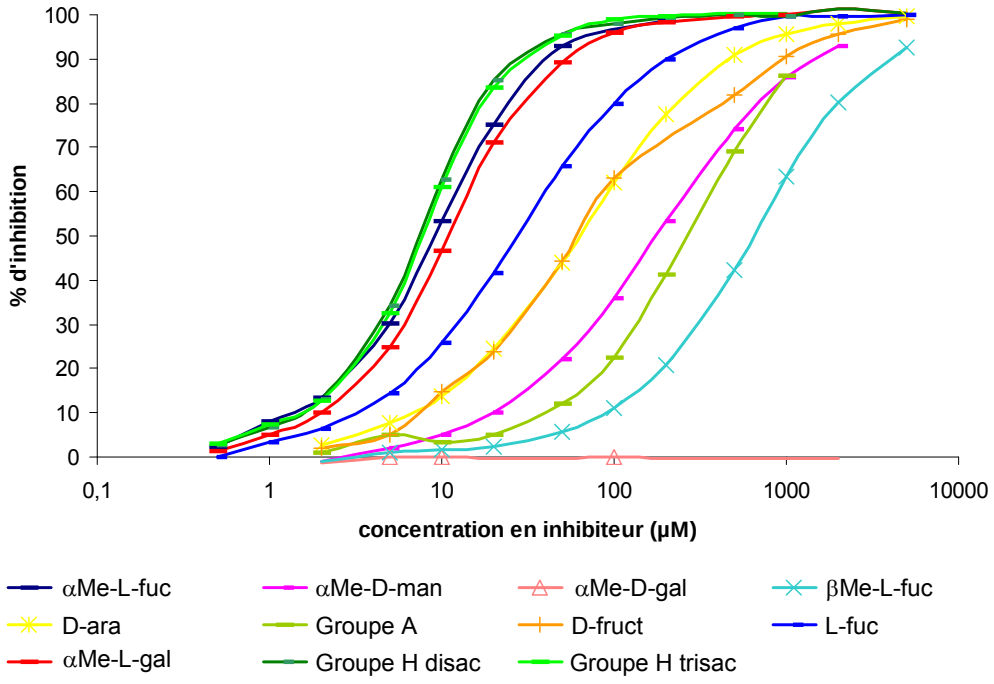


Figure 6.12. Expériences de compétition entre inhibiteurs mono-, di- et trisaccharidiques et la lectine BclB (c = 7,84 µM), évaluée par RPS.

Le methyl-fucoside d'anomérie α est 7 fois plus efficace que l'anomère β et 3 fois meilleur que le fucose non substitué (*Tableau 6.9*). L'αMe-L-galactoside est également un très bon ligand, sa valeur d'IC_{50} étant proche de celle obtenue avec l'αMe-L-fucoside. En revanche, l'αMe-D-galactoside ne provoque aucune inhibition, confirmant ainsi le choix de ce sucre comme référence pour les expériences de SPR sur Biacore®. L'αMe-D-mannoside est aussi un mauvais inhibiteur.

Inhibiteurs	IC50 (µM)	Potentiel d'inhibition
L-Fucose	28	1
αMe-L-Fucose	9,2	3,04
βMe-L-Fucose	70	0,40
D-fructose	80	0,35
D-arabinose	93	0,30
αMe-D-Mannose	180	0,16
αMe-D-Galactose	N.D.	N.D.
αMe-L-Galactose	12	2,33
Groupe sanguin H disac : Fucα1-2Gal	7,5	3,73
Groupe sanguin H trisac : Fucα1-2Galβ1-4Glc	7,9	3,54
Groupe sanguin A trisac : GalNAcα1-3(Fucα1-2)Gal	280	0,1

Tableau 6.9. Valeurs d'IC50 et potentiel d'inhibition des mono-, di- et trisaccharides testés.

Les lettres N.D. signifient non détécté.

Les oligosaccharides de l'épitope du groupe sanguin H possédant comme structure de base Fucα1-2Gal à laquelle s'ajoute un Glc lié en β1-4 pour le trisaccharide, présentent des valeurs d'IC_{50} proches de celles de l'αMeFuc. L'élongation de la chaîne oligosaccharidique n'améliore donc pas l'affinité. En revanche, la valeur d'IC_{50} du trisaccharide de l'épitope du groupe sanguin A est 30 fois plus faible que celle de l'αMeFuc. Ce ligand ramifié occupant un volume plus important qu'un ligand linéaire, provoque probablement un encombrement stérique qui nuit à une bonne liaison avec BclB.

Etude de la spécificité de BclB par puce à sucres

Les expériences en SPR ont confirmé que BclB est bien une lectine à fucose, qui tient très probablement ses propriétés du domaine lectine (*C*-term) très similaire à d'autres lectines bactériennes. Cette étude préliminaire a permis de déposer une demande pour utiliser les puces à sucre accessibles par le Consortium for Functional Glycomics, qui permettent d'évaluer l'interaction avec 374 composés glucidiques. Pour obtenir des résultats les plus clairs possibles, nous avons fait tester la lectine entière, mais également le domaine lectine (*C*-ter).

Les résultats obtenus pour la protéine BclB entière n'ont pas été très satisfaisants. Dans un premier essai la lectine a été utilisée à une concentration de 200 µg/ml ce qui a donné un marquage peu spécifique de tous les glucides contenant du mannose et du fucose. Un

deuxième essai à plus basse dilution a été décevant, probablement à cause du mauvais état de la protéine qui avait subi un temps de stockage trop long.

En revanche, des résultats de très bonne qualité ont été obtenus avec le domaine BclBLec à faible concentration (10 µg/ml). La lectine reconnaît spécifiquement les glucides contenant un résidu fucose. La plus petite unité monosaccharidique reconnue est le L-Fucose. (*Figure 6.13*). Ces données sont en excellent accord avec celles obtenues par SPR.

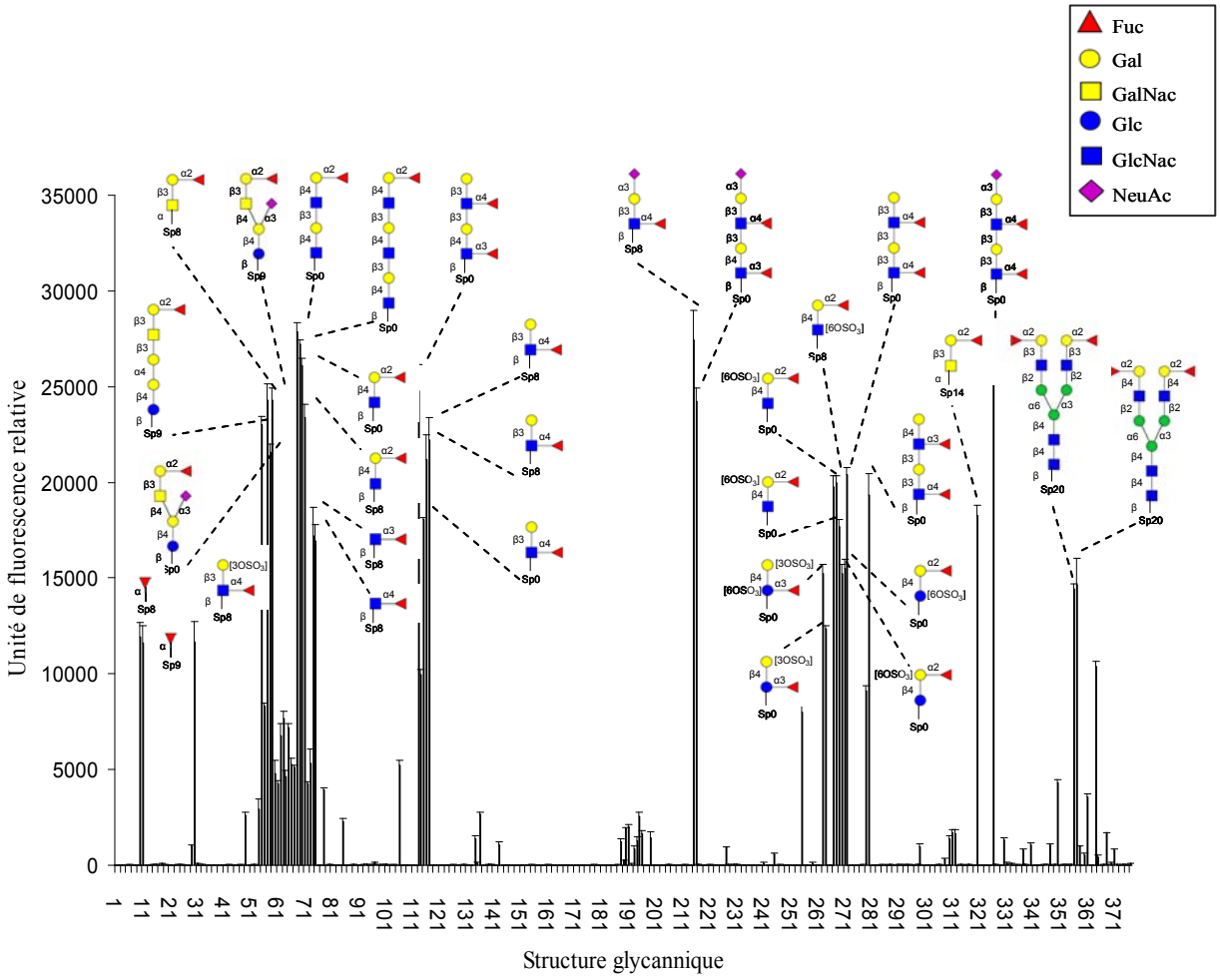


Figure 6.13. Résultats de printed arrays obtenus pour la protéine BclBLec à 10µg.ml^{-1}

Le disaccharide Fucα1-2Gal ne donne pas une interaction très forte mais certains trisaccharides linéaires contenant ce motif sont très bien reconnus tels que le H-type 1 (Fucα1-2βGal1-3βGlcNAc) ou les H-type 3 et H-type 4 (Fucα1-2βGal1-3GalNAcα/β). Par contre le H-type2 (Fucα1-2βGal1-4βGlcNAc) ou son analogue fucosyllactose (Fucα1-2βGal1-4Glc) sont moins bien reconnus. Pour le H-type 1 et H-type 3/4, les substitutions par

un galactose ou un GalNAc (groupe sanguin B et A) ne sont pas tolérées comme cela avait été déjà observée en SPR.

Le trisaccharide branché Lewis a (Galβ1-3[Fucα1-4]GlcNAc) est également un excellent ligand de BclB. Le tétrasaccharide sialylé (sialyl Lewis a) se lie de manière comparable alors que le tétrasaccharide difucosylé de la même famille (Lewis b) n'est pas reconnu. De même l'analogue Lewis x Galβ1-4[Fucα1-3]GlcNAc et ses dérivés n'interagissent pas avec BclB.

La lectine a donc une activité proche de celle de PA-IIL qui est spécifique pour le Lewis a (Perret *et al.* 2005) mais présente des différences puisqu'elle peut également reconnaître les trisaccharides H-type 1 et H-type 3.

6.2.4 Mesure de l'affinité de BclB pour l'αMeFuc

L'affinité pour l'αMethyl-L-Fucoside (αMeFuc) de BclB entière et du domaine BclBLec isolé a été étudiée par microcalorimétrie isotherme de titration. Les courbes sont mesurées dans les deux cas pour des concentrations en sucres dix fois plus élevées que celles en protéine à une température de 25°C. La protéine BclB présente une forte affinité pour l'αMe-L-Fucoside, avec une constante de dissociation de $2.82.10^{-6}$M. Le domaine lectine présente une affinité très similaire avec une constante de dissociation de $2.80.10^{-6}$M. Les constantes d'affinité et grandeur thermodynamiques mesurées sont récapitulées dans le *Tableau 6.10*. Le domaine *N*-terminal ne semble donc pas intervenir dans l'affinité de BclB pour ce ligand monosaccharidique.

	K_a (10^5 M^{-1})	K_d (10^{-6} M)	n	-ΔG (kJ/mol)	-ΔH (kJ/mol)	-TΔS (kJ/mol)
BclB	3.54 (± 0.06)	2.82 (± 0.05)	0.89 (± 0.11)	31.67 (± 0.04)	31.27 (± 0.39)	-0.41 (± 0.42)
BclBLec	3.65 (±0.55)	2.80 (±0.43)	0.98 (±0.01)	31.72 (± 0.38)	33.45 (± 3.55)	1.73 (± 3.17)

Tableau 6.10. Mesure de l'affinité de BclB pour l'αMeFuc par microcalorimétrie isotherme de titration selon le modèle à un site de fixation.

Les résultats sont moyennés sur trois et deux expériences pour BclB et BclBLec respectivement. Les déviations standards sont indiquées entre parenthèses.

Du point de vue des contributions énergétiques, l'interaction de BclB ou BclBLec avec l'αMeFuc est de nature clairement enthalpique, avec une contribution entropique presque nulle. L'absence du domaine *N*-terminal ne semble donc pas affecter la signature énergétique de l'interaction avec l'αMeFuc. La variation d'enthalpie, ΔH, très favorable est caractéristique de la formation de liaisons hydrogène entre le ligand saccharidique et la protéine. La contribution entropique est presque nulle dans le cas de ce monosaccharide. Un tel profil thermodynamique n'est pas caractéristique des interactions protéine-sucre qui montrent en général une contribution entropique très défavorable (TΔS < 0) (Dam and Brewer 2002), mais il correspond aux observations précédentes sur cette famille de lectine (Pokorná *et al.* 2006). Il semble que la présence de deux ions calicum entraîne une variation d'entropie moins défavorable.

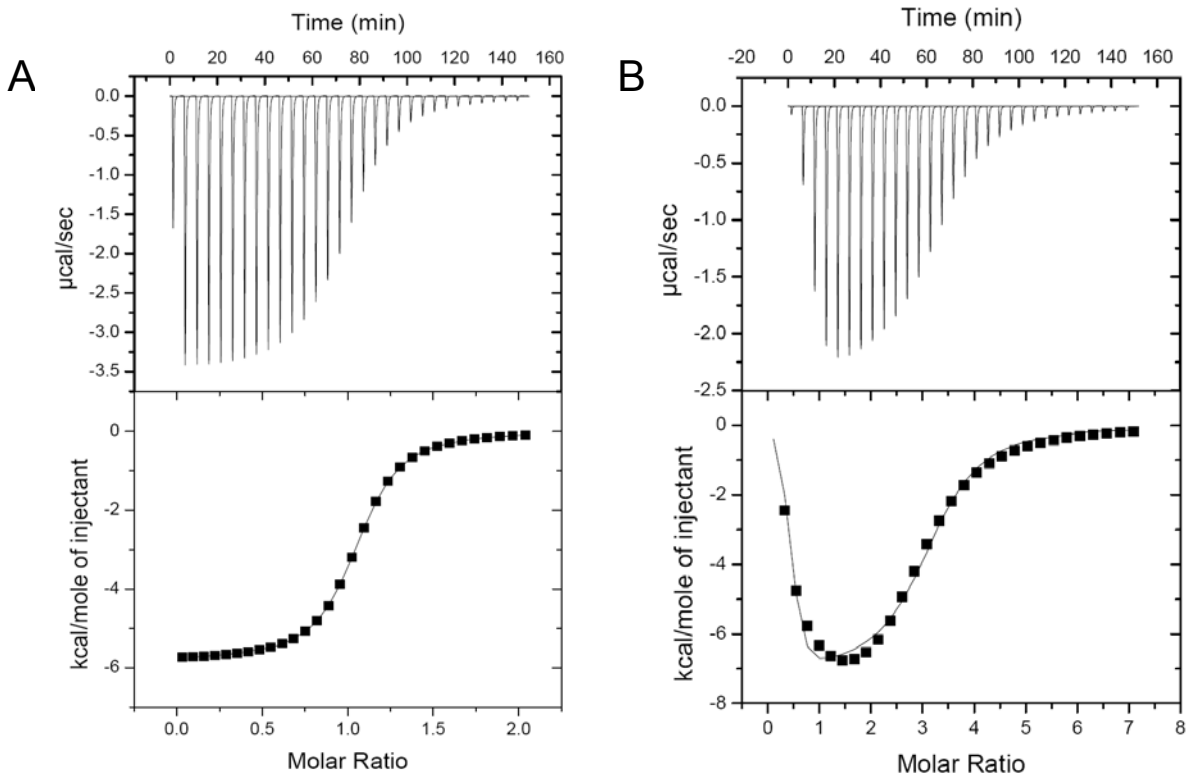


Figure 6.14. Courbes de microcalorimétrie obtenues pour BclB et BclBLec avec l'αMeFuc

L'expérience est réalisée à 25°C avec des concentrations en ligand 10 fois supérieures à celles en protéine.
A. Interaction BclB – αMeFuc pour 0.3mM de lectine. La courbe est ajustée avec le modèle théorique à un site. **B.** Interaction BclBLec – αMeFuc pour 0.12mM de lectine. La courbe est ajustée avec le modèle théorique à deux sites coopératifs.

En revanche, l'allure générale des courbes enregistrées au cours du temps est très différente pour BclB entière et pour le domaine lectine. Le thermogramme de l'interaction αMeFuc avec BclB peut être interprété avec un modèle classique à un site de liaison (*Figure 6.14.A*). Par contre la courbe expérimentale obtenue pour le domaine lectine de BclB montre une forte augmentation de la chaleur libérée pendant les 5 premiers pics (*Figure 6.14.B*). Ce profil de courbe, associé à un phénomène coopératif a été décrit pour la lectine BclA (chapitre précédent). Si on traite le thermogramme BclCLec/αMeFuc par un modèle à 2 sites coopératifs, on obtient des constantes de dissociation de $3.4.10^{-4}$ M et $7.7.10^{-7}$ M pour les sites 1 et 2 respectivement.

Pour BclA, l'hypothèse d'un mouvement de boucle lors de la liaison du ligand a été proposé pour expliquer le phénomène coopératif. Un tel modèle est en accord avec le fait que le domaine lectine isolé de BclB se comporte de la même façon, alors que la lectine entière, probablement rigidifiée par l'existence du domaine supplémentaire ne présente pas de coopérativité. Le dimère de domaine lectine étant probablement moins stable que le dimère de lectine entière, on peut supposer que la fixation du premier sucre stabilise le dimère, et donc le deuxième site actif et aide ainsi à la fixation du deuxieme sucre.

6.2.5 Etudes fonctionnelle du domaine N-terminal

Le domaine *N*-terminal de BclB ne possède aucune homologie avec une protéine de fonction connue. Il pourrait s'agir d'un deuxième domaine lectine, d'une enzyme, ou de tout autre chose.

Nous avons testé l'activité glycosylhydrolase sur un panel de dérivés saccharidiques de type para-nitro-phenyl glycsoside, classiquement utilisés pour tester les activités exo-glucanases. Après une incubation d'une nuit aucune activité enzymatique n'a été détectée.

Afin de tester la possibilité d'une activité lectine pour ce domaine, des expériences d'hémagglutination sur des érythrocytes natifs de lapin ont été réalisées. Les résultats montrent que cette protéine possède des propriétés d'hémagglutination. Avec une concentration de départ définie à $4mg.ml^{-1}$, l'unité d'hémagglutination (UH), correspondant au facteur de dilution à partir duquel l'hemagglutination n'est plus observée est de $7.8\mu g.ml^{-1}$ (*Figure 6.15*).

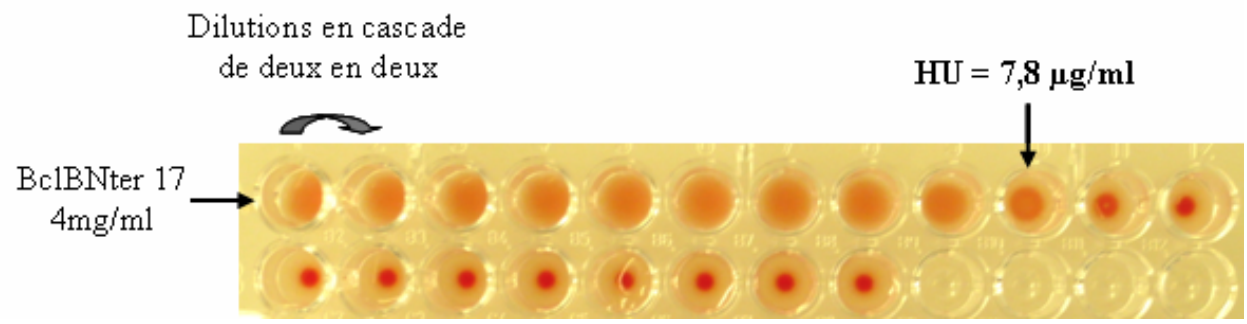


Figure 6.15. Mises en évidence des propriétés hémagglutinantes de BclBNter17

Des tests d'inhibition d'hémagglutination en présence de différents sucres ont été effectués. Les monosaccharides suivants: D-mannose, L-fucose, D-Galactose, D-glucose, D-Arabinose, L-Rhamnose, D-N-acétylgalactosamine (GalNac), D-N-acétylglucosamine (GlcNac) et l'acide N-acétylneuraminique ont été testés. Malheureusement, aucun inhibiteur n'a pu être mis en évidence par cette technique. L'expérience a été renouvelée plusieurs fois et les résultats obtenus pour trois sucres (L-Rhamnose, L-Fucose et D-Mannose) sont présentés ici à titre d'exemple (*Figure 6.16*). Les expérience d'hémagglutination ont montré les propriétés agglutinantes de BclBNter, mais n'ont pas permis de mettre en évidence les ligands de la protéines.

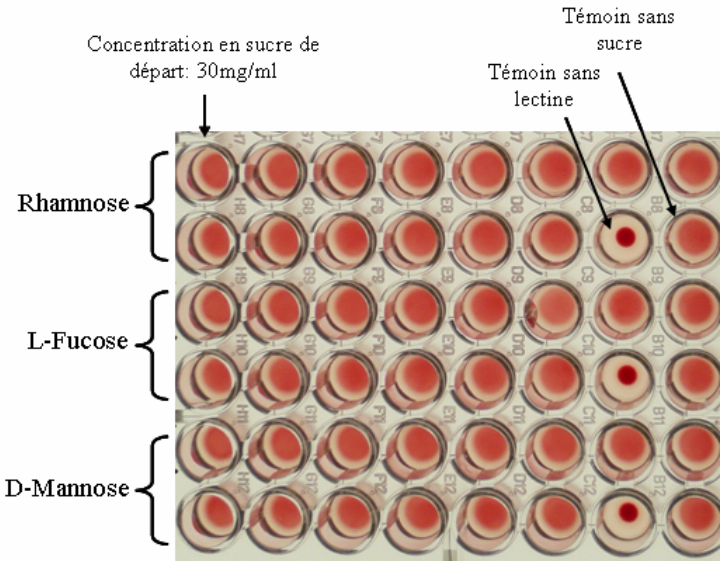


Figure 6.16. Tests d'inhibition d'hémagglutination de BclBNter avec le D-man, le L-Fuc et le Rham.

Les concentrations en sucre de départ est de 30mg/ml. Les autres concentrations sont obtenues par dilutions en cascade de deux en deux. Les deux derniers puits de chaque série sont les témoins négatifs sans protéine et sans sucre. La concentration en protéine correspond à la valeur de 4UH

Le domaine *N*-term a également été testé sur les oligosaccharides des Printed Array disponibles au Consortium for Functional Glycomics afin de déterminer si des sucres complexes pouvaient être reconnus par ce domaine. Les résultats (*Figure 6.17*) montrent que là encore les structures Fucα1-2Gal et NeuAcα1-3Gal semblent être les structures les motifs les mieux reconnues par ce domaine. Cependant, si on prend en compte que ces données ont été obtenus avec de forte concentration en protéine (200µg/ml) et que les intensités de fluorescence obtenues sont relativement faible, il se pourrait ces résultats ne soient pas significatifs mais résultent d'une contamination par l'autre domaine. Il sera nécessaire de les confirmer ou de les invalider avec d'autres approches.

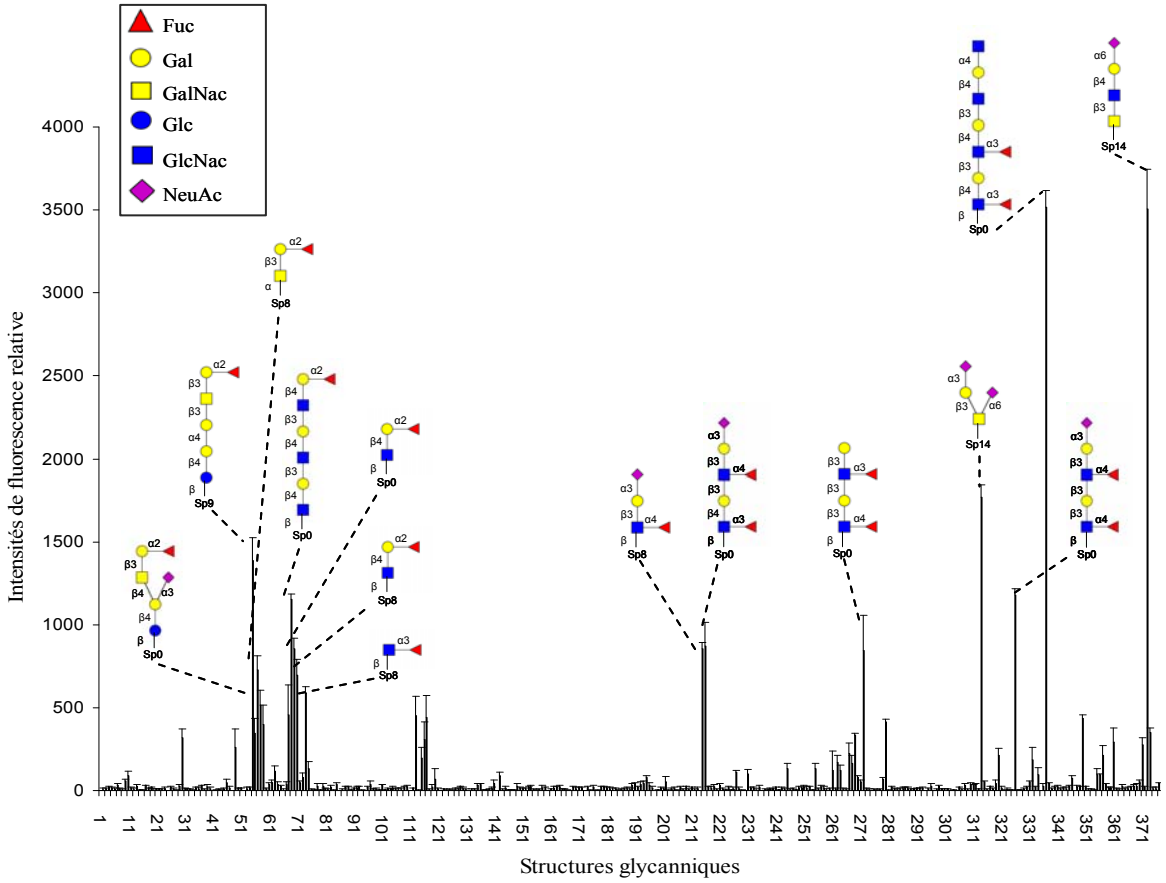


Figure 6.17. Résultats de printed arrays obtenus pour le domaine BclBNter17 à 200µg/ml

6.2.6. Cristallogenèse

Les expériences de cristallogénèse de la protéine BclB et des domaines BclBNter et BclBLec, ont été réalisées avec différents kits commerciaux, CSS-I et CSS-II (Clear Strategy Screen de Molecular Dimension Ltd), kit Cristal Screen 1 et 2 (Hampton Research®) et JBScreen2 4 et 5 (Jena Bioscience). Les expériences ont été répétées pour différentes concentrations en protéines (de 5 à 20mg.ml^{-1}) et des expériences de co-cristallisation de BclB et BclBLec avec l'αMeFuc ont été réalisées dans les même conditions. Malheureusement, aucune de ces conditions n'a permis d'obtenir des cristaux de protéine.

Un accès à la Plateforme de cristallisation automatisée des protéines de l'l'EMBL (http://www.ibs.fr/content/ibs/presentation/Platform/plateform_Xtallo/), nous a permis de cribler de nombreuses condition de cristallogenèse pour BclB et les deux domaines BclBLec et BclBNter. Plusieurs conditions permettant l'obtention de cristaux de BclBNter17 ont ainsi pu être déterminées (*Figure 6.18*).

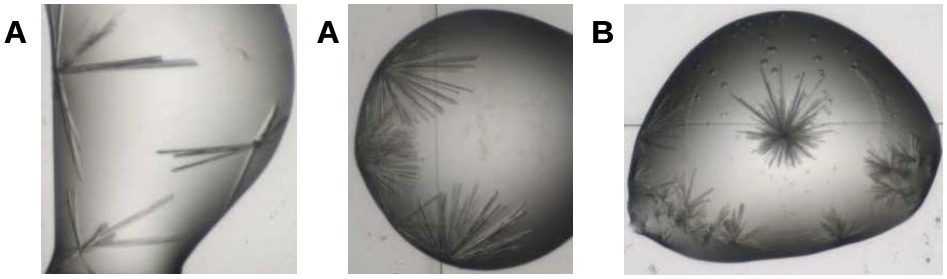


Figure 6.18. Cristaux de BclBNter17 séléniée obtenus dans différentes conditions.

Les cristaux sont obtenus pour des concentrations de 15mg/ml de protéine dans les conditions suivantes **A.** 0.05M Hepes-NaOH / pH7, 0.01M Magnesium sulfate, 1.8M Lithium sulfate. **B.** 0.1M Tris-HCl / pH8.5, 0.1M di-Ammonium phosphates, 0.5M di-Sodium phosphate, 0.5M di-Potassium phosphate. **C.** 0.1M Hepes-NaOH / pH 7.5, 1.0M tri-Sodium citrate.

Les premières collectes à l'ESRF ont permis de vérifier l'incorporation du sélénium dans la protéine Jusqu'à présent, aucun de ces cristaux n'a permis de collecter un jeu de données de données de bonne qualité, mais les expériences d'optimisation seront poursuivies.

6.3. Conclusions

Les résultats concernant BclB exposent la mise au point des stratégies d'expression et de purification des protéines recombinantes (BclB et ses deux domaines BclBLec, BclBNter). Ces protéines ont été produites et purifiées en quantité et avec une pureté suffisante pour envisager les essais de cristallogenèse.

Malheureusement, aucun cristal de la protéine entière n'a été obtenu. Les expériences de microcalorimétrie montrent que la lectine est active et la monodispersité de la protéine a été vérifiée par DLS. Les difficultés rencontrées dans la cristallogenèse ne sont donc pas dues à la qualité de la protéine recombinante, mais probablement à la flexibilité de la lectine composée de deux domaines.

Très récemment, des premiers cristaux du domaine *N*-terminal (BclBNter) ont été obtenus. Malgré les essais d'optimisation ces cristaux ne sont pas de qualité suffisante pour permettre la collecte d'un jeu de données. Des étapes d'optimisation sont donc encore nécessaires pour améliorer ces conditions.

Sur un plan fonctionnel, l'activité de lectine de cette protéine a été confirmée. Avec une affinité de 2.82 10^{-6}M, BclB est une lectine à Fucose. Les expériences de printed array ont montré qu'elle reconnaît en particulier les oligosaccharides fucosylés H-type 1, H-type 3 et H-type 4 ainsi que le Lewis a présents à la surface des cellules.

Le rôle du domaine *N*-terminal reste inconnu. Nous avons montré qu'il est capable d'agglutiner les érythrocytes, mais aucun bon inhibiteur n'a pu être identifié par la technique d'hémagglutination. Les expériences de « printed array » ont montré une interaction de la lectine avec les Fucα1-2Gal et NeuAcα1-3Gal. Mais la faiblesse du signale laisse penser que ces résultats ne sont pas significatifs.

Chapitre 7

Localisation et fonction des lectines solubles de *B. cenocepacia*

Les fonctions des lectines solubles identifiées chez les bactéries opportunistes *P. aeruginosa*, *R. solanacearum* et *C. violacearum* ne sont pas encore clairement établies. Pour comprendre leur rôle biologique, il est tout d'abord important de connaître la localisation cellulaire de ces protéines qui ne présentent pas de signaux de sécrétion. La détermination de leurs ligands naturels (glycanne de l'hôte, alginates de biofilm ou LPS bactérien) est également un élément important. Enfin, dans l'étude de la fonction des lectines, la construction de mutants déficients dans la production de ces protéines constitue un outil primordial.

N'ayant pas la possibilité de cultiver *B. cenocepacia* au CERMAV (nécessité d'un laboratoire P2), une collaboration a été mise en place avec le Professeur Miguel Valvano du "Department of Microbiology and Immunology" de l'université de Western Ontario. Au cours d'un séjour de trois mois dans ce laboratoire, partiellement financé par le Centre Jacques Cartier, des expériences de fractionnements cellulaires ont pu être entreprises. Un mutant délété des trois gènes codant pour les lectines BclA, B et C a été construit et sa virulence chez *Caenorhabditis elegans* a pu être étudiée après mon départ par le Docteur Daniel Aubert. Les tests mesurant la capacité de ce même mutant à former des biofilms grâce au Biofilm Ring Test ont été effectués dans l'équipe de Thierry Jouennes, au laboratoire Polymères, Biopolymères, Membranes de la Faculté des sciences à Rouen.

7.1. Matériel et méthodes

7.1.1. Clonages

Matériel biologique

Les souches bactériennes et plasmides utilisés dans cette étude sont listés dans le *Tableau 7.1 ;* les amorces nucléotidiques sont quand à elles répertoriées dans le *Tableau 7.2*.

Souches - Plasmide	Caractéristiques importantes▲	Référence
Plasmides		
pDA-12	ori_{pBBR1}, Tet^R	D. Aubert
pBADNTF	ori_{ColE1}, Flag (pour fusion N-terminale), Amp^R	C. Marolda
pDA17	ori_{pBBR1}, Flag (pour fusion C-terminale), Tet^R	D. Aubert
pEL-1	ori_{pBBR1}, Flag (pour fusion N-terminale), Tet^R	Cette étude
pRK2013	$ori_{colE1,}$ RK2 derivative, Km^R, mob^+, tra^+	(Figurski and Helinski 1979)
pGpISceI	$ori_{R6K,}$ ΩTp^R , mob^+, comprenant un site de restriction ISceI	(Flannagan, *et al.* 2008)
pDAIsceI	pDA12 exprimant l'endonuclease ISceI	(Flannagan, *et al.* 2008)
pBclA	pEL1-$BclA_{Flag}$	Cette étude
pBclB	pEL1-$BclB_{Flag}$	Cette étude
pBclC	pEL1-$BclC_{Flag}$	Cette étude
pDelBcl	pGpISceI avec fragments flanquants l'opéron contenant *bclA*, *bclB* et *bclC*	Cette étude
pRF132	pDA17-$HtrA_{Flag}$	(Flannagan, *et al.* 2008)
B. cenocepacia		
K56-2	Isolat clinique de patient CF	BCRRC*
Δlec	K56-2, Δ*bclA,B,C*	Cette étude
E. coli		
DH5α	F^-, Φ80 *lacZ*ΔM15 (Δ*lacZYA-arg*F)U169 *endA1 recA1 hsdR*17(rK^-mK^+) *supE*44 *thi*-1 Δ*gyrA*96 *relA*1	Stock laboratoire
SY327	*araD,* Δ*(lac pro) argE(Am) recA56* rif^R *nalA,* λ *pir*	Stock laboratoire

Tableau 7.1. Matériel biologique utilisé dans cette étude

▲Tp, triméthoprime; Km, kanamycine; Amp, ampicilline; Tet, tetracycline ; Gen, gentamycin ; Rif, rifampicine ; HtrA, protéine périplasmique impliquée dans la réponse au stress chez *B. cenocepacia* ; FLAG, épitope (DYKDDDDK) reconnu par les anticorps antiFlag : M2.

* BCRRC, *B. cepacia* Complex Research and Referral Repository for Canadian CF Clinics

Les réactions en chaîne par polymérase (PCR) sont faites sur l'appareil CTP-0200 ADN moteur (MJ Research) avec la ProofStart polymérase ou la Taq polymérase (Qiagen).

Pour les PCR faites sur l'ADN de *B. cenocepacia*, la "Q solution" (Qiagen) est ajoutée au milieu réactionnel selon les instructions du fabricant. Les séquençages sont réalisés par le" Robarts Research Institute DNA sequencing Facility" de l'université de Western Ontario. Les bactéries *B. cenocepacia* et *E. coli* sont cultivées à 37°C en milieu Luria Bertani (LB) ou LB agar solidifié avec 1.6% (w/v) de Bacto agar. La conjugaison des plasmides chez *B. cenocepacia* K56-2 est faite grâce à un système triparental utilisant une souche *E. coli* portant le plasmide helper pRK2013.

Construction du plasmide pEL1

Afin de pouvoir détecter les lectines clonées et exprimées chez *B. cenocepacia*, un plasmide pEL-1 a été construit à partir de pDA-12 et pBADNTF (*Figure 7.1*).

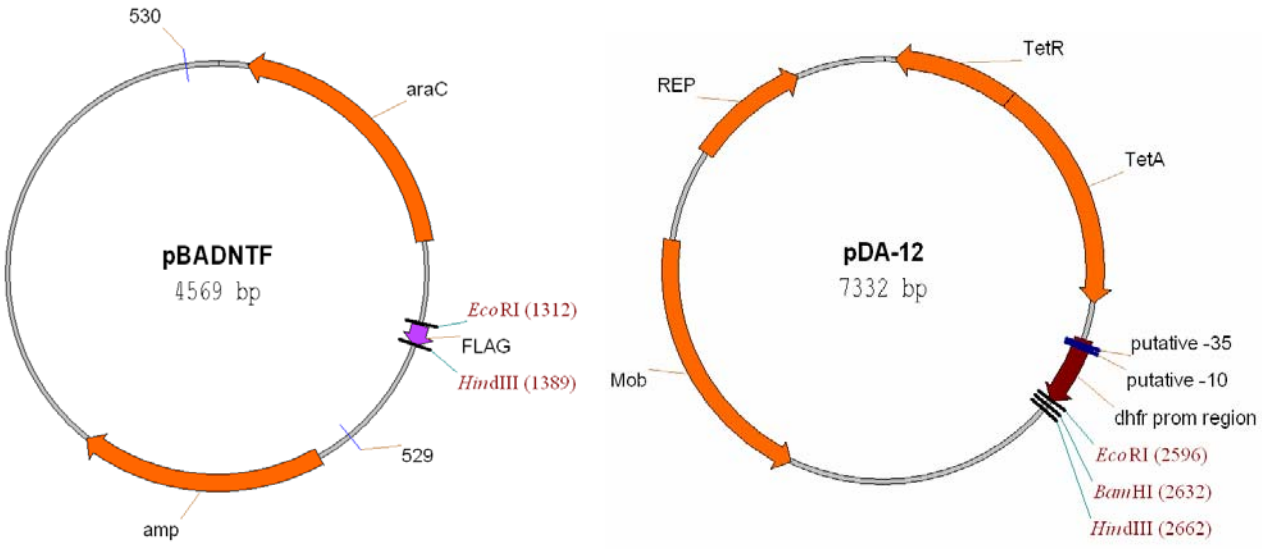


Figure 7.1. Représentation schématique des vecteurs pBADNTF (à gauche) et pDA-12 (à droite)

pBADNTF : *Amp*, gène de résistance à l'ampicilline ; *AraC*, activateur transcriptionnel en présence d'arabinose ; *FLAG*, code pour l'épitope (DYKDDDDK) reconnu par les anticorps anti-Flag : M2.
pDA-12 : *TetR* et *TetA* confèrent la résistance à la Tétracycline; REP, origine de réplication à large spectre d'hôte; *Mob*, gène de mobilisation [pDA-12 est mobilisable (mob+)].

La séquence codant pour l'étiquette « Flag » est amplifiée par PCR grâce aux amorces 529 et 530 (*Tableau 7.2*) en utilisant comme matrice le plasmide pBADNTF. Le produit PCR obtenu est digéré par les enzymes de restriction *EcoRI* et *HindIII* puis le fragment de 83 paires de bases contenant la séquence de l'étiquette « Flag » est ensuite extrait sur gel d'agarose 4% grâce au kit QIAquick Gel Extraction (Qiagen).

Le plasmide (pDA-12) est également digéré par les enzymes de restriction *EcoR1* et *HindIII* puis purifiés avec le kit QIAquick Gel Extraction. Le vecteur pDA-12 est ensuite déphosphorylé par traitement à l'Antarctic Alkaline phosphatase (New England Biolab, NEB) pendant 1h30 a 37°C puis inactivation de l'enzyme (10 min à 65°C). L'insert codant pour l'étiquette Flag est introduit dans pDA-12 par ligation avec la T4 DNA ligase (Roche) pendant 16h a 4°C. La souche d'amplification DH5α d'*E. coli* compétente est transformée avec le produit de ligation par choc thermique 2 minutes à 42°C. Les bactéries contenant le plasmide recombinant sont sélectionnées sur tétracycline 30µg/ml. Après vérification de la présence de l'insert par PCR sur colonies, le plasmide d'un clone positif est extrait à l'aide du QiaPrep spin Miniprep kit (Qiagen) et envoyé au séquençage.

Amorces	Séquence	Sites de restriction
529	AAGAGTTTGTAGAAACGCAAAAAGGCCATCCGTCAGG	
530	CCCATCGGTGATGTCGGCGATATAGGC	
3221	TTTTGGATCCTTCACAACCCTTTACCCACGA	*BamH1*
3215	TTTTAAGCTTGTGATGTAACGGCGAAGACC	*HindIII*
3217	TTTTGGATCCTCCCCTCCTTTCGGCTTCGAT	*BamH1*
3224	TTTTAAGCTTGTACAGCAGTGGGACTGCAA	*HindIII*
3220	TTTAGGATCCTGCTGATTCTCAAACGTCATCCA	*BamH1*
3218	TTTTAAGCTTAACGGCGTCAGGTCAGC	*HindIII*
3230	TTTTCTCGAGAAGACCTACTTCGGGCTGGT	*Xho1*
3234	TACGTCTAGAATCAGTCCAGTCCGATAGGC	*Xba1*
3231	TTCTCTCGAGAGCCATGCTTTTCTCCTGTG	*Xho1*
3232	TTTTGAATTCGGATGGAACAACGGCAATAC	*EcoR1*

Tableau 7.2. Amorces utilisées au cours de cette étude

Les séquences soulignées correspondent aux sites de restriction

Clonage des lectines possédant une étiquette N-terminale chez B. cenocepacia K56-2

Les gènes codant pour les trois lectines : *Bcam0186* pour BclA, *Bcam0184* pour BclB et *Bcam0185* pour BclC sont amplifiées par PCR grâce aux amorces 3220/3218, 3221/3215 et 3217/3224 respectivement (*Tableau 7.2*). L'ADN de *B. cenocepacia* K56-2 disponible à l'université de Western Ontario (BCRRC, *B. cepacia* Complex Research and Referral Repository for Canadian CF Clinics) sert de matrice pour les PCR. Les produits PCR sont

ensuite purifiés avec le kit de purification (Roche). Le plasmide pEL-1 ainsi que les produits PCR sont digérées par les enzymes adéquates (*Tableau 7.2*). La ligation est effectuée comme décrit précédemment par traitement à l'Antarctic Alkaline phosphatase (NEB) puis ligation avec la T4 DNA ligase (Roche). Les produits de ligation : pEL1-$BclA_{Flag}$, pEL1-$BclB_{Flag}$ et pEL1-$BclC_{Flag}$ sont introduits dans la bactérie *E. coli* DH5α compétente par choc thermique 2minutes à 42°C et sélectionnées sur Tet 30µg/ml. Après vérification de la présence des inserts par PCR sur colonies, les plasmides des clones positifs correspondant sont extraits puis séquencés.

Ces constructions, appelées pBclA, pBclB et pBclC dans la suite de l'étude (*Tableau 7.2*), permettent l'expression chez *B. cenocepacia* de protéines fusion comprenant 15 résidus supplémentaires, incluant un Flag épitope (DYKDDDDK) à leur extrémité *N*-terminale. Les plasmides d'intérêt (pBcl) sont ensuite transférés d'*E. coli* DH5α (souche donneuse) à *B. cenocepacia* K56-2 (souche receveuse) par conjugaison triparentale à l'aide d'une souche d'*E. coli* "helper" portant le plasmide self-conjugatif pRK2013 (plasmide à haute fréquence de conjugaison qui code pour les fonctions de transfert permettant la mobilisation des plasmides dérivés de pEL-1). Le plasmide pRK2013 est transféré dans les souches: *E. coli* DH5α (pEL-1), *E. coli* DH5α (pBclA), *E. coli* DH5α (pBclB), *E. coli* DH5α (pBclC), *E. coli* DH5α (pRF132) de la façon suivante. Les souches *E. coli* DH5α contenant les plasmides pEL-1, pBclA, pBclB, pBclC, pRF132 (receveuses) et *E. coli* DH5α pRK2013 (helper, donneuse) sont mises en culture en milieu LB liquide sans antibiotique à 37°C sans agitation pendant 3 heures. Un volume de 200µl d'*E. coli* DH5α pRK2013 est ajouté à 1ml de cultures d'*E. coli* DH5α receveuses contenant chaque construction. Les cultures sont incubées 3h à 37°C. Les bactéries contenant les deux plasmides (pRK2013 et l'un des plasmides pEL-1, pBclA, pBclB, pBclC ou pRF132) sont sélectionnées sur kanamycine et tétracycline 30µg/ml. Grâce au plasmide pRK2013 les plasmides pEL-1, pBclA, pBclB, pBclC et pRF132 deviennent mobilisables. Dans un deuxième temps, les souches *E. coli* DH5α contenant le plasmide pRK2013 ainsi que l'un des plasmides : pEL-1, pBclA, pBclB, pBclC ou pRF132 obtenues lors de la première conjugaison sont utilisées comme souches donneuses et *B. cenocepacia* K56-2 comme souche receveuse. Les bactéries sont mises en culture dans 1ml de milieu LB liquide sans antibiotique. Après 3 heures de culture à 37°C sans agitation, 200µl de la culture de chaque souche donneuse sont ajoutés aux cultures de *B. cenocepacia* K56-2. Après incubation une nuit à 37°C sans agitation, le produit de la conjugaison est étalé sur tétracycline 150µg/ml afin de sélectionner *B. cenocepacia* K56-2 ayant reçu les plasmides d'intérêts. Une concentration de 50µg/ml de gentamycine (auquel *B. cenocepacia* est

naturellement résistante) est également ajoutée afin d'éliminer les cellules d'*E. coli*. Après vérification de la présence de l'insert par PCR sur colonies, le plasmide d'un clone positif est extrait pour chaque construction et envoyé au séquençage.

Construction du mutant B. cenocepacia K56-2

La délétion de trois gènes *Bcam0184*, *Bcam0185* et *Bcam0186* codant respectivement pour BclC, BclB et BclA est obtenue grâce à un système utilisant l'endonucléase *ISceI* (Flannagan*, et al.* 2008). Brièvement, ce système permet la création de mutations non marquées, non polaires et nécessite deux plasmides. Tout d'abord des fragments d'ADN flanquants le gène d'intérêt à éliminer sont clonés dans le vecteur pGpI-Sce-I. Ce plasmide contient un site de restriction pour l'endonucléase I-SceI. Une fois introduit par conjugaison dans *B. cenocepacia* K56-2, ce plasmide ne peut se répliquer mais a la possibilité d'être intégré dans le chromosome par recombinaison homologue à l'aide d'une des deux régions d'homologie clonées (ADN flanquant). Les bactéries ayant intégré le plasmide sont sélectionnées sur Triméthoprime. Un second plasmide (pDAI-SceI) est introduit par conjugaison. Ce plasmide code pour l'enzyme de restriction *I-SceI*. Une fois exprimée, I-SceI réalise une coupure ADN double brin au niveau du site I-SceI préalablement intégré dans le chromosome par l'intermédiaire du plasmide pGpI-SceI. Une coupure d'ADN double brin est létale pour la bactérie si elle n'est pas réparée. Seules les bactéries qui réparent leur ADN par recombinaison homologue utilisant l'un des deux fragments d'ADN clones vont survivre, parmi elles se trouveront des bactéries ayant perdu le gène d'intérêt.

Les régions situées de part et de l'opéron comprenant *bclA,B,C* sont amplifiées par PCR a l'aide des amorces 3230/3232 et 3231/3234 et l'ADN (*Tableau 7.2*) en utilisant l'ADN de *B. cenocepacia* K56-2 comme matrice. Les produits de PCR sont purifiés, et digérés respectivement par les enzymes *Xho*I-*Eco*RI et *Xba*I-*Xho*I. Les inserts sont clonés ensemble dans le plasmide pGPI-SceI digéré avec *XbaI* et *EcoRI*. Le plasmide recombinant résultant est nommé: pDelBcl (*Tableau 7.1*). La délétion de *bclA, B, C* a été réalisée selon la méthode précédemment décrite. La perte des gènes codant pour les lectines chez *B. cenocepacia* K56-2 a été vérifiée par PCR et confirmée par Southern Blot. Dans la suite de l'étude, le mutant est appelé *B. cenocepacia* Δlec (*Tableau 7.1*).

7.1.2. Localisation cellulaire des lectines

Les bactéries *B. cenocepacia* K56-2 contenant chaque construction (pBclA, pBclB, pBclC, pEL1 et pRF132) sont cultivées à 37°C sous agitation (200rpm) en milieu liquide LB liquide contenant 100µg/ml de tetracycline. Le plasmide pRF132 (Flannagan*, et al.* 2008), codant pour la protéase périplasmique HtrA, est utilisé comme témoin de lyse cellulaire.

Isolement des fractions cellulaires : cytoplasme, membrane interne, membrane externe

Des précultures liquides de 5ml sont ensemencées pour chaque construction (pBclA, pBclB, pBclC) ainsi que pour le témoin négatif pEL-1 et incubées toute la nuit comme décrit ci-dessus. Des cultures de 25ml sont ensuite inoculées avec un volume de 20µl de préculture à 37°C toute la nuit afin d'avoir des densités optiques équivalentes pour chaque construction. Les cellules sont ensuite centrifugées 15min à 7000rpm, 4°C et les culots résultant sont resuspendus dans 5ml de 10mM Tris-HCl / pH 8 additionné de 50µl de cocktail anti-protéases 25x (Roche). Les cellules sont soniquées dans la glace à raison de pulses de 4 fois 30sec à 40% d'amplitude. Les échantillons sont ensuite centrifugés 1min à 13200rpm et 4°C. Le culot contient les membranes externes. Le surnageant est ensuite centrifugé 30min à 18000rpm et 4°C. Le surnageant contient alors les protéines périplasmiques et cytoplasmiques. Les culots correspondant sont ensuite resuspendus dans 500µl de Sarkosyl 20mM Tris-HCl / pH 8, 1.5% Sarkosyl et incubés 20min à température ambiante. Après 30 min de centrifugation à 18000rpm et 4°C, le surnageant contient les membranes internes. La concentration en protéine pour chaque fraction est déterminée par dosage de Bradford.

Isolement de la fraction périplasmique

Des précultures liquides de 5ml sont ensemencées pour chaque construction (pBclA, pBclB, pBclC) ainsi que pour le témoin négatif pEL-1 et incubées toute la nuit comme décrit ci-dessus. Des cultures liquides de 50ml sont inoculées avec les 5ml de préculture et incubées à 37°C pour 4 heures. Après centrifugation 10min à 8000rpm et 4°C, les culots sont resuspendus dans 2ml de tampon contenant 30mM Tris-Hcl / pH8, 0 1mM EDTA contenant 20% de sucrose et 3mg/ml de lysozyme. L'EDTA déstabilise la membrane externe et permet au lysozyme d'accéder au périplasme. Après incubation 3 heures à 37°C, l'apparition de sphéroplastes est contrôlée au microscope et l'incubation est prolongée d'une heure si nécessaire. Les cellules sont ensuite centrifugées 3min à 13200rpm et 4°C. Le surnageant

contient les protéines périplasmiques. La concentration protéique est déterminée par dosage de Bradford.

Extraction des protéines de surface

Des précultures liquides sont ensemencées pour chaque construction (pBclA, pBclB, pBclC) ainsi que pour pEL1 (le témoin négatif) et pRF132 (témoin de lyse cellulaire) et incubées toute la nuit comme décrit ci-dessus. Les cultures de nuit sont ensuite diluées dans 5ml de LB de manière à obtenir une Densité Optique à 600nm (DO_{600}) de 0,003 et cultivées à 37°C pour 5h00. Apres 5h00 de culture, la DO_{600} est mesurée et approximativement 3 x 10^9 bactéries (une DO_{600} de 1 correspondant à 8.10^8 bactéries.ml^{-1}) sont prélevées pour chaque culture et centrifugées 5min à 5000g et 20°C. Le culot bactérien est lavé une fois avec 1.5ml de PBS contenant un cocktail d'anti-protéases EDTA free (Roche) et à nouveau centrifugé 5min à 5000g et 6°C. Le culot est finalement resuspendu dans 120µl de SDS 0.05% ou 120µl de monosaccharide (L-Fucose ou D-mannose) à 25 ou 50mM. Après incubation 3min à température ambiante, les cellules sont centrifugées 5min à 5000g et 20°C puis 50µl de surnageant contenant les lectines de la surface cellulaire sont prélevés délicatement sans toucher le culot bactérien. La concentration en protéine est déterminée par dosage de Bradford.

Protéines sécrétées

Les précultures liquides sont ensemencées pour chaque construction (pBclA, pBclB, pBclC) ainsi que pour les deux témoins négatifs pEL-1 et pRF132, et incubées toute la nuit comme décrit ci-dessus. Chaque préculture est diluée à une DO_{600} finale de 0.03 dans 50ml de LB sans antibiotique préalablement chauffé à 37°C. Après incubation 4 heures à 37°C, sous agitation (200 rpm), les cellules sont centrifugées 15min à 10000g et 4°C. Les surnageants de chaque culture sont collectés, passés sur des filtres de 0.22µm et conservés dans la glace. Les extraits protéiques sont ensuite concentrés par l'ajout d'acide trichloroacetique (TCA) à une concentration finale de 10%. Un volume de 3ml de TCA est ajouté à des aliquots de 22ml. Après une nuit à 4°C, les échantillons sont centrifugés 30min à 10000g et 4°C. Un volume de 5 ml d'acétone refroidi dans la glace est ajouté à chaque culot. Après une nouvelle centrifugation de 30min à 10000g et 4°C l'acétone est éliminé par pipetage. Les culots sont séchés à l'aire libre puis resuspendus dans 150µl de tampon phosphate 0.1M à pH7. Après

centrifugation rapide, le surnageant contenant les protéines sécrétées est analysé par western blot.

Dosage de Bradford

Les protéines sont dosées soit selon la méthode de Bradford à l'aide du Kit Biorad Protein Assay, selon les instructions du fabricant.

Western Blot avec les anticorps anti-cepacia

Des cultures de 5ml de LB liquide d'*E. coli* contenant les constructions pBclA, pBclB, pBclC ainsi que le témoin négatif pEL-1 sont incubées toute la nuit sous agitation à 37°C, en présence de 30µg/ml de tétracycline. Les cultures sont diluées au 1/200ème dans 50ml de LB et incubées à 37°C pendant 3h. Trente ml de chaque culture sont centrifugés a 7000rpm pendant 15min. Le culot est resuspendu avec 5ml de 10mM Tris-HCl / pH7.5 contenant l'inhibiteur de protéases. Les cellules sont soniquées dans la glace à raison de pulses de 2 fois 10 sec à 40% d'amplitude. Après centrifugation 1 min à 13200 rpm et 4°C, le surnageant est transféré dans de nouveaux eppendorfs puis centrifugé à 18000 rpm pendant 30 min. Le surnageant contenant les protéines cytosoliques est déposé sur un gel SDS-PAGE 16%. Les bandes correspondant aux protéines fusion sont révélées grâce à l'anticorps monoclonal primaire de souris antiFlag : M2 reconnu par l'anticorps secondaire de chèvre Alexa Fluor 680 - immunoglobuline G anti-souris (Molecular Probes, Eugene, OR). Les bandes correspondant aux protéines immunogéniques de *B. cenocepacia* K56-2 sont détectées par les anticorps primaires de lapin antiK56-2 (pour deux variants de K-56-2: "white" ou "red"), eux même reconnus par l'anticorps secondaire de chèvre IRDye800 immunoglobuline G anti-lapin (Rockland, Pa.). Les membranes sont analysées à l'aide du système d'imagerie infrarouge Odyssey (LI-COR Biosciences, Lincoln, Neb.).

Western Blot

La présence des protéines fusion $BclA_{Flag}$, $BclB_{Flag}$ et $BclC_{Flag}$ dans les différentes fractions de *B. cenocepacia* K56-2 est analysée par Western blot. Les bandes correspondant aux protéines fusion sont révélées grâce à l'anticorps monoclonal primaire de souris anti-Flag : M2 reconnu par l'anticorps secondaire Alexa Fluor 680 - immunoglobuline G de chèvre anti-souris (Molecular Probes, Eugene, OR).

Les échantillons protéiques séparés par SDS-PAGE 16% sont transférés sur une membrane de nitrocellulose pendant 1 heure à 210mA. Après transfert, les membranes sont lavées deux fois 15 min dans 15ml de tampon TBS et bloquées une nuit à 4°C dans la solution de blocage à 10% (western blocking reagent, Roche). Elles sont ensuite incubées avec l'anticorps primaire dilué au 10000ème dans la solution de blocage à 5 %, pendant 2 heures à température ambiante. Puis, les membranes sont lavées trois fois 15 min dans 30ml de TBS, et incubées 2 heures à température ambiante avec l'anticorps secondaire dilué au 100000ème dans la solution de blocage à 5 %. Le signal est détecté grâce au système d'imagerie infrarouge (LI-COR Biosciences, Lincoln, NE).

7.1.3. Etude fonctionnelle des lectines

Courbes de croissance

Des précultures de 5 ml de LB liquide sont ensemencées pour *B. cenocepacia* K56-2 WT et ΔLec et incubées toute la nuit sous agitation à 37°C. Les cultures sont ensuite diluées de manière à avoir une DO_{600} finale de 0.05. La croissance est suivie par mesure de la DO_{600} toutes les demi-heures grâce au système d'analyse de croissance microbiologique automatiseé (Bioscreen C, MTX lab Systems, Inc., Vienna, VA).

Tests de virulence sur Caenorhabditis elegans

La virulence de la souche *B. cenocepacia* K56-2 Δlec est testée grâce au modèle d'infection utilisant *C. elegans* dans les conditions du « slow-killing assay » (Kothe, *et al.* 2003). Brièvement, 1 ml de culture de *B. cenocepacia* K56-2 Δlec et K56-2 WT sont ajustés à une DO_{600} de 1, et 50 µl de cette suspension bactérienne sont étalés sur des boites de pétri de 3,5 cm de diamètre (Falcon), contenant 4 ml de milieu solide NGMII (Nematode Growth Medium). Les boites sont incubées 24 h à 37°C afin de permettre la formation d'un tapis bactérien. Environ 25 larves de *C. elegans* DH26 (provenant du *Caenorhabditis* Genetics Centre, Universite du Minnesota, Minneapolis) sont déposées dans chaque boite de pétri puis incubées à 25°C. La souche *C. elegans* DH26 possède une mutation thermo-sensible dans le gène de spermatogénèse *fer-15*. Les nématodes peuvent donc se reproduire lorsqu'ils sont incubés à 20°C, mais sont stériles à 25°C. Apres infection et incubation à 25°C, il nous est donc possible de compter précisément le nombre de nématodes présents sur chaque boite de pétri et cela sur de longues périodes de temps sans avoir d'interférence avec leur progéniture.

Le nombre de nématodes présents dans chaque puit au moment de l'inoculation est déterminé à l'aide d'un microscope Stemi SV6 (Zeiss). Les boites de pétri sont incubées à 25°C et observées toutes les 24 h pendant 1 semaine afin de déterminer le nombre de nématodes survivants. Les expériences ont été réalisées en duplicatas en utilisant la souche non pathogénique *E. coli* OP50 comme témoin négatif de virulence et la souche *B. cenocepacia* K56-2 WT pathogénique comme témoin positif.

Cinétiques de formation du biofilm : le Biofilm Ring Test

Les expériences sont réalisées avec le Kit Biofilm Ring Test®-version Strip-Well (BioFilm Control), selon les indications du fabricant.

A partir de colonies isolées sur boîte de pétri, les bactéries *B. cenocepacia* K56-2 WT et ΔLec sont mises en culture dans 5ml de milieu BHI (Brain Heart Infusion) liquide stérile sans antibiotique. Les précultures PC1 sont incubées sur la nuit à 37°C et 150 rpm pour atteindre la phase stationnaire de croissance. Des cultures C1 sont ensuite inoculées avec les PC1, de manière à obtenir une DO_{600} de départ de 0.1 ± 0.01. Après incubation toute la nuit à 37°C et 150 rpm, du milieu BHI stérile est ajouté à la culture C1 afin d'ajuster la DO_{600} à 1 ± 0,05 et d'obtenir la culture C2. Enfin, la Solution Bactérienne Initiale (SBI) est obtenue par dilution au 1/250ème des cultures C2 avec du BHI stérile. La solution de particules magnétisables (Toner) est vortexée vigoureusement pendant 1min, puis un volume de 10µl de toner par ml de SBI est ajouté (SBI-Toner). Un témoin négatif constitué de milieu BHI stérile contenant 10µl de toner par ml de milieu (BHI-Toner) est également préparé.

Les microplaques stériles de 96 puits, sont constituées de 12 barrettes amovibles de 8 puits chacune. Les puits sont remplis par 200µl de culture et chaque mesure est réalisée en triplicata. La microplaque est incubée à 37°C en chambre humide. Pour chaque temps T d'incubation, les barrettes correspondantes sont récupérées et un volume d'environ 100 µl de liquide de contraste est ajouté à la surface de chaque puit. L'image, avant et après aimantation, est acquise grâce au logiciel "BFC Elements®".

Expériences de microcalorimétrie isotherme de titration.

Le microcalorimètre utilisé au cours de cette étude est le VP-ITC de MicroCal®. Les solutions de protéine et de fragments lipopolysaccharidiques sont préparées dans le même tampon : 0.1 M Tris-HCl / pH 7.5, 0.3 µM $CaCl_2$, de manière à avoir une concentration en ligand environ 10 fois supérieure à celle en protéine. Des aliquots de 10 µl de sucre sont

injectés toutes les 5 min dans la cellule de mesure contenant la protéine. Des séries de 30 injections sont faites à 25°C.

7.2. Résultats

7.2.1. Préparation des bactéries recombinantes

Les premiers tests effectués avec les anticorps anti-BclA, anti-BclB et anti-BclC (Sérum ont démontré un manque de spécificité. Un grand nombre de protéines de *B. cenocepacia* sont marquées lors de ces essais en addition des lectines. Une approche de biologie moléculaire a donc été utilisée. Nous avons dans un premier temps cloné les trois lectines : BclA, BclB et BclC de manière à leur ajouter une étiquette (Flag) en position N terminale, avant de les exprimer chez *B. cenocepacia* K56-2.

Deux plasmides étaient alors disponibles au laboratoire. Le premier, pBADNTF permet d'insérer une étiquette "Flag" en position *N*-terminale de la protéine clonée mais ne peut se répliquer chez *B. cenocepacia*. Le second, pDA-12, peut quant à lui être répliqué chez *B. cenocepacia* mais ne permet pas d'ajouter de Flag à la protéine clonée. Un vecteur pEL-1 a donc été crée à partir de ces deux plasmides, de manière à pouvoir exprimer les lectines possédant une étiquette en position *N*-terminale chez *B. cenocepacia* (*Figure 7.2*).

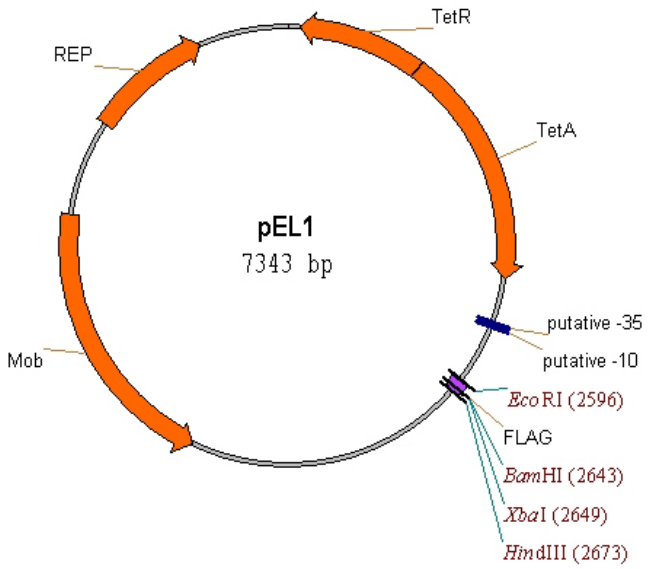


Figure 7.2. Représentation schématique du vecteur pEL-1

TetR et TetA : confèrent la résistance à la Tétracycline. REP : origine de réplication à large spectre d'hôte. *FLAG :* code pour l'épitope (DYKDDDDK) reconnu par les anticorps antiFlag : M2. *Mob* : gène de mobilisation.
pEL-1 est mobilisable (mob+) mais n'est pas auto-conjugatif (tra-). Les gènes *tra* sont apportés par pRK2013 au cours de la conjugaison).

Les constructions pEL1-*Bcl*$_{Flag}$ (pBclA, pBclB, pBclC) obtenues par clonage des lectines dans le plasmide pEL-1 ont été introduites chez *B. cenocepacia* K56-2 grâce à un système de conjugaison triparental utilisant le plasmide helper pRK2013. Un clone de *B. cenocepacia* K56-2 délété des trois gènes codant pour les lectines a également été obtenu. Les différentes cellules utilisées dans cette étude sont listées dans le *Tableau 7.3*.

Cellules	**Description**	**Etudes**		
		Expression naturelle et Immunogénicité	Localisation	Fonction
E. coli DH5α **pBclA**	Cellules contenant le plasmide pEL1-BclA$_{Flag}$	X		
E. coli DH5α **pBclB**	Cellules contenant le plasmide pEL1-BclB$_{Flag}$	X		
E. coli DH5α **pBclC**	Cellules contenant le plasmide pEl1-BclC$_{Flag}$	X		
E. coli DH5α **pEL-1**	Cellules contenant le plasmide pEL1$_{Flag}$	Témoin négatif		
B. cenocepacia K56-2 **pBclA**	Cellules contenant le plasmide pEL1-BclA$_{Flag}$		X	
B. cenocepacia K56-2 **pBclB**	Cellules contenant le plasmide pEL1-BclB$_{Flag}$		X	
B. cenocepacia K56-2 **pBclC**	Cellules contenant le plasmide pEl1-BclC$_{Flag}$		X	
B. cenocepacia K56-2 **pEL-1**	Cellules contenant le plasmide pEL1$_{Flag}$		Témoin négatif	
B. cenocepacia K56-2 **RF132**	Cellules contenant le plasmide pDA17-HtrA$_{Flag}$		Témoin de lyse cellulaire	
B. cenocepacia K56-2 **WT**	*B. cenocepacia* K56-2 sauvage			Témoin positif
B. cenocepacia K56-2 **ΔLec**	Cellules délétées des trois gènes codant pour les lectines BclA, BclB et BclC			X

Tableau 7.3. Cellules utilisées au cours de l'étude de la localisation et de la fonction des lectines chez *B. cenocepacia* K56-2

HtrA, protéase localisée dans le périplasme de *B. cenocepacia*. pDA17, plasmide permettant l'ajout d'un Flag en position *C*-terminale des protéines. WT, Wild Type

7.2.2. Expression et immnunogénicité des lectines BclA, BclB et BclC.

Des anticorps anti_cenocepacia K56-2 (anti-K56-2) ont permis d'étudier l'expression naturelle ainsi que la localisation des lectines chez la bactérie sauvage. Deux types d'anticorps dirigés contre deux variants différents de la souche K56-2 de *B. cenocepacia* ont été utilisés. Le variant "Red" produit une matrice extracellulaire de nature inconnue, tandis que le variant "white" en est dépourvu. Générés à partir de la bactérie entière, ces anticorps reconnaissent les protéines immunogéniques de la surface cellulaire de la bactérie. Des extraits cytosoliques d'*E. coli* contenant les constructions pBclA, pBclB, pBclC et pEL-1 (*Tableau 7.3*) ont été déposés sur un gel SDS-PAGE 16%, à raison de 20µg de protéines total pour chaque échantillon. Le gel a été incubé avec 2 anticorps primaire différents : anti-Flag (souris) et anti-K56-2 (lapin) puis 2 anticorps secondaires différents. L'anti-souris reconnaît l'anti-Flag ; l'autre anti-lapin reconnaît l'anti-K56-2. Les deux anticorps secondaires sont couplés à des alexa fluor différents : l'un émettant dans le rouge et l'autre dans le vert.

Les lectines recombinantes exprimées chez *E. coli.* sont reconnues par les anticorps anti-Flag et sont révélées en rouge par les anticorps secondaires murins. Le faible signal correspondant à BclA indique que chez *E.coli*, cette protéine est moins bien exprimée que les deux autres (*Figure 7.3A*). Les protéines de la surface cellulaire de *B. cenocepacia* K56-2 sont reconnues par les anticorps anti-K56-2 ("white" ou "red") et révélés en vert par les anticorps secondaires de lapin. On distingue les bandes correspondant aux trois lectines mais de nombreuses bandes non spécifiques gênent l'observation (*Figure 7.3B*).

Les bandes jaunes indiquent un double marquage anti-Flag et anti-K56-2 (*Figure 7.3C*). Les trois lectines recombinantes produites chez *E.coli* sont donc reconnues par les anticorps anti-K56-2. Les trois lectines sont reconnues par les anticorps anti-K56-2. Ces anticorps étant générés chez la bactérie *B. cenocepacia* K56-2 sauvage, la reconnaissance des trois protéines par ceux-ci indique que les lectines sont naturellement exprimées chez la bactérie. De plus, les anticorps anti-K56-2 étant produits à partir des bactéries intactes, la reconnaissance des lectines recombinantes suggère que les protéines sont localisées à la surface des bactéries. Enfin la présence d'anticorps dirigés contre les lectines dans les anticorps anti-K56-2 montre que ces protéines sont des éléments immunogéniques chez la bactérie. Il ne sembles pas y avoir de différence d'expression des trois lectines chez les deux variants white" et "red.

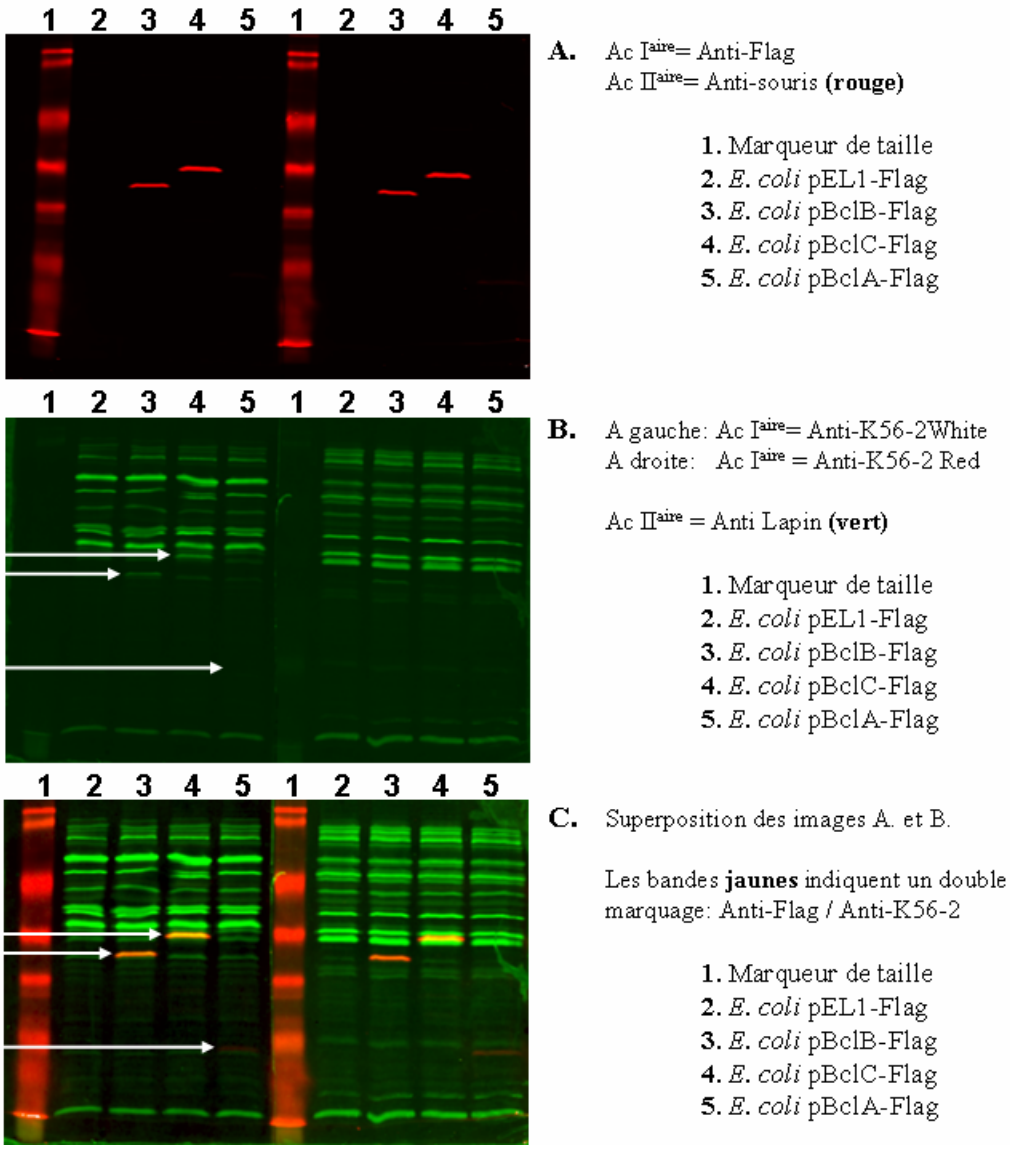


Figure 7.3. Reconnaissance des lectines recombinantes $BclA_{Flag}$, $BclB_{Flag}$ et $BclC_{Flag}$ produite chez *E. coli* par des anticorps anti K-56-2

A. Les protéines recombinantes, $BclA_{Flag}$, $BclB_{Flag}$ et $BclC_{Flag}$ sont reconnues par des anticorps anti-Flag. Les anticorps secondaires murins émettent dans le rouge. **B.** Les protéines de la surface cellulaire de *B. cenocepacia* K56-2 sont reconnues par des anticorps anti-K56-2 "white" à gauche et "red" à droite. Les anticorps secondaires produits chez le lapin émettent dans le vert. **C.** Les bandes jaunes indiquent un double marquage anti-Flag et anti K56-2.

7.2.3. Localisation des lectines solubles de B. cenocepacia.

La localisation des protéines fusion $BclA_{Flag}$, $BclB_{Flag}$ et $BclC_{Flag}$ a été étudiée par fractionnement cellulaire. Après détermination de leur concentration en protéines par dosage de Bradford, les échantillons protéiques correspondant aux différentes fractions : cytoplasme, membrane interne, périplasme et membrane externe, sont déposées sur gels SDS-PAGE 16%. La présence des lectines est ensuite détectée par western blot grâce aux anticorps Anti-Flag M2.

Les résultats montrent que les lectines sont majoritairement présentes dans le cytoplasme ainsi que dans la membrane externe de *B. cenocepacia* K56-2 (*Figures 7.4A* et *7.4D*). On les retrouve également, mais en quantité dix fois moins importante, dans la membrane interne et dans le périplasme (*Figure 7.4B et 7.4C*).

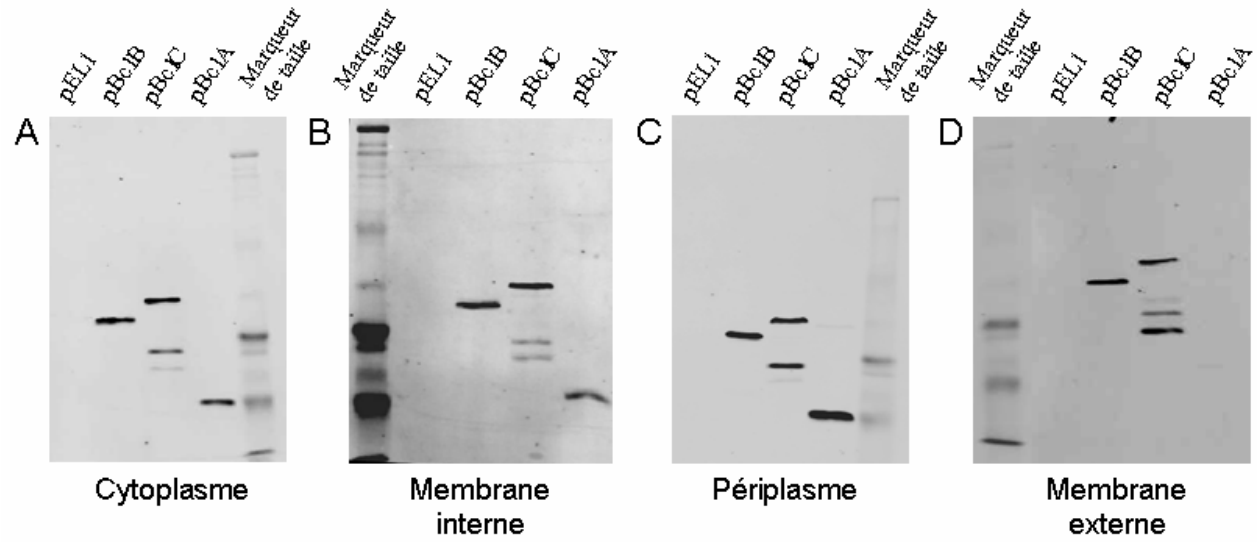


Figure 7.4. Localisation cellulaire des protéines $BclA_{Flag}$, $BclB_{Flag}$ et $BclC_{Flag}$ chez *B. cenocepacia* K56-2.

Les gels sont chargés par des quantités variables de protéines en fonction de la fraction analysée. **A.** cytoplasme : 1.3µg. **B.** membrane interne : 14µg. **C.** périplasme : 12µg. **D.** membrane externe : 2.7µg. Les échantillons sont séparés par SDS-PAGE 16% et transférés sur une membrane de nitrocellulose. Les protéines possédant une étiquette (Flag) sont détectées grâce aux anticorps anti-Flag. Les cellules pEL1, contenant le plasmide vide servent de témoin négatif à l'expérience.

Deux produits de dégradations apparaissent pour la protéine BclC avec des tailles d'environ 20 kDa. Ces bandes ne correspondent pas à la taille des deux domaines BclCNter et

BclCLec séparés qui sont de 15.7 et 12.4 kDa respectivement. La bande correspondant à BclA dans la membrane externe est très faible en comparaison avec les deux autres protéines.

Différentes approches ont été utilisées pour tester la présence des lectines à la surface des cellules, mais aussi dans le milieu extracellulaire. Les témoins négatifs : cellules pEL1 et pRF132 permettent de tester la validité de nos protocole expérimentaux. En effet, la protéase HtrA impliquée dans la réponse au stress chez *B. cenocepacia* est une protéine périplasmique.

Dans un premier temps, les protéines ont été décrochées de la surface cellulaire grâce à une concentration de 0.05% de SDS a été testée (Huber et al. 2004). L'analyse des protéines extraites par Western Blot révèle la présence de la protéine périplasmique HtrA dans les protéines de surface, indiquant que les cellules ont été endommagées au cours de l'extraction par le SDS et que notre protocole n'est pas valable pour cette étude (*Figure 7.4A*).

Une étude précédente avait démontré que la lectine PA-IIL pouvait être extraite à partir de la membrane externe de *P. aeruginosa* par un traitement avec le p-nitrophenyl α-L-fucose (Tielker, *et al.* 2005). De plus, la lectine mutée au niveau de son site de liaison (D104A) n'était plus détectée dans la membrane externe de la bactérie. Ces résultats suggèrent que la protéine PA-IIL est attachée à un ligand saccharidique de la surface cellulaire. Une deuxième approche reposant sur un système de compétition avec les sucres de la surface cellulaire a donc été mise au point (*Figure 7.5*).

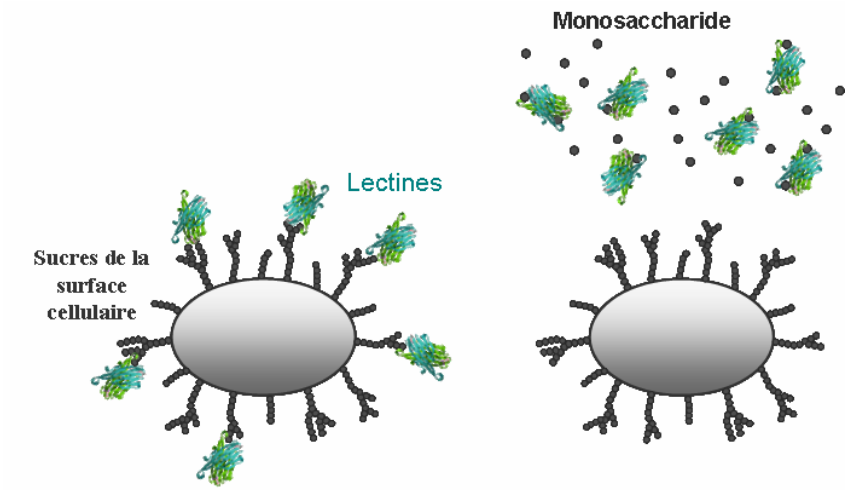


Figure 7.5. Extraction des lectines solubles de la surface cellulaire de *B. cenocepacia* par un système de compétition avec un monosaccharide (D-mannose ou L-fucose).

Les expériences sont réalisées pour deux monosaccharides : D-Mannose et de L-Fucose pour deux concentrations différentes : 25 et 50 mM. Les expériences ont été reproduites plusieurs fois sur des extractions différentes. Les résultats sont identiques pour le L-fucose et le D-mannose, et ce, quelque soit la concentration en sucre. Seul le western Blot d'analyse de

la fraction protéique extraite par 25mM de D-mannose est présenté dans ce manuscrit (*Figure 7.6B*). Les trois protéines sont détectées dans les protéines extraites de la surface cellulaire. Là encore, la quantité de BclA est moins importante que pour les deux autres lectines.

Les protéines du milieu extracellulaire ont-elles aussi été analysées par western blot. Les résultats présentés dans la *Figure 7.6C* montrent que les trois lectines sont présentes dans cette fraction. L'absence de bande dans le puit correspondant aux protéines extracellulaires de pRF132 indique qu'il n'y a pas eu de lyse cellulaire au cours de l'expérience. Les lectines sont donc sécrétées dans le milieu extracellulaire.

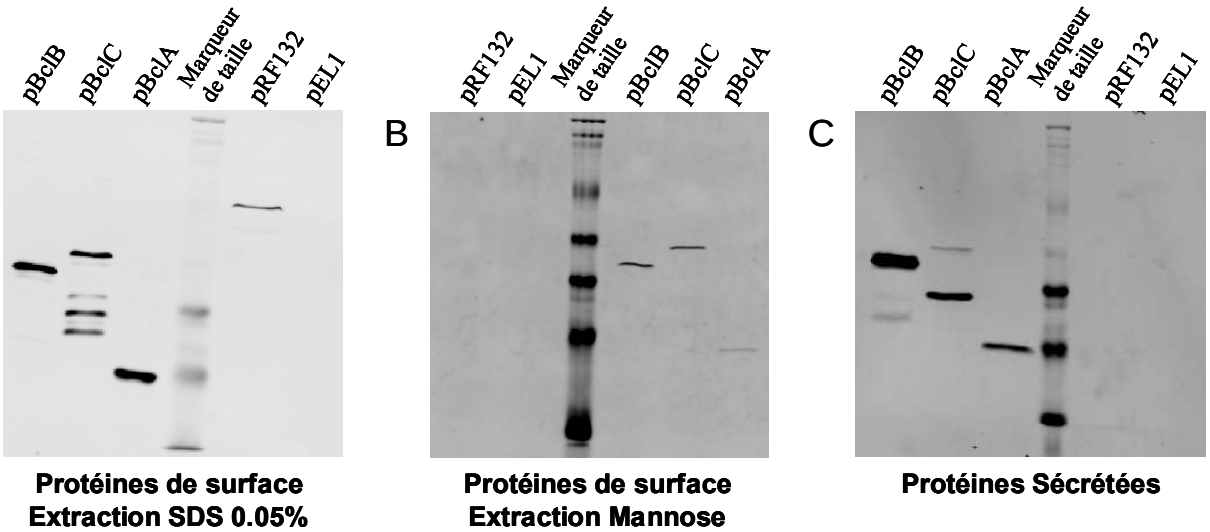


Figure 7.6. Identification des protéines $BclA_{Flag}$, $BclB_{Flag}$ et $BclC_{Flag}$ à la surface des cellules et dans le milieu extracellulaire de *B. cenocepacia* K56-2.

A. Protéines de surface extraites par 0.05% de SDS. **B.** protéines de surface extraites par 25mM de D-Mannose. **C.** Protéines sécrétées.
Les échantillons sont séparés par SDS-PAGE 16% et transférés sur une membrane de nitrocellulose. Les protéines possédant une étiquette (Flag) sont détectées grâce aux anticorps anti-Flag. Les cellules pEL1, contenant le plasmide vide et pRF132, contenant la protéine périplasmique HtrA, servent de témoin négatif à l'expérience.

7.2.4. Etude de l'affinité de BclA et BclB pour des fragments lipopolysaccharidiques.

Les résultats précédents nous ont montré que les lectines solubles de *B. cenocepacia* sont attachées à la surface des cellules et peuvent être décrochée par un monosaccharide (L-fucose ou D-mannose). Nous pouvons donc proposer que les lectines soient attachées aux LPS de *B. cenocepacia* particulièrement abondants chez cette bactérie. Ces LPS contiennent

une grande variété de monosaccharides mais ceux qui paraissent le plus intéressants sont les D-heptose car ils présentent une configuration *D-manno* avec deux groupements hydroxyles équatoriaux et un axial, ce qui est la stéréochimie requise pour se lier aux lectines à deux calciums.

Une série de monosaccharides et de disaccharides synthétisés par le Professeur Paul Kosma (Institute of Chemistry, Vienne, Autriche) ont été testés par microcalorimétrie isotherme de titration avec la lectine BclA. Les quatre heptoses suivants ont pu être étudiés : le 2-deoxy-heptose, l'allyl L-glycero-α-D-manno-heptopyranoside, le D-glycéro-D-manno-heptose et enfin le L-glycéro-D-manno-heptose (*Figure 7.7*). Les interactions ont été mesurées à 25°C, pour des concentrations en sucre dix fois supérieures à celles en protéines.

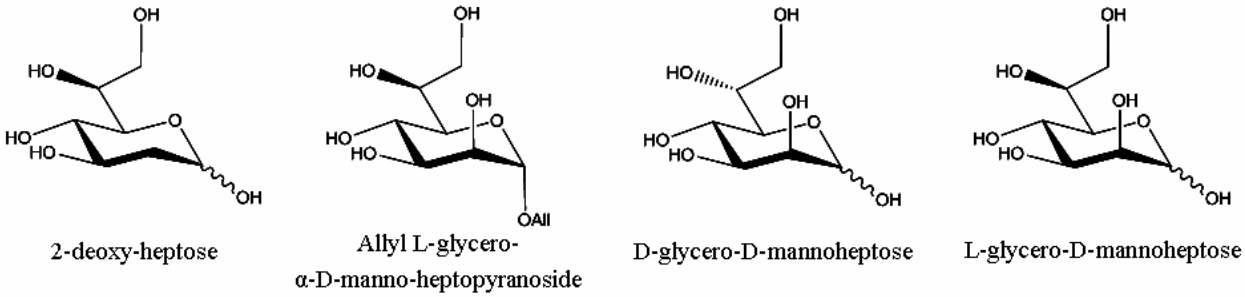


Figure 7.7. Structure des heptoses testés en microcalorimétrie isotherme de titration.

La lectine BclA interagit avec le 2-deoxy-heptose, mais avec une affinité très faible de 0.56 mM. C'est le cas également du L- glycéro-D-manno-heptose reconnu par BclA et BclB avec une affinité de 0.12 mM et 0.79 mM respectivement. En revanche, aucune interaction n'a été détectée pour le D- glycéro-D-manno-heptose. La plus forte affinité a été déterminée pour l'interaction de BclA avec l'allyl L-glycero-α-D-manno-heptopyranoside (Kd = 74 μM) (*Tableau 7.4*)

K_a (10^4 M^{-1})	K_d (10^{-5} M)	N	-ΔG (kJ/mol)	-ΔH (kJ/mol)	-TΔS (kJ/mol)
1.35 (± 0.08)	7.4 (± 0.5)	0.94 (± 0.05)	23.6 (± 0.1)	12.8 (± 1)	10.8 (± 1)

Tableau 7.4. Constantes thermodynamiques de l'interaction entre BclA et l'Allyl L-glycero-α-D-manno-heptopyranoside mesurées par microcalorimétrie isotherme de titration.

Les déviations standards sont calculées à partir de trois expériences indépendantes

Dans un deuxième temps, l'interaction des lectines BclA et BclB avec des fragments lipopolysaccharidiques purifiés à partir d'extraits bactériens (Collaboration Antonio Molinaro de l'Université de Naples, Italie) a été étudiée. Deux fragments ont été testés: un fragment oligosaccharidique constitué de deux heptoses, un Kdo, un Ko ainsi qu'un ara4N et un LPS O-acétylé contenant un lipideA O-acétylé contenant les mêmes sucres que le fragment précédent. Aucune interaction n'a été détectée, que ce soit avec BclA ou BclB. Les interactions testées dans cette d'étude sont résumées dans le tableau (*Tableau 7.5*).

Ligand testé	BclA	BclB
Monosaccharides		
2-deoxy-Heptose	-	
D-Glycero-D-mannoheptose	-	
L-Glycero-D-mannoheptose	+	+
Allyl L-glycero-α-D-manno-heptopyranoside	+	
Fragments Lipopolysaccharidiques		
Partie oligosaccharidique	-	-
LPS o-acetylé	-	-

Tableau 7.5. Résumé des interactions mesurées par microcalorimétrie isotherme de titration.

– Aucune interaction mesurée
+ Mesure d'une interaction

7.2.5. Etude du rôle des lectines chez B. cenocepacia K56-2

Afin d'étudier le rôle des lectines solubles chez la bactérie, un mutant de *B. cenocepacia* K56-2 délété des gènes codant pour les trois lectines a été construit. La croissance de ce mutant, sa virulence chez *C. elegans* ainsi que sa capacité à former des biofilms ont été étudiées.

Les courbes de croissance de *B. cenocepacia* K56-2 WT et ΔLec ont été mesurées sur une période de 24 heures. Les résultats indiquent que la croissance de la bactérie n'est pas affectée par la délétion des trois gènes codant pour les lectines (*Figure 7.8*).

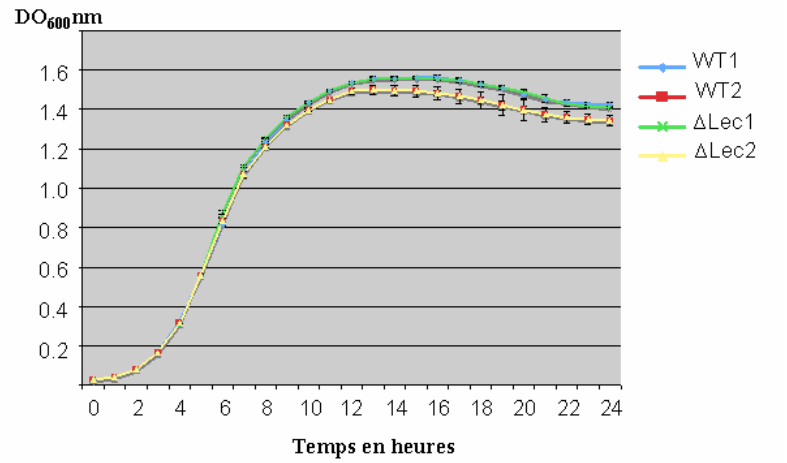


Figure 7.8. Croissance de *B. cenocepacia* K56-2 sauvage (WT) et délété des lectines (ΔLec)

Les cellules sont cultivées en milieu LB liquide sans antibiotique à 37°C sur une période de 24 heures. Les courbes sont mesurées pour deux cultures indépendantes de *B. cenocepacia* K56-2 sauvage (WT1 et WT2) ainsi que pour deux cultures du mutant (Δlec1 et Δlec2).

Des tests de virulence chez *C. elegans* ont été effectués pour *B. cenocepacia* WT et ΔLec. Les mesures sont comparées à celles obtenues pour la bactérie non pathogénique *E. coli* (OP50). La souche *B. cenocepacia* K56-2 WT est pathogénique pour *C. elegans*. En quelques jours les nématodes ingérant *B. cenocepacia* K56-2 deviennent malades. Ils présentent une motilité réduite et leur intestin est abondamment colonisé par *B. cenocepacia*. En 4 jours 100% des nématodes meurent du fait de l'infection. Il ne semble pas y avoir de différence dans la virulence de *B. cenocepacia* K56-2 sauvage et muté dans ce modèle (*Figure 7.9*).

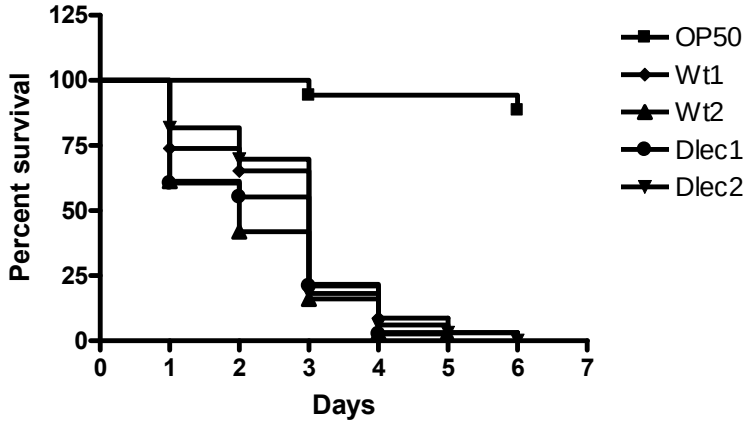


Figure 7.9. Test de la virulence de la bactérie *B. cenocepacia* K56-2 délétée chez *C. elegans.*

Le test est réalisé pour deux cultures indépendantes de *B. cenocepacia* K56-2 sauvage (WT1 et WT2) ainsi que pour deux cultures du mutant (Δlec1 et Δlec2), ainsi que pour la bactérie non pathogénique *E. coli* (OP50). Chaque test est réalisé sur une période d'une semaine.

Enfin, la cinétique de formation du biofilm par le mutant ΔLec a été étudiée et comparée à celle du WT grâce à une nouvelle technologie appelée le "Biofilm Ring Test". Dans cette technique, les bactéries sont mélangées à des particules magnétisable et mises en culture dans des plaques multipuits. Les puits sont soumis à une aimantation après différents temps de culture. L'image du fonds des puits est analysée et un score appelé « Indice BioFilm » (BFI) est attribué à chaque puit. Lorsque le biofilm se forme, les particules magnétisables ne peuvent plus se concentrer dans champ magnétique et on observe une absence de formation des spots (*Figure 7.10A*). L'expérience a été renouvelée plusieurs fois et la *Figure 7.10* représente les résultats classiquement obtenus. Que ce soit pour la souche sauvage ou mutée, le biofilm se forme après quatre heures de culture en moyenne (2<BFI<7) pour être total après 5 heures de culture (2>BFI). Il ne semble pas y avoir de différence dans la formation de biofilm chez B. cenocepacia K56-2 sauvage et muté (*Figure 7.10B*).

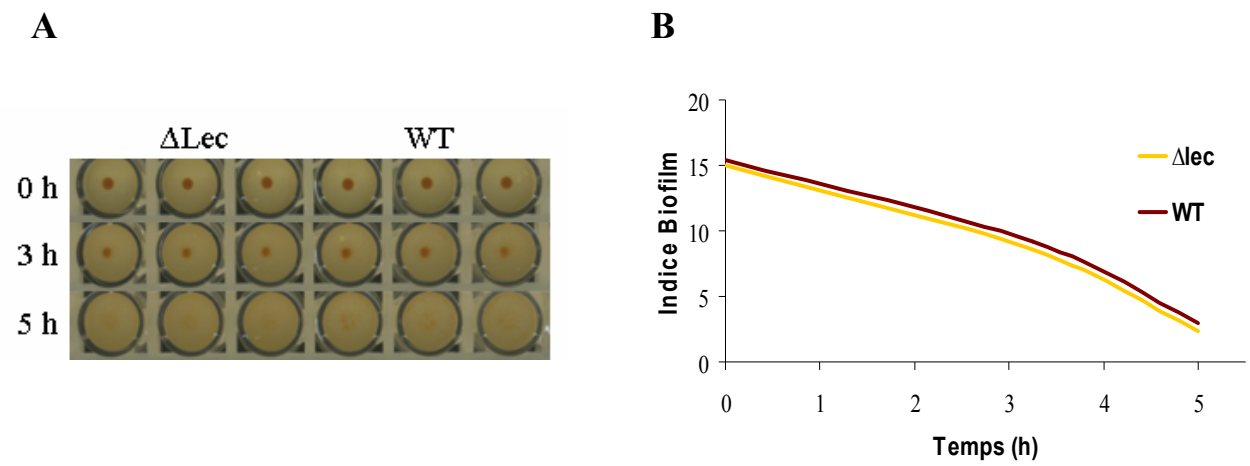


Figure 7.10. Comparaison de la formation du biofilm chez *B. cenocepacia* K56-2 WT et ΔLec grâce au biofilm Ring Test.

A. Images brutes obtenues après aimantation des puits de *B. cenocepacia* K56-2 WT et ΔLec pour 0, 3 et 5heures de culture. Les mesures sont réalisées en triplicata sur trois cultures indépendantes
B. Cinétique de formation du biofilm de *B. cenocepacia* K56-2 WT et ΔLec.

7.3. Conclusions et perspectives

Les expériences de western blot effectuées avec les anticorps anti-K56-2 ont montré que les protéines BclA, BclB et BclC sont présentes à la surface des bactéries. L'extraction par des monosaccharides (L-Fuc et D-Man) des protéines de la surface cellulaire a confirmé cette localisation extracellulaire des lectines. Nous avons émis l'hypothèse que les lectines sont attachés aux LPS de la bactérie. Les premiers essais de mesure de l'affinité de ces protéines pour des fragments lipopolysaccharidiques n'ont pas permis de valider cette hypothèse. Toutefois, ces premières mesures sont sont prometteuses. Les essais seront poursuivis avec des fragments lipopolysaccharidiques purs.

La reconnaissance des lectines par les anticorps anti-K56-2 a démontré les propriétés immunogéniques de ces protéines. La virulence ainsi que la croissance des bactéries ne semblent pas affectées par la délétion des trois gènes codant pour les lectines. Les expériences de « Biofilm Ring Test » ont montré que la quantité de biofilm formé n'est pas modifiée par la délétion des lectines.

.

Chapitre 8

Conclusion et perspectives

Dans le cadre de ces travaux de thèse je me suis intéressée à la caractérisation biochimique et structurale de deux lectines solubles de *B. cenocepacia* : BclA et BclB possédant des homologies de séquences avec la lectine PA-IIL de *P. aeruginosa*. La localisation ainsi que la fonction de ces protéines chez *B. cenocepacia* ont également pu être étudiées.

8.1. Etude de la lectine BclA

BclA, une lectine à mannose

La lectine la plus courte, BclA est la plus proche de PA-IIL en terme d'homologie de séquence. La résolution de sa structure cristallographique en complexe avec l'α-methyl-mannoside (αMeMan) a montré que contrairement à PA-IIL qui est organisée en tétramère, BclA est un homodimère. Comme pour PA-IIL, chaque monomère de BclA est capable de lier un sucre par l'intermédiaire de deux ions calcium. Les expériences de printed array ont permis de mettre en évidence la spécificité stricte de cette lectine pour les structures glycanniques de type oligo-mannoside. (Article n°1).

Les lectines sont aujourd'hui utilisées en tant qu'outil dans le secteur biomédical, mais aussi dans la recherche. Certaines d'entre elles sont utilisées comme marqueurs histo-chimiques pour identifier les modifications de la glycolysation liées à certaines maladies comme le cancer. De par leur spécificité, les lectines immobilisées sur colonne peuvent aussi être utilisées pour l'identification, la purification et la caractérisation des glycoconjugués (Hirabayashi 2004). Avec une affinité de 2.75 10^{-6}M, BclA est aujourd'hui la lectine à mannose possédant la plus forte affinité pour ce monosaccharides. Associée à cette affinité micro-molaire, la spécificité stricte de cette nouvelle lectine pour le mannose offre des perspectives d'utilisation en biotechnologie.

BclA a déjà été utilisée comme outils pour tester la disponibilité de résidus mannose à la surface de copolymères mannosylés (Jutta Rieger, 2007) (Article n°3 en Annexe). Le même type d'étude est actuellement en cours dans notre laboratoire, avec des nanoparticules d'or sur lesquels des résidus mannose ont été greffées (Travaux de thèse de Michael Reynolds, 2006-2009.

Effet coopératif de l'interaction entre BclA et ses ligands

L'allure des courbes mesurées en microcalorimétrie pour l'interaction entre BclA et ses ligands correspond à une interaction à deux sites coopératifs. La lectine PA-IIL ainsi que les autres protéines de la même famille ne présentent pas un tel effet.

Dans la structure de BclA complexée à l'αMeMan, l'histidine 112 est impliquée dans un contact hydrophobe avec le mannose. La densité électronique indique que cette histidine qui est absente chez PA-IIL (remplacée par une Thréonine) se trouve dans une double conformation (*Figure 8.1*). Nous avons donc suggéré que cet acide aminé pourrait jouer un rôle dans l'effet coopératif. Mais les courbes de microcalorimétrie mesurées avec la lectine BclA, mutée (H112A) sont identiques à celles obtenues pour la lectine sauvage, invalidant ainsi cette hypothèse (Données non publiées - Expériences réalisées par Michaela Wimmerova, Brno).

BclA possède également une séquence de huit acides aminés (95 à 102) qui sont absents chez PA-IIL (*Figure 8.1*). Ces acides aminés étant en contact avec ceux du monomère adjacent, nous avons proposé que la fixation du sucre dans le premier site induise un changement de conformation qui serait transmis à l'autre monomère par l'intermédiaire de ces acides aminés supplémentaires. La deuxième lectine, BclB ne possède pas ces acides aminés supplémentaires. Les courbes de microcalorimétrie obtenues pour cette protéine peuvent être ajustées avec un modèle à un site de liaison. En revanche, les courbes obtenues pour le domaine lectine BclBLec seul présentent les caractéristiques d'une interaction coopérative. La séquence de huit acides aminés, absente de BclB, ne peut donc pas expliquer l'effet coopératif pour BclA et BclBLec.

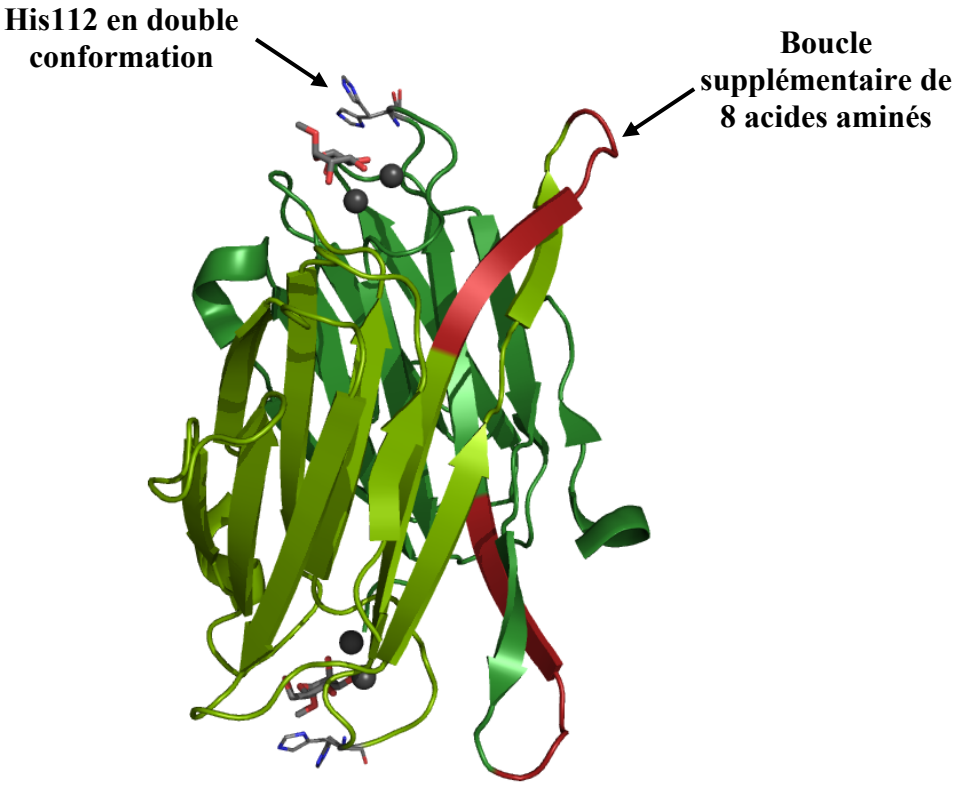


Figure 8.1. Homodimère de BclA en complexe avec l'αMeMan

Chaque monomère, reconnaît un sucre par l'intermédiaire de deux ions calcium (sphères grises). Les huit acides aminés absents chez PA-IIL sont représentés en rouge

8.2. Etude de la lectine BclB

La protéine BclB, organisée en homotétramère, reconnaît le Fucose avec une affinité de 2 82.10^{-6}M. Sa spécificité pour les oligosaccharide fucosylés: H-type 1, H-type 3, H-type 4 et Lewis a été déterminée par « printed array ».

Les difficultés rencontrées dans la cristallogenèse de la protéine entière étant probablement dues à la flexibilité entre les deux domaines, les efforts de cristallisation seront poursuivis avec les deux domaines BclBLec et BclBNter.

La fonction du domaine lectine qui pourrait être un deuxième domaine lectine, une enzyme, ou tout autre chose n'a pu être déterminée au cours de cette étude. Ces résultats sont très récents et les expériences seront poursuivies pour déterminer le rôle de ce domaine. Les informations structurales apportées par la résolution de la structure pourraient nous aider à répondre à cette question.

8.3. Rôle des lectines de surface chez B. cenocepacia

Les résultats obtenus au cours de ces travaux ont montré que les lectines solubles, BclA, BclB et BclC de *B. cenocepacia* sont présentes à la surface des cellules mais sont aussi sécrétées dans le milieu extracellulaire.

L'extraction des protéines de la surface cellulaire par des monosaccharides a montré que les protéines sont attachées à des ligands saccharidiques. Des premières mesures de l'interaction entre les lectines et des fragments de LPS (heptoses) ont été effectuées. Ces résultats préliminaires étant prometteurs l'étude d'autres fragments sera poursuivie. La répétition des expériences de décrochage des lectines de la surface cellulaire sur une des souches de *B. cenocepacia* exprimant des molécules de LPS tronquées à différents niveaux (sans o-antigène, sans heptose…) disponible au "Department of Microbiology and Immunology" permettrait de tester cette hypothèse. De plus, des clones de chaque lectine dans lesquels les acides aminés impliqués dans la liaison au sucre sont mutés ont récemment été construits par le Dr Daniel Aubert qui continue aujourd'hui ce travail. L'analyse des protéines secrétées pour ces mutants nous permettrons de savoir si la fonction de lectine est nécessaire à la sécrétion des protéines à l'extérieur de la cellule.

Il a été suggéré que la lectine PA-IIL joue un rôle dans la formation du biofilm (Tielker S., 2005). Dans cette étude aucun effet de la délétion des trois lectines n'a été observé sur la quantité de biofilm formée. Des expériences de microscopie confocale nous permettraient de voir si la structure elle-même du biofilm est altérée par la délétion des lectines.

La capacité de BclA à former un réseau, observée dans la structure cristallographique et confirmée par AFM illustre bien les propriétés d'hémagglutination des lectines (Article n°2). Cette protéine pourrait donc par exemple être impliquée dans la stabilisation du biofilm.

De part sa spécificité, La lectine BclB pourrait quand à elle servir de point d'ancrage à *B. cenocepacia* sur les voies aériennes des patients atteints de mucoviscidose possédant des glycoconjugués particulièrement fucosylées et initier l'infection. Elle pourrait également être impliquée dans la formation des biofilms.

Finalement, le rôle exact de ces protéines chez la bactérie n'a pas encore pu être déterminé. Il est possible que les lectines jouent un rôle dans certaines phases du cycle de vie des bactéries, ou dans certaines conditions d'interaction avec le milieu extérieur et des expériences supplémentaires seront nécessaires pour analyser ces différentes possibilités.

Bibliographie

A

Adam, J., Pokorna, M., Sabin, C., Mitchell, E.P., Imberty, A. and Wimmerova, M. (2007) Engineering of PA-IIL lectin from *Pseudomonas aeruginosa* - Unravelling the role of the specificity loop for sugar preference. *BMC Struct. Biol.*, **7** 36.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389-3402.

Amaral, M.D. and Kunzelmann, K. (2007) Molecular targeting of CFTR as a therapeutic approach to cystic fibrosis. *Trends Pharmacol. Sci.*, **28**, 334-341.

B

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.

Baker, H.M., Basu, I., Chung, M.C., Caradoc-Davies, T., Fraser, J.D. and Baker, E.N. (2007) Crystal structures of the staphylococcal toxin SSL5 in complex with sialyl Lewis X reveal a conserved binding site that shares common features with viral and bacterial sialic acid binding protein. *J. Mol. Biol.*, **374**, 1298-1308.

Baldwin, A., Sokol, P.A., Parkhill, J. and Mahenthiralingam, E. (2004) The *Burkholderia cepacia* epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*. *Infect. Immun.*, **72**, 1537-1547.

Banerjee, D. and Stableforth, D. (2000) The treatment of respiratory pseudomonas infection in cystic fibrosis: what drug and which way? *Drugs*, **60**, 1053-1064.

Barasch, J., Kiss, B., Prince, A., Saiman, L., Gruenert, D. and al-Awqati, Q. (1991) Defective acidification of intracellular organelles in cystic fibrosis. *Nature*, **352**, 70-73.

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 (PA-IL). *J. Mol. Biol.*, **383**, 837-853.

Blixt, O., Head, S., et al. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A*, **101**, 17033-17038.

Bochner, B.S., Alvarez, R.A., Mehta, P., Bovin, N.V., Blixt, O., White, J.R. and Schnaar, R.L. (2005) Glycan array screening reveals a candidate ligand for Siglec-8. *J. Biol. Chem.*, **280**, 4307-4312.

Boucher, R.C. (2002) An overview of the pathogenesis of cystic fibrosis lung disease. *Adv. Drug Deliv. Rev.*, **54**, 1359-1371.

Brisse, S., Cordevant, C., Vandamme, P., Bidet, P., Loukil, C., Chabanon, G., Lange, M. and Bingen, E. (2004) Species distribution and ribotype diversity of *Burkholderia cepacia* complex isolates from French patients with cystic fibrosis. *J. Clin. Microbiol.*, **42**, 4824-4827.

Brito, C.F., Carvalho, C.B., Santos, F., Gazzinelli, R.T., Oliveira, S.C., Azevedo, V. and Teixeira, S.M. (2004) *Chromobacterium violaceum* genome: molecular mechanisms associated with pathogenicity. *Genet. Mol. Res.*, **3**, 148-161.

Brockhausen, I., Vavasseur, F. and Yang, X. (2001) Biosynthesis of mucin type O-glycans: lack of correlation between glycosyltransferase and sulfotransferase activities and CFTR expression. *Glycoconj. J.*, **18**, 685-697.

Bukholder, W.H. (1950) Sour skin, a bacterial rot of onion bulbs. *Phytopathology*, **40**, 115-117.

Busch, R. (1990) On the history of cystic fibrosis. *Acta Univ Carol [Med] (Praha)*, **36**, 13-15.

Buts, L., Bouckaert, J., De Gents, E., Loris, R., Oscarson, S., Lahmann, M., Messens, J., Brosens, E., Wyns, L. and De Greve, H. (2003) The fimbrial adhesin F17-G of enterotoxigenic *Escherichia coli* has an immunoglobulin-like lectin domain that binds N-acetylglucosamine. *Mol. Microb.*, **49**, 705-715.

Bylund, J., Burgess, L.A., Cescutti, P., Ernst, R.K. and Speert, D.P. (2006) Exopolysaccharides fro*m Burkholderia cenocepacia* inhibit neutrophil chemotaxis and scavenge reactive oxygen species. *J. Biol. Chem.*, **281**, 2526-2532.

Byres, E., Paton, A.W., Paton, J.C., Lofling, J.C., Smith, D.F., Wilce, M.C., Talbot, U.M., Chong, D.C., Yu, H., Huang, S., Chen, X., Varki, N.M., Varki, A., Rossjohn, J. and Beddoe, T. (2008) Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. *Nature*, **456**, 648-652.

C

Chen, J.S., Witzmann, K.A., Spilker, T., Fink, R.J. and LiPuma, J.J. (2001) Endemicity and inter-city spread of *Burkholderia cepacia* genomovar III in cystic fibrosis. *J. Pediatr.*, **139**, 643-649.

Chua, K.L., Chan, Y.Y. and Gan, Y.H. (2003) Flagella are virulence determinants of *Burkholderia pseudomallei*. *Infect. Immun.*, **71**, 1622-1629.

Chung, M.C., Wines, B.D., Baker, H., Langley, R.J., Baker, E.N. and Fraser, J.D. (2007) The crystal structure of staphylococcal superantigen-like protein 11 in complex with sialyl

Lewis X reveals the mechanism for cell binding and immune inhibition. *Mol. Microbiol.*, **66**, 1342-1355.

Cioci, G., Mitchell, E.P., Gautier, C., Wimmerova, 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 Lett.*, **555**, 297-301.

Coenye, T., Spilker, T., Van Schoor, A., LiPuma, J.J. and Vandamme, P. (2004) Recovery of *Burkholderia cenocepacia* strain PHDC from cystic fibrosis patients in Europe. *Thorax*, **59**, 952-954.

Comolli, J.C., Waite, L.L., Mostov, K.E. and Engel, J.N. (1999) Pili binding to asialo-GM1 on epithelial cells can mediate cytotoxicity or bacterial internalization by *Pseudomonas aeruginosa*. *Infect. Immun*, **67**, 3207-3214.

Compant, S., Nowak, J., Coenye, T., Clement, C. and Ait Barka, E. (2008) Diversity and occurrence of *Burkholderia spp*. in the natural environment. *FEMS Microbiol. Rev.*, **32**, 607-626.

Conway, B.A. and Greenberg, E.P. (2002) Quorum-sensing signals and quorum-sensing genes in *Burkholderia vietnamiensis*. *J. Bacteriol.*, **184**, 1187-1191.

Cox, A.D. and Wilkinson, S.G. (1991) Ionizing groups in lipopolysaccharides of *Pseudomonas cepacia* in relation to antibiotic resistance. *Mol. Microbiol.*, **5**, 641-646.

D

Dalmastri, C., Chiarini, L., Cantale, C., Bevivino, A. and Tabacchioni, S. (1999) Soil type and maize cultivar affect the genetic diversity of maize root-associated *Burkholderia cepacia* populations. *Microb. Ecol.*, **38**, 273-284.

Dam, T.K. and Brewer, C.F. (2002) Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.*, **102**, 387-429.

Davidson, D.J. and Dorin, J.R. (2001) The CF mouse: an important tool for studying cystic fibrosis. *Expert Rev. Mol. Med.*, **2001**, 1-27.

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.

de Bentzmann, S., Roger, P., Dupuit, F., Bajolet-Laudinat, O., Fuchey, C., Plotkowski, M.C. and Puchelle, E. (1996) Asialo GM1 is a receptor for Pseudomonas aeruginosa adherence to regenerating respiratory epithelial cells. *Infect. Immun.*, **64**, 1582-1588.

De Greve, H., Wyns, L. and Bouckaert, J. (2007) Combining sites of bacterial fimbriae. *Curr Opin Struct Biol*, **17**, 506-512.

De Soyza, A., Silipo, A., Lanzetta, R., Govan, J.R. and Molinaro, A. (2008) Chemical and biological features of *Burkholderia cepacia* complex lipopolysaccharides. *Innate Immun.*, **14**, 127-144.

Delmotte, P., Degroote, S., Lafitte, J.J., Lamblin, G., Perini, J.M. and Roussel, P. (2002) Tumor necrosis factor alpha increases the expression of glycosyltransferases and sulfotransferases responsible for the biosynthesis of sialylated and/or sulfated Lewis x epitopes in the human bronchial mucosa. *J. Biol. Chem.*, **277**, 424-431.

Desai, M., Buhler, T., Weller, P.H. and Brown, M.R. (1998) Increasing resistance of planktonic and biofilm cultures of *Burkholderia cepacia* to ciprofloxacin and ceftazidime during exponential growth. *J. Antimicrob. Chemother.*, **42**, 153-160.

Di Sant'Agnese, P., Darling, R.C., Perara, G.A. and Shea, E. (1953) Abnormal electrolyte composition of sweat in cystic fibrosis of the pancreas. *AMA Am. J. Dis. Child*, **86**, 618-619; discussion, 619.

Diggle, S.P., Stacey, R.E., Dodd, C., Camara, 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 ist human kideny receptor. *Cell*, **105**, 733-743.

Dosanjh, A., Lencer, W., Brown, D., Ausiello, D.A. and Stow, J.L. (1994) Heterologous expression of delta F508 CFTR results in decreased sialylation of membrane glycoconjugates. *Am. J. Physiol.*, **266**, C360-366.

Doublie, S. (1997) Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.*, **276**, 523-530.

Drake, D. and Montie, T.C. (1988) Flagella, motility and invasive virulence of *Pseudomonas aeruginosa*. *J. Gen. Microbiol.*, **134**, 43-52.

Duchaud, E., et al. (2003) The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*. *Nat.Biotechnol.*, **21**, 1307-1313.

Ducruix, A. and Giégé, R. (1992) Crystallization of nucleic acids and proteins, a practical approach. The praticals approach series.

Duran, N. and Menck, C.F. (2001) *Chromobacterium violaceum*: a review of pharmacological and industiral perspectives. *Crit. Rev. Microbiol.*, **27**, 201-222.

E

Eaves-Pyles, T.D., Wong, H.R., Odoms, K. and Pyles, R.B. (2001) *Salmonella* flagellin-dependent proinflammatory responses are localized to the conserved amino and carboxyl regions of the protein. *J. Immunol.*, **167**, 7009-7016.

Elliott, G.N., Chen, W.M., Bontemps, C., Chou, J.H., Young, J.P., Sprent, J.I. and James, E.K. (2007) Nodulation of *Cyclopia* spp. (Leguminosae, Papilionoideae) by B*urkholderia tuberum*. *Ann. Bot. (Lond)*, **100**, 1403-1411.

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

F

Figurski, D.H. and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc. Natl. Acad. Sci. USA, 76, 1648–1652.

Flannagan, R.S., Linn, T. and Valvano, M.A. (2008) A system for the construction of targeted unmarked gene deletions in the genus *Burkholderia*. Environ. Microbiol., 10, 1652-1660.

Fotinou, C., Emsley, P., Black, I., Ando, H., Ishida, H., Kiso, M., Sinha, K.A., Fairweather, N.F. and Isaacs, N.W. (2001) The crystal structure of tetanus toxin Hc fragment complexed with a synthetic GT1b analogue suggests cross-linking between ganglioside receptors and the toxin. *J. Biol. Chem.*, **276**, 32274-32281.

G

Gerlach, R.G. and Hensel, M. (2007) Protein secretion systems and adhesins: the molecular armory of Gram-negative pathogens. *Int. J. Med. Microbiol.*, **297**, 401-415.

Gibson, G.A., Hill, W.G. and Weisz, O.A. (2000) Evidence against the acidification hypothesis in cystic fibrosis. *Am. J. Physiol. Cell Physiol.*, **279**, C1088-1099.

Gilboa-Garber, N. (1982) *Pseudomonas aeruginosa* lectins. *Methods Enzymol.*, **83**, 378-385.

Gilboa-Garber, N., Katcoff, D.J. and Garber, N.C. (2000) Identification and characterization of *Pseudomonas aeruginosa* PA-IIL lectin gene and protein compared to PA-IL. *FEMS Immunol. Med. Microbiol.*, **29**, 53-57.

Gillis, M., Van Tan, M., Bardin, R., Goor, M., Hebbar, P., Willems, A.S., P., Kersters, K., Heulin, T. and Fernandez, M.P. (1995) Polyphasic taxonomy in the genus *Burkholderia* leading to an amended description of the genus and proposition of *Burkholderia vietnamensis* sp. nov. for N2-fixing isolates from rice in Vietnam. *Int. J. Syst. Bacteriol.*, **45**, 274-289.

Glendinning, K.J., Parsons, Y.N., Duangsonk, K., Hales, B.A., Humphreys, D., Hart, C.A. and Winstanley, C. (2004) Sequence divergence in type III secretion gene clusters of the *Burkholderia cepaci*a complex. *FEMS Microbiol. Lett.*, **235**, 229-235.

Goldstein, I.J., Hughes, R.C., Monsigny, M., Osawa, T. and Sharon, N. (1980) What should be called a lectin? . *Nature*, **285**, 66.

Goldstein, R., Sun, L., Jiang, R.Z., Sajjan, U., Forstner, J.F. and Campanelli, C. (1995) Structurally variant classes of pilus appendage fibers coexpressed from *Burkholderia (Pseudomonas)* cepacia. *J. Bacteriol.*, **177**, 1039-1052.

Gotschlich, A., et al. (2001) Synthesis of multiple N-acylhomoserine lactones is wide-spread among the members of the *Burkholderia cepacia* complex. *Syst. Appl. Microbiol.*, **24**, 1-14.

Govan, J.R. and Deretic, V. (1996) Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiol. Rev.*, **60**, 539-574.

Govan, J.R., Brown, A.R. and Jones, A.M. (2007) Evolving epidemiology of *Pseudomonas aeruginosa* and the *Burkholderia cepacia* complex in cystic fibrosis lung infection. *Future Microbiol.*, **2**, 153-164.

Govan, J.R., Brown, P.H., Maddison, J., Doherty, C.J., Nelson, J.W., Dodd, M., Greening, A.P. and Webb, A.K. (1993) Evidence for transmission of *Pseudomonas cepacia* by social contact in cystic fibrosis. *Lancet*, **342**, 15-19.

Greco, A., Ho, J.G., Lin, S.J., Palcic, M.M., Rupnik, M. and Ng, K.K. (2006) Carbohydrate recognition by Clostridium difficile toxin A. *Nature Struct. Biol.*, **13**, 460-461.

Groux-Degroote, S., Krzewinski-Recchi, M.A., Cazet, A., Vincent, A., Lehoux, S., Lafitte, J.J., Van Seuningen, I. and Delannoy, P. (2008) IL-6 and IL-8 increase the expression of glycosyltransferases and sulfotransferases involved in the biosynthesis of sialylated and/or sulfated Lewisx epitopes in the human bronchial mucosa. *Biochem. J.*, **410**, 213-223.

H

Haase, G., Skopnik, H., Groten, T., Kusenbach, G. and Posselt, H.G. (1991) Long-term fungal cultures from sputum of patients with cystic fibrosis. *Mycoses*, **34**, 373-376.

Hales, B.A., Morgan, J.A., Hart, C.A. and Winstanley, C. (1998) Variation in flagellin genes and proteins of *Burkholderia cepacia*. *J. Bacteriol.*, **180**, 1110-1118.

Hill, W.G., Harper, G.S., Rozaklis, T. and Hopwood, J.J. (1997) Sulfation of chondroitin/dermatan sulfate by cystic fibrosis pancreatic duct cells is not different from control cells. *Biochem. Mol. Med.*, **62**, 85-94.

Hirabayashi, J. (2004) Lectin-based structural glycomics : glycoproteomics and glycan profiling. *Glycoconj J.* **21**, 35-40

Holden, M.T., et al. (2004) Genomic plasticity of the causative agent of melioidosis, *Burkholderia pseudomallei*. *Proc. Natl. Acad. Sci USA*, **101**, 14240-14245.

Huber, B., Feldmann, F., Kothe, M., Vandamme, P., Wopperer, J., Riedel, K. and Eberl, L. (2004) Identification of a novel virulence factor in *Burkholderia cenocepacia* H111 required for efficient slow killing of Caenorhabditis elegans. *Infect. Immun.*, **72**, 7220-7230.

Hughes, J.E., Stewart, J., Barclay, G.R. and Govan, J.R. (1997) Priming of neutrophil respiratory burst activity by lipopolysaccharide from *Burkholderia cepacia*. *Infect. Immun.*, **65**, 4281-4287.

Hung, C.S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C.G., Strouse, R., Langermann, S., Waksman, G. and Hultgren, S.J. (2002) Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.*, **44**, 903-915.

I

Imberty, A. and Varrot, A. (2008) Microbial recognition of human cell surface glycoconjugates. *Curr. Opin. Struct. Biol.*, **18**, 567-576.

Imberty, A., Mitchell, E.P. and Wimmerová, M. (2005) Structural basis for high affinity glycan recognition by bacterial and fungal lectins. *Curr. Opin. Struct. Biol.*, **15**, 525-534.

Isles, A., Maclusky, I., Corey, M., Gold, R., Prober, C., Fleming, P. and Levison, H. (1984) Pseudomonas cepacia infection in cystic fibrosis: an emerging problem. *J. Pediatr.*, **104**, 206-210.

J

Jiang, X., Hill, W.G., Pilewski, J.M. and Weisz, O.A. (1997) Glycosylation differences between a cystic fibrosis and rescued airway cell line are not CFTR dependent. *Am. J. Physiol.*, **273**, L913-920.

Johnson, W.M., Tyler, S.D. and Rozee, K.R. (1994) Linkage analysis of geographic and clinical clusters in *Pseudomonas cepacia* infections by multilocus enzyme electrophoresis and ribotyping. *J. Clin. Microbiol.*, **32**, 924-930.

Johnston, R.B., Jr. (2001) Clinical aspects of chronic granulomatous disease. *Curr. Opin. Hematol.*, **8**, 17-22.

Jones, A.M., Dodd, M.E., Govan, J.R., Barcus, V., Doherty, C.J., Morris, J. and Webb, A.K. (2004) *Burkholderia cenocepacia* and *Burkholderia multivorans*: influence on survival in cystic fibrosis. *Thorax*, **59**, 948-951.

K

Kerem, B., Rommens, J.M., Buchanan, J.A., Markiewicz, D., Cox, T.K., Chakravarti, A., Buchwald, M. and Tsui, L.C. (1989) Identification of the cystic fibrosis gene: genetic analysis. *Science*, **245**, 1073-1080.

Knowles, M.R. and Boucher, R.C. (2002) Mucus clearance as a primary innate defense mechanism for mammalian airways. *J. Clin. Invest.*, **109**, 571-577.

Kostlanová, N., Mitchell, E.P., Lortat-Jacob, H., Oscarson, S., Lahmann, M., Gilboa-Garber, N., Chambat, G., Wimmerová, M. and Imberty, A. (2005) The fucose-binding lectin from *Ralstonia solanacearum*: a new type of -propeller architecture formed by oligomerisation and interacting with fucoside, fucosyllactose and plant xyloglucan. *J. Biol. Chem.*, **280**, 27839-27849.

Kothe, M., Antl, M., Huber, B., Stoecker, K., Ebrecht, D., Steinmetz, I. and Eberl, L. (2003) Killing of *Caenorhabditis elegans* by *Burkholderia cepacia* is controlled by the cep quorumsensing system. *Cell. Microbiol.*, **5**, 343–351.

L

Lamblin, G., et al. (2001) Human airway mucin glycosylation: A combinatory of carbohydrate determinants which vary in cystic fibrosis. *Glycoconj. J.*, **18**, 661-684.

Lamblin, G., Lhermitte, M., Klein, A., Houdret, N., Scharfman, A., Ramphal, R. and Roussel, P. (1991) The carbohydrate diversity of human respiratory mucins: a protection of the underlying mucosa? *Am. Rev. Resp. Dis.*, **144**, S19-24.

Lameignere, E., Malinovská, L., Sláviková, M., Duchaud, E., Mitchell, E.P., Varrot, A., Šedo, O., Imberty, A. and Wimmerová, M. (2008) Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia*. *Biochem. J.*, **411**, 307-318.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283-291.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283-291.

Leir, S.H., Parry, S., Palmai-Pallag, T., Evans, J., Morris, H.R., Dell, A. and Harris, A. (2005) Mucin glycosylation and sulphation in airway epithelial cells is not influenced by cystic fibrosis transmembrane conductance regulator expression. *Am. J. Respir. Cell Mol. Biol.*, **32**, 453-461.

Lessie, T.G., Hendrickson, W., Manning, B.D. and Devereux, R. (1996) Genomic complexity and plasticity of *Burkholderia cepacia*. *FEMS Microbiol. Lett.*, **144**, 117-128.

Lewenza, S., Conway, B., Greenberg, E.P. and Sokol, P.A. (1999) Quorum sensing i*n Burkholderia cepacia*: identification of the LuxRI homologs CepRI. *J. Bacteriol.*, **181**, 748-756.

Lewenza, S., Visser, M.B. and Sokol, P.A. (2002) Interspecies communication between *Burkholderia cepacia* and *Pseudomonas aeruginosa*. *Can. J. Microbiol.*, **48**, 707-716.

Lewis, H.A., et al. (2004) Structure of nucleotide-binding domain 1 of the cystic fibrosis transmembrane conductance regulator. *EMBO J.*, **23**, 282-293.

Ling, H., Boodhoo, A., Hazes, B., Cummings, M.D., Armstrong, G.D., Brunton, J.L. and Read, R.J. (1998) The structure of shiga-like toxin I B-pentamer complexed with an analogue of its receptor Gb3. *Biochemistry*, **37**, 1777-1788.

Lipuma, J.J. (2002) Preventing *Burkholderia cepacia* complex infection in cystic fibrosis: is there a middle ground? *J. Pediatr.*, **141**, 467-469.

Lipuma, J.J., Dasen, S.E., Nielson, D.W., Stern, R.C. and Stull, T.L. (1990) Person-to-person transmission of Pseudomonas cepacia between patients with cystic fibrosis. *Lancet*, **336**, 1094-1096.

Loris, R., Tielker, D., Jaeger, K.-E. and Wyns, L. (2003) Structural basis of carbohydrate recognition by the lectin LecB from *Pseudomonas aeruginosa*. *J. Mol. Biol.*, **331**, 861-870.

Loutet, S.A., Flannagan, R.S., Kooi, C., Sokol, P.A. and Valvano, M.A. (2006) A complete lipopolysaccharide inner core oligosaccharide is required for resistance of *Burkholderia cenocepacia* to antimicrobial peptides and bacterial survival in vivo. *J. Bacteriol.*, **188**, 2073-2080.

Lutter, E., Lewenza, S., Dennis, J.J., Visser, M.B. and Sokol, P.A. (2001) Distribution of quorum-sensing genes in the *Burkholderia cepacia* complex. *Infect. Immun.*, **69**, 4661-4666.

Lyczak, J.B., Cannon, C.L. and Pier, G.B. (2002) Lung infections associated with cystic fibrosis. *Clin. Microbiol. Rev.*, **15**, 194-222.

M

Mahenthiralingam, E., Baldwin, A., Drevinek, P., Vanlaere, E., Vandamme, P., LiPuma, J.J. and Dowson, C.G. (2006) Multilocus sequence typing breathes life into a microbial metagenome. *PLoS ONE*, **1**, e17.

Mahenthiralingam, E., Campbell, M.E., Henry, D.A. and Speert, D.P. (1996) Epidemiology of *Burkholderia cepacia i*nfection in patients with cystic fibrosis: analysis by randomly amplified polymorphic DNA fingerprinting. *J. Clin. Microbiol.*, **34**, 2914-2920.

Mahenthiralingam, E., et al. (2001) Infection with *Burkholderia cepacia* complex genomovars in patients with cystic fibrosis: virulent transmissible strains of genomovar III can replace Burkholderia multivorans. *Clin. Inf. Dis.*, **33**, 1469-1475.

Mahenthiralingam, E., Urban, T.A. and Goldberg, J.B. (2005) The multifarious, multireplicon *Burkholderia cepacia* complex. *Nat. Rev. Microbiol.*, **3**, 144-156.

Mathee, K., et al. (1999) Mucoid conversion of *Pseudomonas aeruginosa* by hydrogen peroxide: a mechanism for virulence activation in the cystic fibrosis lung. *Microbiology*, **145**, 1349-1357.

Matsui, H., Grubb, B.R., Tarran, R., Randell, S.H., Gatzy, J.T., Davis, C.W. and Boucher, R.C. (1998) Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease. *Cell*, **95**, 1005-1015.

McCoy, A.J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr*, **63**, 32-41.

McKenney, D., Brown, K.E. and Allison, D.G. (1995) Influence of *Pseudomonas aeruginosa* exoproducts on virulence factor production in *Burkholderia cepacia*: evidence of interspecies communication. *J. Bacteriol.*, **177**, 6989-6992.

McPherson, A. (1990) Current approaches to macromolecular crystallization. *Eur J Biochem*, **189**, 1-23.

Merckel, M.C., Tanskanen, J., Edelman, S., Westerlund-Wilkström, B., Korhonen, T.K. and Goldman, A. (2003) The structural basis of receptor-binding by *Escherichia coli* associaed with diarrhea and septicemia. *J. Mol. Biol.*, **331**, 897-905.

Merritt, E.A. and Hol, W.G.J. (1995) AB5 toxins. *Curr. Opin. Struct. Biol.*, **5**, 165-171.

Merritt, E.A., Sarfaty, S., Van Den Akker, F., L'hoir, C., Martial, J.A. and Hol, W.G.J. (1994) Crystal structure of cholera toxin B-pentamer bound to receptor GM1 pentasaccharide. *Protein Sci.*, **3**, 166-175.

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 Struct. Biol.*, **9**, 918-921.

Mitchell, E.P., Sabin, C., Šnajdrová, L., Pokorná, M., Perret, S., Gautier, C., Hofr, C., Gilboa-Garber, N., Koča, J., Wimmerová, M. and Imberty, A. (2005) High affinity fucose binding of *Pseudomonas aeruginosa* lectin PA-IIL: 1.0 Å resolution crystal structure of the complex combined with thermodynamics and computational chemistry approaches. *Proteins: Struct. Funct. Bioinfo.*, **58**, 735-748.

Moens, S. and Vanderleyden, J. (1996) Functions of bacterial flagella. *Crit. Rev. Microbiol.*, **22**, 67-100.

Montie, T.C. and Stover, G.B. (1983) Isolation and characterization of flagellar preparations from *Pseudomonas* species. *J. Clin. Microbiol.*, **18**, 452-456.

Mougous, J.D., Gifford, C.A., Ramsdell, T.L. and Mekalanos, J.J. (2007) Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*. *Nat. Cell Biol.*, **9**, 797-803.

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.

N

Navaza, J. (1994) AMoRE - an Automated Package For Molecular Replacement. *Acta Crystallogr A*, **50**, 157-163.

Navaza, J. and Saludjian, P. (1997) AMoRe: An automated molecular replacement program package. In C W Carter, J. and Sweet, R.M. (eds), *Methods Enzymol*. Academic Press, London & New York, pp. 581-594.

Nierman, W.C., et al. (2004) Structural flexibility in the *Burkholderia mallei* genome. *Proc. Natl. Acad. Sci. USA*, **101**, 14246-14251.

Nzula, S., Vandamme, P. and Govan, J.R. (2002) Influence of taxonomic status on the in vitro antimicrobial susceptibility of the *Burkholderia cepacia* complex. *J. Antimicrob. Chemother.*, **50**, 265-269.

Otag, F., Ersoz, G., Salcioglu, M., Bal, C., Schneider, I. and Bauernfeind, A. (2005) Nosocomial bloodstream infections with *Burkholderia stabilis*. *J. Hosp. Infect.*, **59**, 46-52.

O'Toole, G.A. and Kolter, R. (1998) Flagellar and twitching motility are necessary for *Pseudomonas aeruginosa* biofilm development. *Mol. Microb.*, **30**, 295-304.

Palleroni, N.J., Kunisawa, R., Contopoulo, R. and Doudoroff, M. (1973) Nucleic acid homologies in the genus *Pseudomonas*. *Int. J. Syst. Bact.*, **23**, 333-339.

Parke, J.L. and Gurian-Sherman, D. (2001) Diversity of the *Burkholderia cepacia* complex and implications for risk assessment of biological control strains. *Annu. Rev. Phytopathol;*, **39**, 225-258.

Payne, G.W., Ramette, A., Rose, H.L., Weightman, A.J., Jones, T.H., Tiedje, J.M. and Mahenthiralingam, E. (2006) Application of a recA gene-based identification approach to the maize rhizosphere reveals novel diversity in *Burkholderia* species. *FEMS Microbiol. Lett.*, **259**, 126-132.

Perret, S., Sabin, C., Dumon, C., Pokorná, M., Gautier, C., Galanina, O., Ilia, S., Bovin, N., Nicaise, M., Desmadril, M., Gilboa-Garber, N., Wimmerova, M., Mitchell, E.P. and Imberty, A. (2005) Structural basis for the interaction between human milk oligosaccharides and the bacterial lectin PA-IIL of *Pseudomonas aeruginosa*. *Biochem. J.*, **389**, 325-332.

Pokorná, M., Cioci, G., Perret, S., Rebuffet, E., Kostlánová, N., Adam, J., Gilboa-Garber, N., Mitchell, E.P., Imberty, A. and Wimmerová, M. (2006) Unusual entropy driven affinity of *Chromobacterium violaceum* lectin CV-IIL towards fucose and mannose. *Biochemistry*, **45**, 7501-7510.

Pokorná, M., et al. (2006) Unusual entropy driven affinity of *Chromobacterium violaceum* lectin CV-IIL towards fucose and mannose. *Biochemistry*, **45**, 7501-7510.

Porchet, N. and Aubert, J.P. (2004) MUC genes: mucin or not mucin? That is the question. *Med. Sci. (Paris)*, **20**, 569-574.

Pukatzki, S., Ma, A.T., Sturtevant, D., Krastins, B., Sarracino, D., Nelson, W.C., Heidelberg, J.F. and Mekalanos, J.J. (2006) Identification of a conserved bacterial protein secretion system in *Vibrio cholerae* using the Dictyostelium host model systemer. *Proc. Natl. Acad. Sci USA*, **103**, 1528-1533.

R

Rasolomampianina, R., et al. (2005) Nitrogen-fixing nodules from rose wood legume trees (*Dalbergia* spp.) endemic to Madagascar host seven different genera belonging to alpha- and beta-Proteobacteria. *Mol. Ecol.*, **14**, 4135-4146.

Reddi, K., Phagoo, S.B., Anderson, K.D. and Warburton, D. (2003) *Burkholderia cepacia*-induced IL-8 gene expression in an alveolar epithelial cell line: signaling through CD14 and mitogen-activated protein kinase. *Pediatr. Res.*, **54**, 297-305.

Reik, R., Spilker, T. and Lipuma, J.J. (2005) Distribution of *Burkholderia cepacia* complex species among isolates recovered from persons with or without cystic fibrosis. *J. Clin. Microbiol.*, **43**, 2926-2928.

Rhim, A.D., Kothari, V.A., Park, P.J., Mulberg, A.E., Glick, M.C. and Scanlin, T.F. (2000) Terminal glycosylation of cystic fibrosis airway epithelial cells. *Glycoconj. J.*, **17**, 385-391.

Richau, J.A., Leitao, J.H., Correia, M., Lito, L., Salgado, M.J., Barreto, C., Cescutti, P. and Sa-Correia, I. (2000) Molecular typing and exopolysaccharide biosynthesis of *Burkholderia cepacia* isolates from a Portuguese cystic fibrosis center. *J. Clin. Microbiol.*, **38**, 1651-1655.

Riedel, K., et al. (2001) N-acylhomoserine-lactone-mediated communication betwe*en Pseudomonas aeruginosa* and *Burkholderia cepacia* in mixed biofilms. *Microbiology*, **147**, 3249-3262.

Rieger, J. Stoffelbach, F., Cui, D., Imberty, A., Lameignere, E., Putaux, J.L., Jérôme, R. Jérôme, C. and Auzély-Velty, R. (2007) Mannosylated poly(ethylene oxide)-b-poly(ε-caprolactone) diblock copolymers: synthesis, characterization, and interaction with a bacterial lectin. *Biomacromolecules* **8**, 2717-2725

Riordan, J.R., et al. (1989) Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science*, **245**, 1066-1073.

Rogers, G.B., Carroll, M.P., Serisier, D.J., Hockey, P.M., Jones, G. and Bruce, K.D. (2004) Characterization of bacterial community diversity in cystic fibrosis lung infections by use of 16s ribosomal DNA terminal restriction fragment length polymorphism profiling. *J. Clin. Microbiol.*, **42**, 5176-5183.

Rommens, J.M., et al. (1989) Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science*, **245**, 1059-1065.

Rose, M.C. and Voynow, J.A. (2006) Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol. Rev.*, **86**, 245-278.

Rüdiger, H. (1993) Purification of plant lectins. In Gabius, H.J.G., S. (ed), *Lectins and Glycobiology*. Springer, Berlin, pp. 31-46.

S

Saiman, L. and Siegel, J. (2004) Infection control in cystic fibrosis. *Clin. Microbiol. Rev.*, **17**, 57-71.

Saiman, L., Cacalano, G. and Prince, A. (1990) *Pseudomonas cepacia* adherence to respiratory epithelial cells is enhanced by *Pseudomonas aeruginosa*. *Infect. Immun.*, **58**, 2578-2584.

Sajjan, S.U. and Forstner, J.F. (1992) Identification of the mucin-binding adhesin of *Pseudomonas cepacia* isolated from patients with cystic fibrosis. *Infect. Immun.*, **60**, 1434-1440.

Sajjan, U.S. and Forstner, J.F. (1993) Role of a 22-kilodalton pilin protein in binding of *Pseudomonas cepacia* to buccal epithelial cells. *Infect. Immun.*, **61**, 3157-3163.

Sajjan, U.S., Sylvester, F.A. and Forstner, J.F. (2000) Cable-piliated *Burkholderia cepacia* binds to cytokeratin 13 of epithelial cells. *Infect. Immun.*, **68**, 1787-1795.

Salanoubat, M., et al. (2002) Genome sequence of the plant pathogen *Ralstonia solanacearum*. *Nature*, **415**, 497-502.

Scanlin, T.F. and Glick, M.C. (1999) Terminal glycosylation in cystic fibrosis. *Biochim. Biophys. Acta*, **1455**, 241-253.

Schell, M.A., et al. (2007) Type VI secretion is a major virulence determinant in *Burkholderia mallei*. *Mol. Microbiol.*, **64**, 1466-1485.

Schuster, M., Lostroh, C.P., Ogi, T. and Greenberg, E.P. (2003) Identification, timing, and signal specificity of *Pseudomonas aeruginosa* quorum-controlled genes: a transcriptome analysis. *J. Bacteriol.*, **185**, 2066-2079.

Seksek, O., Biwersi, J. and Verkman, A.S. (1996) Evidence against defective trans-Golgi acidification in cystic fibrosis. *J. Biol. Chem.*, **271**, 15542-15548.

Shaw, D., Poxton, I.R. and Govan, J.R. (1995) Biological activity *of Burkholderia (Pseudomonas) cepacia* lipopolysaccharide. *FEMS Immunol. Med. Microbiol.*, **11**, 99-106.

Shimomura, H., Matsuura, M., Saito, S., Hirai, Y., Isshiki, Y. and Kawahara, K. (2003) *Unusual interaction of a lipopolysaccharide isolated from* Burkholderia cepacia with polymyxin B. *Infect. Immun.*, **71**, 5225-5230.

Sixma, T.K., Pronk, S.E., Kalk, K.H., Van Zanten, B.A.M., Berghuis, A.M. and Hol, W.G.J. (1992) Lactose binding to heat-labile enterotoxin revealed by X-ray crystallography. *Nature*, **355**, 561-564.

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.

Sokol, P.A., Sajjan, U., Visser, M.B., Gingues, S., Forstner, J. and Kooi, C. (2003) The CepIR quorum-sensing system contributes to the virulence of *Burkholderia cenocepacia respiratory infections*. *Microbiology*, **149**, 3649-3658.

Speert, D.P., Henry, D., Vandamme, P., Corey, M. and Mahenthiralingam, E. (2002) Epidemiology of *Burkholderia cepacia* complex in patients with cystic fibrosis, Canada. *Emerg. Infect. Dis.*, **8**, 181-187.

Stanier, R.Y., Palleroni, N.J. and Doudoroff, M. (1966) The aerobic pseudomonads: a taxonomic study. *J. Gen. Microbiol.*, **43**, 159-271.

Steffan, R.J., Sperry, K.L., Walsh, M.T., Vainberg, S. and Condee, C.W. (1999) Field-scale evaluation of in situ bioaugmentation for remediation of chlorinated solvents in groundwater. *Environ. Sci. Technol.*, **33**, 2771-2781.

Stein, P.E., Boodhoo, A., Armstrong, G.D., Heerze, L., Cockle, S.A., Klein, M.H. and Read, R.J. (1994) Structure of a pertussis toxin-sugar complex as a model for receptor binding. *Nat. Struct. Biol.*, **1**, 591-596.

Stenmark, P., Dupuy, J., Imamura, A., Kiso, M. and Stevens, R.C. (2008) Crystal structure of botulinum neurotoxin type A in complex with the cell surface co-receptor GT1b-insight into the toxin-neuron interaction. *PLoS Pathog*, **4**, e1000129.

Sturgess, J. and Imrie, J. (1982) Quantitative evaluation of the development of tracheal submucosal glands in infants with cystic fibrosis and control infants. *Am. J. Pathol.*, **106**, 303-311.

Sudakevitz, D., Imberty, A. and Gilboa-Garber, N. (2002) Production, properties and specificity of a new bacterial L-fucose- and D-arabinose-binding lectin of the plant aggressive pathogen *Ralstonia solanacearum* and its comparison to related plant and microbial lectins. *J. Biochem.*, **132**, 353-358.

Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E.P., Lerrer, B., Wimmerova, M., Katcof, f.D.J., Imberty, A. and Gilboa-Garber, N. (2004) A new *Ralstonia solanacearum* high affinity mannose-binding lectin RS-IIL structurally resembling the *Pseudomonas aeruginosa* fucose-specific lectin PA-IIL. *Mol. Microbiol.*, **52**, 691-700.

Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E.P., Lerrer, B., Wimmerova, M., Katcof, f.D.J., Imberty, A. and Gilboa-Garber, N. (2004) A new *Ralstonia solanacearum*

high affinity mannose-binding lectin RS-IIL structurally resembling the *Pseudomonas aeruginosa* fucose-specific lectin PA-IIL. *Mol. Microbiol.*, **52**, 691-700.

Summer, E.J., Gonzalez, C.F., Carlisle, T., Mebane, L.M., Cass, A.M., Savva, C.G., LiPuma, J. and Young, R. (2004) *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. *J. Mol. Biol.*, **340**, 49-65.

Sun, L., et al. (1995) The emergence of a highly transmissible lineage of cbl+ *Pseudomonas (Burkholderia) cepacia* causing CF centre epidemics in North America and Britain. *Nat. Med.*, **1**, 661-666.

T

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*, **151**, 1313-1323.

Tomich, M., Griffith, A., Herfst, C.A., Burns, J.L. and Mohr, C.D. (2003) Attenuated virulence of a *Burkholderia cepacia* type III secretion mutant in a murine model of infection. *Infect. Immun.*, **71**, 1405-1415.

Tomich, M., Herfst, C.A., Golden, J.W. and Mohr, C.D. (2002) Role of flagella in host cell invasion by *Burkholderia cepacia*. *Infect. Immun.*, **70**, 1799-1806.

Tomlin, K.L., Coll, O.P. and Ceri, H. (2001) Interspecies biofilms of *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Can. J. Microbiol.*, **47**, 949-954.

Tran Van, V., Mavingui, P.B., O. and Balandreau, J.H., T. (1994) Promotion de croissance du riz inoculé par une bactérie fixatrice d'azote, *Burkholderia vietnamensis*, isolée d'un sol sulfaté acide du Vietnam. *Agronomie*, **14**, 697-707.

U

Urban, T.A., Griffith, A., Torok, A.M., Smolkin, M.E., Burns, J.L. and Goldberg, J.B. (2004) Contribution of *Burkholderia cenocepacia* flagella to infectivity and inflammation. *Infect. Immun.*, **72**, 5126-5134.

V

Vagin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J Appl Crystallogr*, **30**, 1022-1025.

Vandamme, P., et al. (1997) Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov. *Int. J. Syst. Bacteriol.*, **47**, 1188-1200.

Vanlaere, E., Lipuma, J.J., Baldwin, A., Henry, D., De Brandt, E., Mahenthiralingam, E., Speert, D., Dowson, C. and Vandamme, P. (2008) *Burkholderia latens* sp. nov., *Burkholderia diffusa* sp. nov., *Burkholderia arboris* sp. nov., *Burkholderia seminalis* sp.

nov. and *Burkholderia metallica* sp. nov., novel species within the *Burkholderia cepacia* complex. *Int. J. Syst. Evol. Microbiol.*, **58**, 1580-1590.

Venter, J.C., et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, **304**, 66-74.

Venturi, V., Friscina, A., Bertani, I., Devescovi, G. and Aguilar, C. (2004) Quorum sensing in the *Burkholderia cepacia* complex. *Res. Microbiol.*, **155**, 238-244.

Vermis, K., Vandekerckhove, C., Nelis, H.J. and Vandamme, P.A. (2002) Evaluation of restriction fragment length polymorphism analysis of 16S rDNA as a tool for genomovar characterisation within the *Burkholderia cepacia* complex. *FEMS Microbiol. Lett.*, **214**, 1-5.

Visser, M.B., Majumdar, S., Hani, E. and Sokol, P.A. (2004) Importance of the ornibactin and pyochelin siderophore transport systems in *Burkholderia cenocepacia* lung infections. *Infect. Immun.*, **72**, 2850-2857.

Vriend, G. (1990) WHAT IF: a molecular modelling and druf design program. *J Mol Graph*, **8**, 52-56.

W

Wagner, V.E., Bushnell, D., Passador, L., Brooks, A.I. and Iglewski, B.H. (2003) Microarray analysis of *Pseudomonas aeruginosa* quorum-sensing regulons: effects of growth phase and environment. *J. Bacteriol.*, **185**, 2080-2095.

Winzer, K., Falconer, C., Garber, N.C., Diggle, S.P., Camara, M. and Williams, P. (2000) The *Pseudomonas aeruginosa lectins* PA-IL and PA-IIL are controlled by quorum sensing and by RpoS. *J. Bacteriol.*, **182**, 6401-6411.

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.

X

Xia, B., Royall, J.A., Damera, G., Sachdev, G.P. and Cummings, R.D. (2005) Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis. *Glycobiology*, **15**, 747-775.

Y

Yabuuchi, E., Kosako, Y., Oyaizu, H., Yano, I., Hotta, H., Hashimoto, Y., Ezaki, T. and Arakawa, M. (1992) Proposal of Burkholderia gen. nov. and transfer of seven species of the genus *Pseudomonas* homology group II to the new genus, with the type species Burkholderia cepacia (Palleroni and Holmes 1981) comb. nov. *Microbiol. Immunol.*, **36**, 1251-1275.

Yabuuchi, E., Kosako, Y., Yano, I., Hotta, H. and Nishiuchi, Y. (1995) Transfer of two *Burkholderia* and an *Alcaligenes* species to *Ralstonia* gen. Nov.: Proposal of *Ralstonia pickettii* (Ralston, Palleroni and Doudoroff 1973) comb. Nov., *Ralstonia solanacearum* (Smith 1896) comb. Nov. and *Ralstonia eutropha* (Davis 1969) comb. Nov. *Microbiol. Immunol.*, **39**, 897-904.

ANNEXE 1

Publication sur l'utilisation de BclA pour la caractérisation de micelles

ANNEXE 2

Données brutes obtenues à partir des glycan arrays sur la partie lectine de BclC

Mannosylated Poly(ethylene oxide)-*b*-Poly(ϵ-caprolactone) Diblock Copolymers: Synthesis, Characterization, and Interaction with a Bacterial Lectin

Jutta Rieger,[‡,†,§] Francois Stoffelbach,[†] Di Cui,[‡] Anne Imberty,[‡] Emilie Lameignere,[‡] Jean-Luc Putaux,[‡] Robert Jérôme,[†] Christine Jérôme,[†] and Rachel Auzély-Velty*[,‡]

Centre de Recherches sur les Macromolécules Végétales (CERMAV-CNRS), BP53, 38041 Grenoble cedex 9, France (affiliated with Université Joseph Fourier, and member of the Institut de Chimie Moléculaire de Grenoble), and Center for Education and Research on Macromolecules (CERM), University of Liége, Sart-Tilman B6, B-4000 Liége, Belgium

Received March 27, 2007; Revised Manuscript Received June 15, 2007

A novel bioeliminable amphiphilic poly(ethylene oxide)-*b*-poly(ϵ-caprolactone) (PEO-*b*-PCL) diblock copolymer end-capped by a mannose residue was synthesized by sequential controlled polymerization of ethylene oxide and ϵ-caprolactone, followed by the coupling of a reactive mannose derivative to the PEO chain end. The anionic polymerization of ethylene oxide was first initiated by potassium 2-dimethylaminoethanolate. The ring-opening polymerization of ϵ-caprolactone was then initiated by the ω-hydroxy end-group of PEO previously converted into an Al alkoxide. Finally, the saccharidic end-group was attached by quaternization of the tertiary amine α-end-group of the PEO-*b*-PCL with a brominated mannose derivative. The copolymer was fully characterized in terms of chemical composition and purity by high-resolution NMR spectroscopy and size exclusion chromatography. Furthermore, measurements with a pendant drop tensiometer showed that both the mannosylated copolymer and the non-mannosylated counterpart significantly decreased the dichloromethane/water interfacial tension. Moreover, these amphiphilic copolymers formed monodisperse spherical micelles in water with an average diameter of ∼11 nm as measured by dynamic light scattering and cryo-transmission electron microscopy. The availability of mannose as a specific recognition site at the surface of the micelles was proved by isothermal titration microcalorimetry (ITC), using the BclA lectin (from *Burkholderia cenocepacia*), which interacts selectively with α-D-mannopyranoside derivatives. The thermodynamic parameters of the lectin/mannose interaction were extracted from the ITC data. These colloidal systems have great potential for drug targeting and vaccine delivery systems.

Introduction

Cell surface carbohydrates from glycoproteins and glycolipids play a key role as recognition sites between cells, but also between cells and microorganisms. The recognition mechanisms are based on specific interactions between the saccharide residues and protein receptors, the so-called "lectins".[1] Thus, numerous polymeric materials carrying saccharide moieties, such as linear polymers,[2,3] dendrimers,[4−6] polymer micelles,[7] and nanoparticles,[8] have been developed for analytical, diagnostic, and therapeutic purposes. It was shown that materials that contain saccharides in a polyvalent array can exhibit enhanced binding capacity with lectins (due to "the cluster glycoside effect"), while the monomeric carbohydrate derivatives exhibit only weak affinity to the same lectins.[9−12] Such materials are promising carrier systems in drug delivery because of the specific cellular targeting of the drugs via the membrane lectins, which often participate in the internalization of their ligands.[13] It should be noted that the binding of a sugar to a specific glycoreceptor closely depends on the configuration of the anomeric (C-1) carbon of the sugar. Therefore, carbohydrates should be incorporated stereoselectively in the polymers in order to interact efficiently with the receptor.

Interestingly, dendritic cells express a variety of lectins at their surface, including specific receptors for mannose. These cells capture, process, and display antigens to native T cells, which initiate the cellular immune response.[14] They have thus been identified as a potent target for vaccine delivery in order to initiate adaptive immune responses.[15] The possible specific targeting of dendritic cells makes mannose-coated colloids attractive delivery systems for vaccines.

The purpose of this study was the synthesis of an amphiphilic biodegradable/bioresorbable mannosylated diblock copolymer as a novel precursor of colloidal (drug) delivery systems. This copolymer consists of a hydrophobic, biodegradable poly(ϵ-caprolactone) (PCL) segment and a hydrophilic, bioeliminable poly(ethylene oxide) (PEO) segment, end-capped by a reactive mannose derivative on the PEO side. The amphiphilic and self-assembling properties of the diblock copolymer were studied together with its recognition by a mannose-specific bacterial lectin.

Experimental Section

Materials. Polymerizations were carried out under dry and oxygen-free argon using standard Schlenk techniques. ϵ-Caprolactone (ϵ-CL) (Aldrich, 99%) was dried over calcium hydride under stirring at room temperature for 48 h and purified by vacuum distillation just before

* Corresponding author. E-mail: rachel.auzely@cermav.cnrs.fr; fax: +33 476 54 72 03.

‡ Centre de Recherches sur les Macromolécules Végétales.

† University of Liége.

§ Present address: Laboratoire de Chimie des Polymères, Université Pierre et Marie Curie-Paris 6, CNRS-UMR 7610, 4 Place Jussieu, 75252 Paris Cedex 05, France.

10.1021/bm070342y CCC: $37.00 © 2007 American Chemical Society
Published on Web 08/17/2007

use. Toluene and tetrahydrofuran (THF) were purified by distillation under nitrogen after drying over sodium benzophenone complex. Dichloromethane (CH_2Cl_2) and pyridine were dried by refluxing over calcium hydride for at least 48 h and distilled prior to use. Acetic acid (AcOH), boron trifluoride diethyl etherate ($BF_3{\cdot}OEt_2$) (Fluka, 96%), 2-bromoethanol (Fluka, >95%), 3,3-diethoxy-1-propanol (Aldrich, 98%), diethyl ether (Vel), 2-dimethylamino ethanol (Aldrich, 99.5%), *N*,*N*-dimethylformamide (DMF) (Fluka, over molecular sieve, >99.8%), ethylene oxide (EO) (Messer), 1-*O*-methyl-α-D-mannopyranoside (Sigma, >99%), D-(+)-mannose (Sigma, >99%), methanol (Fluka, over molecular sieve, >99.5%), 2-propanol (^{i}PrOH), toluene (Fluka), and triethylaluminum ($AlEt_3$) (Fluka, 1.9 M in toluene) were used as received. Milli-Q water was used for all the experiments.

NMR Spectroscopy. ^{1}H NMR and ^{13}C NMR experiments were performed at 25 °C using a Bruker DRX400 spectrometer operating at 400 and 100 MHz, respectively. Chemical shifts (δ in ppm) were reported with respect to external tetramethylsilane (TMS = 0 ppm), and calibration was performed using the signal of the residual protons of the solvent as a secondary reference. Deuterium oxide (D_2O) and deuterated chloroform ($CDCl_3$) were purchased from SDS (Vitry, France). Detailed conditions are given in the figure captions.

Mass Spectrometry (MS). Electrospray mass spectra were recorded in the positive mode on a ZabSpec time-of-flight (Micromass, U.K.) mass spectrometer. The mannose derivatives were dissolved in methanol and injected into the electrospray ion source. The capillary voltage was set to 4 kV. Poly(ethylene glycol) standards were used for external calibration.

Size Exclusion Chromatography (SEC). The number-average molecular weight (M_n) and polydispersity (M_w/M_n) were determined by SEC at 45 °C. The chromatograph was equipped with a refractive index detector and two polystyrene gel columns (columns HP PL gel 5 μm, porosity: 10^2, 10^3, 10^4, and 10^5Å, Polymer Laboratories) that were eluted by THF or DMF at a flow rate of 1 mL/min. The columns were calibrated with polystyrene and PEO standards (Polymer Laboratories), respectively.

Dynamic Light Scattering (DLS). The size and size distribution of the micelles were determined by DLS with an ALV5000 digital correlator in combination with an ALV goniometer and an ALV-SIPC photomultiplier. The incident light source was an ionized argon laser (Spectra Physics 2016) emitting at $\lambda = 488$ nm. Micellar solutions at 25 °C were filtered through Millex-GS filters (porosity 0.22 μm) prior to measurements. The intensity autocorrelation functions ($g_2(t,\theta)$) measured at a given angle (θ) were analyzed with the REPES routine[16] according to a continuous distribution of relaxation times. The average relaxation times components were q^2 dependent in the investigated angular range (from 30 to 140°), such that diffusive motions were probed. The term q is the wave vector defined as $q = 4\pi n/(\lambda \sin(\theta/2))$, where n is the refractive index of the solvent.

Interfacial Tension Measurements. The CH_2Cl_2/H_2O interfacial tension was measured at 20 °C with a drop tensiometer (TRACKER, ITConcept, Longessaigne, France) equipped with a Bioblock Scientific Polystat CC2.[17] Dichloromethane solutions of the copolymers of different concentrations were prepared with CH_2Cl_2 previously saturated with Milli-Q water (mixing for 24 h). A drop of constant volume (5−10 μL) of each solution was formed in water (6 mL, preliminarily saturated with CH_2Cl_2), and the dynamic interfacial tension $\gamma(t)$ was determined by analyzing the axial symmetric shape (Laplacian profile) of the pendant drop in Milli-Q water.

Cryo-Transmission Electron Microscopy (Cryo-TEM). The morphology and size of the micelles were determined from cryo-TEM images. According to protocols reported elsewhere,[18−20] thin liquid films of nanoparticle suspensions (at 0.1 wt % solid content) were prepared on NetMesh (Pelco, U.S.A.) "lacey" carbon membranes and quench-frozen in liquid ethane. Once mounted in a Gatan 626 cryo-holder cooled with liquid nitrogen, the samples were transferred to a microscope and observed at low temperature (−180 °C). Images were recorded on Kodak SO163 film, using a Philips CM200 "Cryo" electron microscope operating at 80 kV.

Isothermal Titration Calorimetry (ITC). The specific recognition properties of the mannosylated micelles were examined by ITC using the BclA lectin produced in *E. coli*. BclA is a bacterial lectin from *Burkholderia cenocepacia* that shows strong sequence similarity with the well-characterized PA-IIL lectin from *Pseudomonas aeruginosa*[21] and binds α-D-mannopyranose (αMan) with high affinity ($K_{a\ (BclA/\alpha MeMan)} = 3.6 \times 10^5\ M^{-1}$),[22] much more strongly than Concanavalin A (ConA) ($K_{a\ (ConA/\alpha MeMan)} = 8.2 \times 10^3\ M^{-1}$).[23] Experiments were carried on a Microcal VP−ITC titration microcalorimeter (Northampton, MA). All titrations were made in 0.1 M tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) buffer, pH 7.5, with 0.03 mM $CaCl_2$ at 25 °C. The reaction cell ($V = 1.4478$ mL) contained either (a) a solution of BclA lectin ([monomer] = 0.04 mM) or (b) the micellar solution ($[\text{Mannose}]_{exp} = 0.1$ mM). A series of 30 injections of 10 μL from the computer-controlled 300-μL microsyringe at an interval of 5 min of (a) the micellar solution ($[\text{Mannose}]_{theo} = 0.5$ mM) or (b) the solution of BclA ([monomer] = 2 mM) were performed into the receptor solution while stirring at 297 rpm at 25 °C. The raw experimental data were reported as the amount of heat produced after each injection of ligand as a function of time. The amount of heat produced per injection was calculated by integrating the area under the individual peaks by the instrument software, while taking into account the heat of dilution. The experimental data were fitted to a theoretical titration curve using the instrument software (ORIGIN software (Microcal)), refining the enthalpy change (in kJ/mol), ΔH, the association constant (in L/mol), K_a, and the stoichiometry of the interaction (number of binding sites per receptor), n.[24] Calculations were performed using the "one set of binding sites" model.

Synthesis. Synthesis of the mannose derivatives was monitored by thin-layer chromatography (TLC) on a precoated plate of silica gel 60 F_{254} (layer thickness 0.2 mm; E. Merck), with detection by UV absorption and/or by charring with a 5% (v/v) ethanolic solution of sulfuric acid. Flash column chromatography was performed on silica gel (Merck Gerduran SI 60, 40−63) with the appropriate eluant. The purity of the compounds was determined by ^{1}H NMR spectroscopy and MS, respectively. When no trace of byproducts was observed by NMR and MS, the products were reported as ">95%" pure (no trace by NMR) and ">99% pure" (no trace by MS).

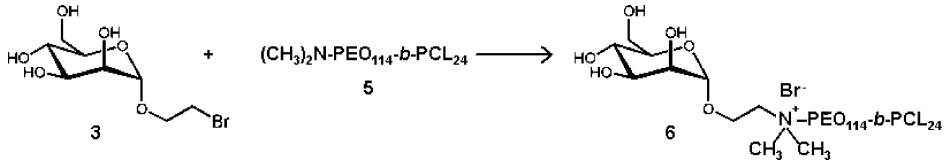
*1,2,3,4,6-Penta-O-acetyl-D-mannopyranose (**1**).* α-D-Mannose pentaacetate was synthesized as commonly reported in the literature.[25] The peracetylated sugar was dried by three azeotropic distillations of toluene just before the next synthesis step.

*2-Bromoethyl-2,3,4,6-tetra-O-acetyl-α-D-mannopyranoside (ManOAc-Br) (**2**).* To a solution of 5.00 g of dried acetylated sugar **1** (0.013 mol) and 1.84 mL of 2-bromoethanol (0.017 mol) in 30 mL of dry CH_2Cl_2, 11.0 mL of $BF_3{\cdot}OEt_2$ (0.090 mol) were added dropwise at 0 °C over a period of 30 min under nitrogen. The progress of the reaction was followed by TLC (7:3 ethyl acetate/ cyclohexane). After 6 h of stirring at room temperature under nitrogen, the reaction mixture was slowly added to 40 mL of ice-cooled water. The aqueous phase was extracted with 20 mL of CH_2Cl_2. The combined organic phases were washed with a saturated aqueous solution of $NaHCO_3$ (20 mL), H_2O (2 × 5 mL), and dried (Na_2SO_4). The solvent was evaporated under vacuum at room temperature, and the crude product was purified by flash chromatography over silica (1:2 ethyl acetate/cyclohexane), to give 4.01 g (8.83 mmol) of **2** as a white crystalline solid.

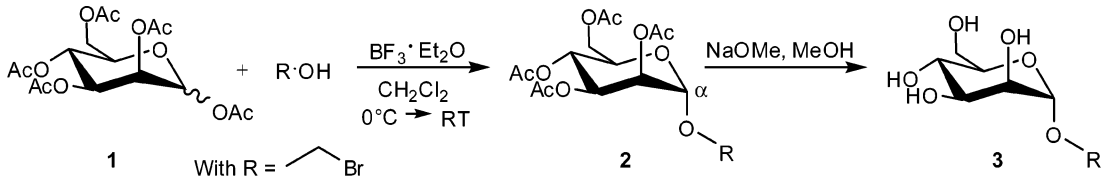
Yield: 69%; purity: >99%; TLC: $R_f = 0.75$ (6.5:3.5 EtOAc/cyclohexane) (H_2SO_4).

^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm): 3.52 (t, *2H*, C*H*$_2$Br), 3.88 (m, *1H*, C*H*HCH$_2$Br), 3.98 (m, *1H*, CH*H*CH$_2$Br), 4.12−4.18 (m, *2H*, H-5, H-6a), 4.29 (m, *1H*, H-6b), 4.88 (d, *1H*, H-1), 5.26 − 5.38 (m, *3H*, H-2, H-4, H-3). ^{13}C NMR ($CDCl_3$, 100 MHz) δ (ppm): 21.98, 22.03, 22.07, 22.19 (4×*C*H$_3$COO), 30.09 (OCH$_2$*C*H$_2$Br), 63.76 (C-6), 67.28 (C-4), 69.78, 70.19, 70.32, 70.69 (*C*H$_2$*C*H$_2$Br, C-5, C-3, C-2), 99.08 (C-1), 171.01, 171.14, 171.32, 171.99 (4 × *C*H$_3$COO). ESI-

Scheme 1. Synthetic Pathway to the Mannosylated Amphiphilic Copolymer **6**



Scheme 2. Synthesis of the 2-Bromoethyl-α-D-mannopyranoside **3**



HRMS: $C_{16}H_{23}O_{10}BrNa$ $[M+Na]^+$, M_{theor} *m/z*: 477.03723, $M_{ESI-HRMS}$ *m/z*: 477.0358.

2-Bromoethyl-α-D-mannopyranoside (ManOH-Br) (3). Compound **2** (1.10 g, 2.41 mmol) was dissolved in 25 mL of dry methanol, and 1.6 mL of 0.82 M NaOMe (1.3 mmol) were added under nitrogen. After 35 h of stirring at room temperature, the solution was neutralized with an Amberlite IR-120 (H^+) resin and filtered off. Solvent was evaporated to give 0.69 g (2.41 mmol) of **3** as a waxy solid.

Yield: 100%; purity: >99%; TLC: R_f = 0.02 (7:3 EtOAc/cyclohexane) (H_2SO_4).

^{1}H NMR (D_2O, 400 MHz) δ (ppm): 3.52−3.60 (m, *3H*, CHHC*H*$_2$-Br, H-4), 3.66−3.95 (m, *7H*, C*HH*CH$_2$Br, H-2, H-3, H-5, H-6a, H-6b), 4.84 (d, *1H*, H-1). ^{13}C NMR (D_2O, 100 MHz) δ (ppm): 32.6 (C_t, OCH$_2$*C*H$_2$Br), 62.2 (C-6), 67.8 (C-4), 68.8, 71.2, 71.8, 74.2 (CH$_2$CH$_2$-Br, C-5, C-3, C-2), 101.0 (C-1). ESI-HRMS: $C_8H_{15}O_6BrNa$ $[M+Na]^+$, M_{theor} *m/z*: 308.99497, $M_{ESI-HRMS}$ *m/z*: 308.9941; $C_8H_{15}O_6BrK$ $[M+K]^+$, M_{theor} *m/z*: 324.96891, $M_{ESI-RHMS}$ *m/z*: 324.9674.

α-N,N-Dimethylaminoethyl Poly(ethylene oxide) (Me$_2$N-PEO-OH) (4). In a flame-dried and argon-purged flask, 200 mL of anhydrous THF, 1.2 mL of *N,N*-dimethylaminoethanol (12 mmol) and 15 mL of a potassium naphthalene/THF solution (0.8 M) were added under argon. After vigorous stirring at room temperature for 15 min, the mixture was added into a 500 mL Parr reactor followed by 60 g of EO (1.36 mol). After 19 h of polymerization at 30 °C, 2-propanol was added, and polymer **4** was precipitated by an excess of diethyl ether and vacuum-dried at 30 °C.

Yield: 95%; $M_{n,NMR}$ = 5470 g/mol; M_w/M_n (SEC) = 1.06; end functionality > 95%.

^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm): 2.22 (s, *6H*, C*H*$_3$−N), 2.47 (m, *2H*, CH$_3$−N−C*H*$_2$), 3.62 (m, *4H*, O−C*H*$_2$−C*H*$_2$−O−).

α-N,N-Dimethylaminoethyl Poly(ethylene oxide)-b-Poly(ε-caprolactone) (Me$_2$N-PEO-b-PCL-OH) (5). In a flame-dried and argon-purged flask, compound **4** (Me$_2$N-PEO-OH) (1.6 g, 0.3 mmol) was dried by azeotropic distillation with toluene (3 × 10 mL). Then 3 mL of dry toluene, one drop of pyridine, and $AlEt_3$ (0.25 mL, 1.9 M, 0.48 mmol) were added. After 20 min at room temperature, 3 mL of CH_2Cl_2 and 0.8 mL of ε-CL (7.2 mmol) were added. After 36 h, the polymerization was stopped by an excess of acetic acid solution (1 M in water), and polymer **5** was recovered by precipitation within a 10-fold excess of heptane and vacuum-dried.

For the preparation of micelles, the polymer was dissolved in acetone, filtered through a 0.2 μm Millex LG (Millipore), and dried in a vacuum.

Yield: 95%; $M_{n,NMR}$ = 8300 g/mol; M_w/M_n (SEC) = 1.10; end functionality > 95%.

^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm): 1.35 (m, *2H*, CH$_2$−C*H*$_2$−CH$_2$), 1.65 (m, *4H*, C*H*$_2$−CH$_2$-C*H*$_2$), 2.27 (s, *6H*, C*H*$_3$−N), 2.30 (m, *2H*, C*H*$_2$−COO), 2.51 (t, *2H*, CH$_3$−N−C*H*$_2$), 3.63 (m, *4H*, O−C*H*$_2$−C*H*$_2$−O−), 4.05 (t, *2H*, COO−C*H*$_2$), 4.22 (t, *2H*, PCL−COO−C*H*$_2$−CH$_2$−O−PEO).

PEO-b-PCL End-Capped by a Mannose Residue via Quaternization (6). A 0.46 g portion of compound **5** (0.06 mmol) was dissolved in 10 mL of dry DMF. Then, 0.28 g of **3** (0.98 mmol) dissolved in 10 mL of DMF was added. After 50 h of stirring under nitrogen at 70 °C, the reaction mixture was concentrated to 5 mL by partial evaporation of the solvent. Then, 10 mL of water was added, and the mixture was introduced in a dialysis membrane (cutoff 3500) and then dialyzed against Milli-Q water for 48 h. The polymer was recovered by freeze-drying as a white solid (0.45 g).

For the preparation of micelles, compound **6** was dissolved in acetone, filtered through a 0.2 μm Millex LG (Millipore), and dried in a vacuum. It was stored at −20 °C.

Yield: 96%; $M_{n,NMR}$ = 8500 g/mol; M_w/M_n (SEC) = 1.19; end-group functionalization ∼ 80%.

^{1}H NMR ($CDCl_3$, 400 MHz) δ (ppm): 1.35 (m, *2H*, CH$_2$−C*H*$_2$−CH$_2$), 1.65 (m, *4H*, C*H*$_2$−CH$_2$−C*H*$_2$), 2.30 (m, *2H*, C*H*$_2$−COO), 3.32 (s, *6H*, C*H*$_3$−N), 3.45−4.05 (m, *10H*, mannose), 3.63 (m, *4H*, O−C*H*$_2$−C*H*$_2$−O−), 4.05 (t, *2H*, COO−C*H*$_2$), 4.22 (t, *2H*, PCL−COO−C*H*$_2$−C*H*$_2$−O−PEO), 4.95 (s, *1H*, H-1, mannose).

Preparation of Micellar Solutions. Typically, 50 mg of amphiphilic copolymer was dissolved in 3 mL of THF (c = 16.6 mg/mL). After complete dissolution of the copolymer, 3 mL of aqueous Tris buffer solution was added dropwise. The mixture was dialyzed (pre-swollen semipermeable membrane: cutoff 3500) for 12 h against Tris buffer that was regularly replaced.

The content of mannose in the micellar solution was determined by acidic hydrolysis of the copolymer. Thus, the micellar solution (1 mL) was heated under acidic conditions (2 M trifluoroacetic acid (0.5 mL) for 4 h at 100 °C). Mannose was quantitatively analyzed after conversion to mannitol acetate by gas chromatography (Hewlett-Packard 5890) and fitted with a flame-ionization detector and a SP2380 column (30 m × 0.53 mm i.d.) using N_2 as a gas carrier.

Results and Discussion

Synthesis of the Targeted Mannosylated Amphiphilic Copolymer. As illustrated in Scheme 1, the strategy for the synthesis of the mannosylated amphiphilic PEO-*b*-PCL diblock copolymer relied on a quaternization reaction between a brominated mannose derivative **3** and a PEO-*b*-PCL diblock copolymer end-capped by a tertiary amine **5**.

The brominated mannose derivative **3** was synthesized by glycosylation of 2-bromoethanol by penta-*O*-acetyl-*β*-D-mannopyranoside **1**,[25] in the presence of an excess of boron trifluoride etherate. The α-mannoside **2** was stereoselectively formed (Scheme 2). After purification by flash chromatography, compound **2** (69% yield) was deacetylated under Zemplén conditions to give quantitatively 2-bromoethyl-α-D-mannopy-

Scheme 3. Synthetic Pathway to Amphiphilic PEO-*b*-PCL Copolymers α-End-Capped by a Tertiary Amine

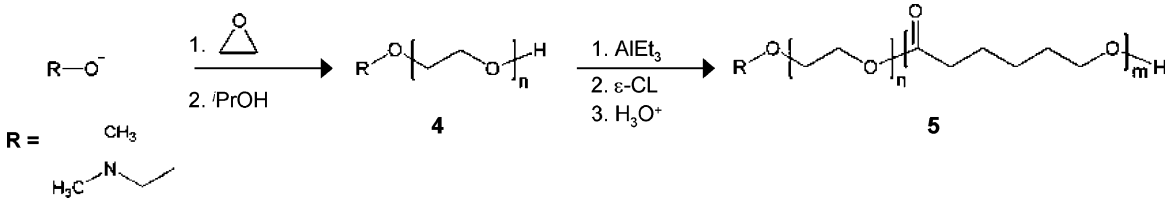


Table 1. Macromolecular Characteristics of α-*N*,*N*-dimethylaminoethyl PEO **4**, α-*N*,*N*-dimethylaminoethyl PEO-*b*-PCL **5,** and the Mannosylated Copolymer **6**

#	polymer	$M_{n,theor}$	$M_{n,NMR}$	$M_{n,SEC}$	M_w/M_n [e]
4	Me_2N−PEO-OH	5090[a]	5470	5050[e]	1.06
5	Me_2N−PEO-*b*-PCL	8210[b]	8300[d]	7100[e]	1.10
6	(Mannose)Me_2N^+− PEO-*b*-PCL	8500[c]	8500	7440[e]	1.19

[a] $M_{n,theor}$ = [[EO]$_0$/ROH × MW_{EO}] + MW (*N*,*N*-dimethylaminoethanol) at 100% of conversion. [b] $M_{n,theor}$ = [[ε-CL]$_0$/[Me_2N−PEO-OH] × $MW_{ε-CL}$] + 5470. [c] $M_{n,theor}$ = MW(1-*O*-ethyl-α-D-mannoside) + 8300. [d] $M_{n,NMR}$ = DP(PCL) × $MW_{ε-CL}$ + 5470; $M_{n,NMR}$ is the number average molecular weight determined by ^{1}H NMR spectroscopy. [e] As determined by SEC in THF calibrated by PEO standards.

roside **3**. The chemical structure of **3** was confirmed by ^{1}H NMR and high-resolution electrospray MS.

The PEO-*b*-PCL copolymer was synthesized by controlled polymerization techniques, in order to tune the molecular weight of the two blocks and to favor a narrow molecular weight distribution.[26] Indeed, for the envisioned PEO-*b*-PCL copolymers to self-assemble into colloidal delivery systems in water, it is essential to precisely control their hydrophilic−lipophilic balance. The hydrophilic PEO block was synthesized first by the "living" anionic polymerization of EO, initiated by potassium *N*,*N*-dimethylaminoethanolate (Scheme 3).

The number average molecular weight (M_n) of the PEO block **4** was determined by ^{1}H NMR from the relative intensity of the proton signals of the PEO chain at 3.62 ppm ($I_{3.62ppm}/4$) and the methyl groups (CH_3N) of the tertiary amine end-group at 2.22 ppm ($I_{2.22ppm}/6$) (Table 1). The polymer **4** was eluted in a single and narrow fraction by SEC (data not shown). The molecular weight determined by NMR and SEC (using a PEO standard) agreed with the value predicted by the monomer/initiator ratio and the monomer conversion (see Table 1, entry 4).

In a second step, the ω-hydroxyl group of the PEO chain was reacted with $AlEt_3$ and converted into an Al alkoxide, which is the actual initiator of the ring-opening polymerization of ε-CL.[27] The length of the PCL block was controlled by the ε-CL over PEO-Al alkoxide molar ratio. The PEO-*b*-PCL diblock copolymer **5** was isolated by precipitation in heptane and characterized by SEC and ^{1}H NMR. The size exclusion chromatogram of the copolymer was narrow (M_w/M_n = 1.10) and unimodal (cf. Supporting Information Figure S1, trace A), indicating the complete initiation of the polymerization of ε-CL by the PEO macroinitiator. The average number molecular weight (M_n) of the PCL block could be determined by ^{1}H NMR spectroscopy (cf. Supporting Information Figure S2). The M_n of the PCL segment was close to the theoretical value, consistent with polymerization control (see Table 1, entry 5).

The final step to the targeted mannosylated copolymer was the quaternization of the tertiary amine-functionalized copolymer by the brominated mannose derivative (Scheme 1). The same reaction was previously used to selectively and quantitatively convert poly(ε-caprolactone)-*co*-poly(γ-bromo-ε-caprolactone) chains into the corresponding pyridinium salts.[28−30] Depending on the copolymer composition, amphiphilic or water-soluble polyesters were prepared in good yields (90%) without degradation of the polyester.

Here, the mannose derivative **3** was covalently bonded to the tertiary amino end-group of the amphiphilic PEO-*b*-PCL copolymer **5** merely by heating the mixture of the two compounds in an appropriate solvent, without assistance of activating agents. For optimization of the reaction, the yield in ammonium salt was determined for different reaction conditions (solvent, bromide/tertiary amine molar ratio, and temperature; cf. Supporting Information Table S1).

Clearly, a large excess of 2-bromoethyl-α-D-mannopyrano side was needed. Moreover, high reaction temperature and long reaction time significantly improved the reaction yield. The best results were obtained within 5 h, in dry DMF at 70 °C, with a large excess of the brominated mannose-derivative (>15 molar equiv) compared to the tertiary amine end-groups (cf. Supporting Information Table S1, entry E). However, longer reaction times (5 days) did not improve the conversion of the quaternization, but led to the degradation of the polymer backbone. (cf. Supporting Information Table S1, entry F)

After reaction, the excess mannose was removed by dialysis against water for 48 h, and the purified product was recovered in good yield by freeze-drying. Figure 1 shows the ^{1}H NMR spectrum of copolymer **6** prepared in 50 h with an excess of 15 equiv of mannose at 70 °C in DMF (cf. Supporting Information Table S1, reaction conditions E).

^{1}H NMR spectroscopy allowed the determination of the degree of glycosylation and the successful elimination of excess mannose; compared to the spectrum of copolymer **5** before derivatization (cf. Supporting Information Figure S2), new resonances in the range of 3.4 to 5.0 ppm were observed, which are characteristic of the protons of the mannose moiety. The signal at 4.95 ppm can be attributed to the anomeric proton of mannose (H_1). Furthermore, the two peaks at 2.27 ppm (6H, CH_3−N) and at 2.51 ppm (2H, N−CH_2), characteristic of the tertiary amino end-group of PEO-*b*-PCL (**5**), disappeared, indicating the total conversion into ammonium salt. Moreover, a resonance observed at 3.32 ppm was found to be 6 times more intense than the anomeric proton H_1 of mannose. It was attributed to the methyl protons of the quaternary ammonium salt, in agreement with the chemical shift of the methyl protons of quaternary ammonium salts found in the literature.[31] The integrals ratio of 1:6 provides evidence for the complete removal of unreacted mannose.

Finally, the same copolymer **6** was analyzed by SEC to confirm that the polymer was not degraded upon quaternization. The chromatogram of the glycosylated polymer remains unimodal and narrow (cf. Supporting Information Figure S1, trace B), consistent with the absence of degradation and side reactions (M_w/M_n = 1.19).

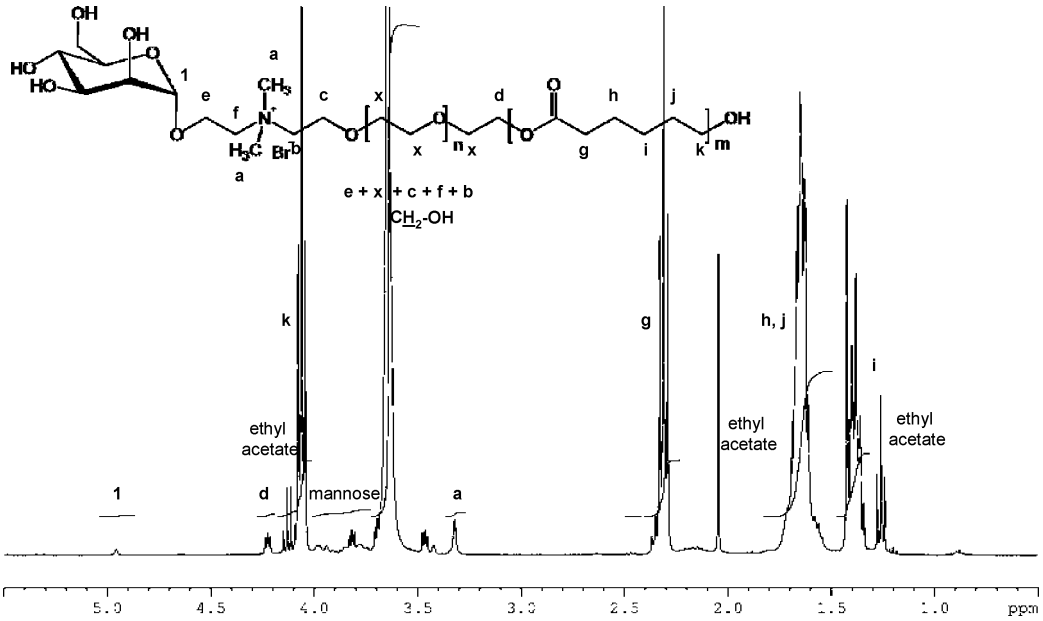


Figure 1. ^{1}H NMR spectrum (400 MHz, 25 °C, $CDCl_3$) of α-mannosylated PEO-*b*-PCL **6** from copolymer **5**.

In conclusion, the tertiary amino end-group of Me_2N-PEO-*b*-PCL was quaternized by brominated-mannose within good yields, without significant chain degradation and without the need of protection/deprotection chemistry.

Amphiphilic Properties of the (Glyco)polymers: Dynamic Interfacial Tension Measurements and Micelle Formation. The amphiphilic properties of the mannosylated copolymer **6** (Table 1) and the parent copolymer **5** (Table 1) were investigated using a drop tensiometer (TRACKER). As the copolymers were barely soluble in water, experiments were performed with solutions in dichloromethane (CH_2Cl_2). The CH_2Cl_2/water interfacial tension (γ) of different copolymer solutions was measured as a function of time. For pure CH_2Cl_2/H_2O (mutually saturated by H_2O and CH_2Cl_2, respectively), the interfacial tension was 28 mN/m and did not vary with time. Figure 2 shows the dynamic interfacial tension isotherms $\gamma(t)$ for solutions of copolymers **5** (Figure 2 A) and **6** (Figure 2 B) at different concentrations, C_S, in CH_2Cl_2.

As a rule, several steps of adsorption were observed with different characteristic relaxation times: the induction step (*lag stage* for $t < \tau_1$), the post-induction step (*post-lag stage* for $\tau_1 < t < \tau_2$), and the final step ($t > \tau_2$). As shown in Figure 2, the *lag time* τ_1 is well defined for relatively low concentrations C_S ($< 5 \times 10^{-3}$ mg/mL), with a slow decrease of $\gamma(t)$ (curves 6, 5, and 4) at the beginning of the adsorption process. This induction time corresponds to the diffusion of the macromolecules from the bulk to the interface. Usually, when at the interface, macromolecules, such as polysoaps and proteins,[32−34] unfold and may adsorb irreversibly by anchoring their hydrophilic part into the polar phase. The *post-lag stage* of adsorption is characterized by a fast decrease in the interfacial tension ($d\gamma/dt$). During this period, the first layer adsorbed at $t < \tau_1$ is "compacted" by newly arriving polymer chains, which penetrate the layer of adsorbed molecules. At the end of the *post-lag stage*, the adsorption layer starts to act as a repulsive barrier against newly arriving macromolecules, as manifested by a remarkable decrease in $d\gamma/dt$ beyond time τ_2. Figure 2 indicates that the mannosylated copolymer **6** has a higher interfacial activity than copolymer **5**. This is confirmed in Figure 3, where the interfacial tension at an aging time of 1000 s is plotted as a function of the copolymer concentration.

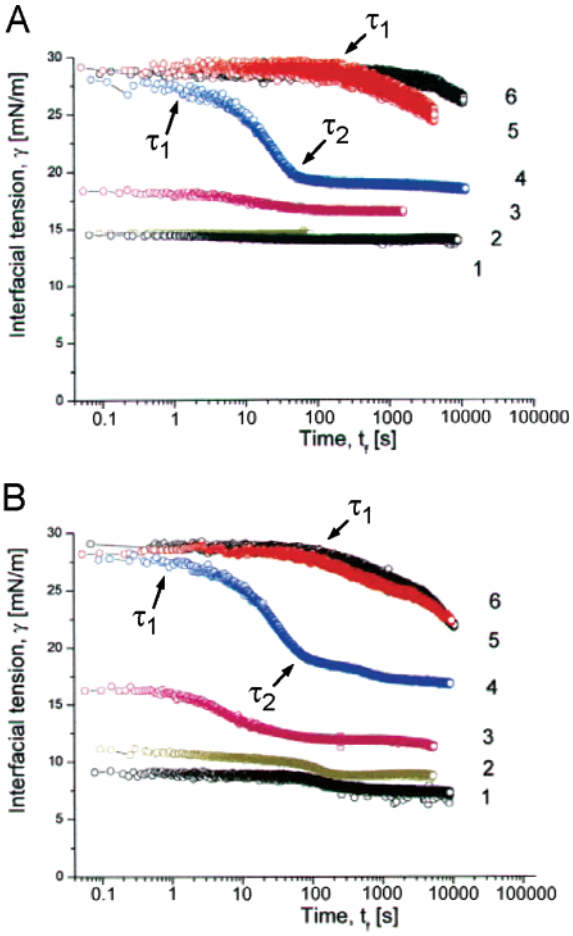


Figure 2. Kinetic curves for (A) copolymer **5**, Me_2N-PEO-*b*-PCL in CH_2Cl_2 at different concentrations C_S: (1) 1; (2) 0.5; (3) 5×10^{-2}; (4) 5×10^{-3}; (5) 5×10^{-4}; and (6) 5×10^{-6} mg/mL, and (B) copolymer **6**, (Mannose)Me_2N^+-PEO-*b*-PCL, Br^- in CH_2Cl_2 at (1) 5; (2) 0.5; (3) 5×10^{-2}; (4) 5×10^{-3}; (5) 5×10^{-4}; and (6) 5×10^{-6} mg/mL.

Both diblock copolymers **5** and **6** decrease the oil−water interfacial tension γ efficiently. At the same copolymer concentration in CH_2Cl_2, the quaternized copolymer **6** (bearing the mannose moiety) is, however, more effective. It must be noted

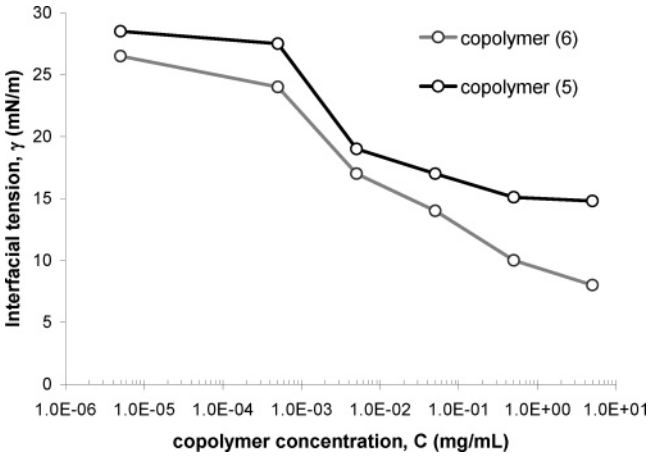


Figure 3. Adsorption isotherms of the interfacial tension γ at the H_2O/CH_2Cl_2 interface for copolymer **5** and its mannosylated counterpart **6**, measured at an aging time of 10^3 s.

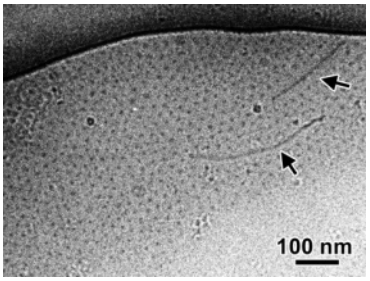


Figure 4. Cryo-TEM image of spherical micelles of the mannosylated copolymer **6** embedded in vitreous ice. The arrows indicate elongated objects that may correspond to cylindrical micelles.

that, at high concentrations ($c > 1$ mg/mL), the γ values are close to the limit of accuracy of the pendant drop method (drops fell down after a few minutes).

As the water solubility of the copolymers is extremely low, micelles from PEO-*b*-PCL diblock copolymers **5** and **6** were prepared by the dialysis method.[35,36] The concentration of the as-prepared micellar aqueous solutions was in the range of 10^{-3} to 10^{-4} M. The critical micelle concentration of PEO-*b*-PCL copolymers of very similar composition (M_n (PEO) = 5000 g/mol, M_n (PCL) = 2500 and 4000 g/mol)[37] has been reported in the range of 10^{-7} M, meaning that, at 10^{-4} M, the large majority of copolymer chains is present as micelles.

The size and shape of the micelles were analyzed by DLS and cryo-TEM. According to DLS analysis, the two copolymers formed monodisperse micelles with an average hydrodynamic radius R_H of ~11 nm. This value is of the same order of magnitude as the R_H reported for micelles of PEO-*b*-PLA[38] diblock copolymers (with M_n(PEO) = 5000, M_n(PLA) = 2000, and R_H = 9.5 nm)[38] and PEO-*b*-PCL[37,39] copolymers (e.g., for M_n(PEO) = 5000 and M_n(PCL) = 2500, and R_H = 11 nm).

The cryo-TEM image in Figure 4 shows micelles of the mannosylated copolymer **6** embedded in a thin film of vitreous ice. A homogeneous distribution of electron-dense particles is observed along with occasional elongated objects. The particles are not in contact, describing in some areas a more or less hexagonal pattern, with an average inter-particle distance of 24−26 nm. Moreover, the particles are at a distance of about 25 nm from the elongated objects. This suggests that the spherical micelles are made of an electron-dense core of PCL blocks surrounded by a more diffuse PEO corona. In the cryo-TEM images, the PCL cores would be at a distance of 25 nm from each other due to the steric repulsion of the PEO coronas. If

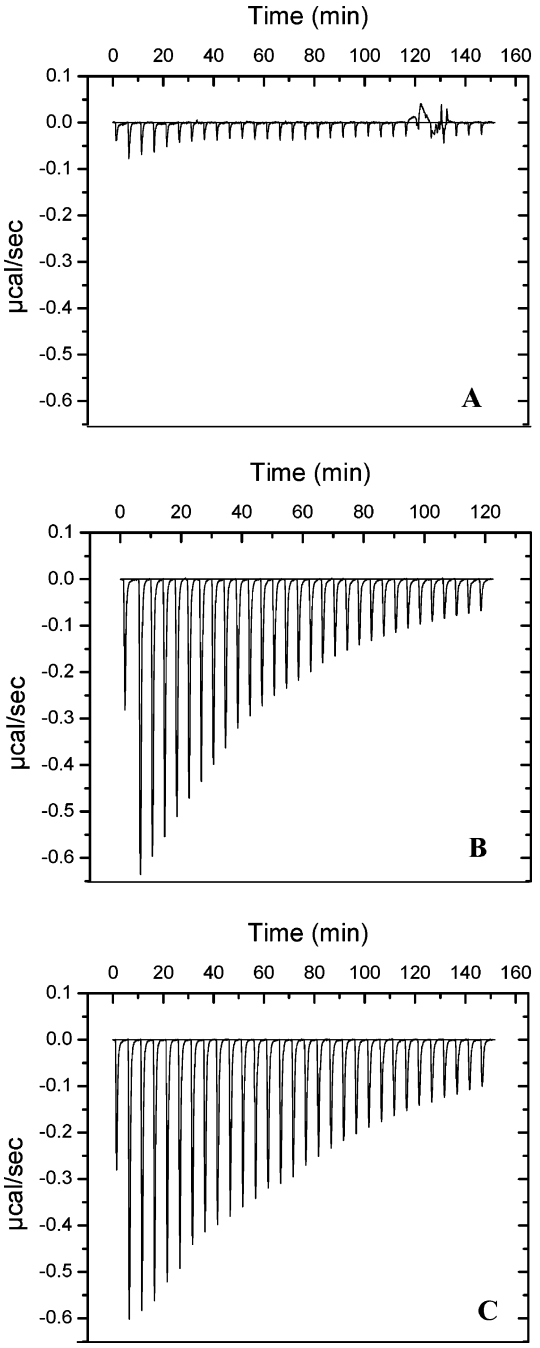


Figure 5. Calorimetric data for the titration of a solution of BclA lectin ([monomer] = 0.04 mM) with (A) a micellar solution of copolymer **5**, (B) a micellar solution of the mannosylated copolymer **6** ($[\text{Mannose}]_{theo}$ = 0.5 mM), and (C) a solution of αMeMan ([Mannose] = 0.5 mM), at 25 °C.

we assume that the particles are not interpenetrating, the radius of the micelles would be approximately 12 nm, in good agreement with the size measured by DLS. The occasionally observed elongated objects may correspond to cylindrical copolymer micelles with a diameter of about 25 nm.

Recognition of Surface-Exposed Mannose: Isothermal Titration Calorimetry. The availability of mannose as a specific recognition site at the surface of the micelles was assessed by ITC. This technique, which has been increasingly used for ligand binding studies during the past 15 years,[24] provides in a single measurement the binding constant, K_a, the enthalpy of binding, ΔH, and the stoichiometry, n, of the interaction. Although ConA is a commonly used lectin to demonstrate specific binding to mannose moieties, the recently

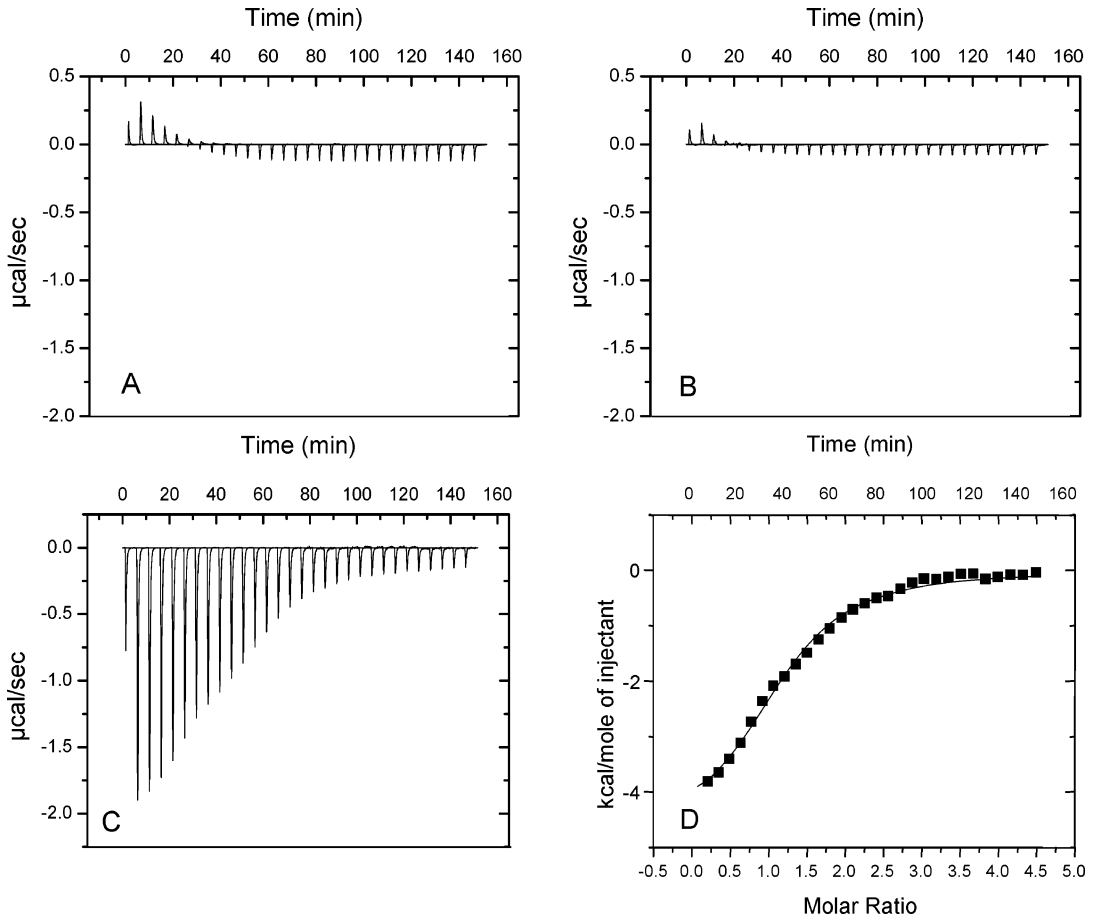


Figure 6. Calorimetric data for (A) the addition of a solution of BclA lectin ([Monomer] = 2 mM) to the micellar solution of the Me_2N−PEO-*b*-PCL copolymer **5**, (B) the addition of a solution of BclA lectin ([Monomer] = 2 mM) to Tris buffer solution, and (C) the titration of the micellar solution of mannosylated copolymer **6** ($[Mannose]_{theo}$ = 0.1 mM) by BclA lectin ([Monomer] = 2 mM) at 25 °C. (D) Integrated values of the peaks of thermogram C (after subtraction of thermogram A) (squares) and fit with the calculated curve (line).

characterized BclA lectin from *Burkholderia cenocepacia* was used in this study. Indeed, because of the low concentration of mannose in the micellar solutions ($[Mannose]_{theo}$ = 0.5 mM) herein prepared by the dialysis method and the relatively low association constant for the α-D-mannopyranoside derivatives (αMeMan) binding to ConA ($K_a = 8.2 \times 10^3$ M^{-1}),[40] the heat evolved during titration was too small for an accurate characterization of the complex formation. In contrast, data could be collected with BclA, which binds αMan much more strongly than ConA ($K_a = 3.6 \times 10^5$ M^{-1}). In a first step, ITC experiments were performed by adding the ligand (micelles with mannose) to the protein receptor in the microcalorimeter cell. Figure 5 compares the thermograms for the calorimetric titration of BclA at 25 °C with micellar solutions of copolymer **5**, mannosylated copolymer **6** ($[Man]_{theo}$ = 0.5 mM), and methyl α-D-mannopyranoside (αMeMan) ([Man] = 0.5 mM).

Compared to the titration with the non-mannosylated micelles, which yielded only small exothermic peaks (thermogram A), a substantial heat release was noted for the titration with the mannosylated micelles (thermogram B). The magnitude of this effect decreased progressively with the progress of the titration, which corresponds to the classically observed increase of occupancy of the receptor sites, similarly to the titration with free αMeMan (thermogram C). Thus, comparison of the thermograms of Figure 5 provides evidence for the selective interaction between mannose at the surface of the micelles and BclA. Furthermore, the similarity observed between thermograms B and C for the injection of similar amounts of mannose residues, strongly suggests that most of the mannose moieties in the micellar assemblies are available for complexation with BclA. In the case of the non-mannosylated micelles (thermogram A), the small heat release can be attributed to the dilution of the micellar solution. Nevertheless, the low concentrations of mannose and lectin used for the titration experiments did not allow the thermodynamic parameters for the mannose binding to BclA to be determined accurately.

Thus, in a second step, reverse titration was carried out. The micellar solution was titrated with the solution of the lectin, such that higher concentrations of ligand and receptor were used, making possible a more reliable determination of the calorimetric parameters. The thermograms for the titration of the mannosylated micelles by BclA (thermogram C), for the addition of BclA to the neat buffer solution (thermogram B), and for the addition of BclA to the Me_2N−PEO-*b*-PCL micelles (thermogram A) are shown in Figure 6. Thermograms A and B are very similar and consistent with the heat of dilution of lectin in the buffer containing micelles or not. Compared to the low endothermic signals in thermograms A and B, a strong exothermic effect is observed in thermogram C as result of the mannose complexation. The very low heat release observed for the last injections indicates that the complexation of mannose at the surface of the micelles is quasi complete. The thermodynamic parameters were extracted from the ITC data on the basis of the simplest model of binding that assumes that individual saccharides bind to individual discrete binding sites in a non-cooperative fashion. For this purpose, the effect of the lectin dilution in the micellar solution (no mannose involved) was taken into account, and thermogram A was subtracted from

Table 2. Thermodynamic Parameters Calculated from ITC[a]

compound	[αMeMan] (mM)	[BclA monomer] (mM)	$K_a \times 10^{-4}$ (M^{-1})	ΔH (kJ/mol)	$T\Delta S$ (kJ/mol)	n (Man: lectin)
αMeMan	3	0.3	36.4 (±1)	−23 (±0.4)	8.8 (±0.3)	0.83 (±0.11)
micelles of copolymer **6**	0.1[b]	2	3.9 (±0.4)	−20 (±0.7)	6.2 (±0.7)	1.18 (±0.03)

[a] Reported values are the mean of at least two experiments. Deviations reported in the fitting procedures are indicated in parenthesis. [b] Concentration of mannose determined by gas chromatography.

thermogram C as reported in Figure 6D. The goodness of fit between the titration data and the calculated curve validates the one-site model (Figure 6D). Table 2 compares the thermodynamic parameters calculated for the complex formation between BclA and free αMeMan, and between BclA and mannose at the surface of the micelles.

ΔH is comparable for the interaction of BclA with free mannose and micelle-bound mannose, consistent with a similar mechanism of binding, although a small loss of enthalpy is observed when the mannose residue is attached on the micelles. The entropy contribution is favorable for both ligands, which is unusual for protein−carbohydrate interactions[23] but was previously reported for this particular family of calcium-dependent bacterial lectins.[41] The fixation of mannose to the micelles results in a less favorable entropy term for binding to BclA. As the result of both enthalpy and entropy loss, the affinity constant of BclA for mannose at the surface of micelles by the lectin is 10 times weaker than that for free mannose. The entropy cost for fixing the ligand was previously reported for the complexation of hydrophobic guests by β-cyclodextrin, free and attached to polymer chains, respectively.[42] Moreover, the stoichiometry n is close to the theoretical value of 1 (one mannose complexed by one lectin monomer), which suggests that nearly all the mannose residues in the micellar solution are available for complexation by BclA.

In contrast to previously described micelles with galactose[43] or mannose[44] at their surface, no cooperative interaction between mannose and the BclA lectin is observed for the micellar aggregates in this study. Nevertheless, as previously pointed out,[5] the behavior of multivalent saccharide ligands is closely related to the assay used to evaluate them. Thus, enhancement in the apparent affinity reported in the literature for saccharide-functionalized micelles, on the basis of agglutination or affinity column inhibition assays, might result from entropically driven protein aggregation phenomena, which strongly depend on the molecular characteristics of the ligand and the protein concentration. Indeed, previous ITC experiments showed that enhanced apparent affinity might be explained by an enthalpy−entropy compensation.[5] Hence, a reduction in the binding enthalpy together with a large increase in entropy imposed by aggregation would lead to an overall increase in the binding constant. Therefore, it would be rash to compare the conclusions about ligand binding in this work with those reported in the literature.

Conclusions

A novel amphiphilic PEO-*b*-PCL diblock copolymer end-capped by a carbohydrate unit was synthesized by coupling a mannose derivative to the α-end-group of a PEO-*b*-PCL diblock copolymer. The length of each block was predetermined by the monomer-to-initiator molar ratio, because controlled polymerization techniques were used. The PEO block was first prepared by living anionic polymerization of EO initiated by potassium *N*,*N*-dimethylaminoethanolate. A tertiary amino α-end-group was accordingly made available for attaching a mannose moiety by quaternization with a brominated mannose derivative. The yield of the end-functionalization was high, and no chain degradation was observed. The interfacial activity of the diblock copolymer, mannosylated or not, was confirmed by measurements of the CH_2Cl_2/water interfacial tension. These amphiphilic polymers formed micelles of similar size, as confirmed by DLS and cryo-TEM. The interaction of the two types of micelles with a mannose-specific lectin (BclA) was investigated for the first time, to the best of our knowledge, by ITC. The thermograms unambiguously confirmed the bioavailability of mannose moieties of the "mannosylated" micelles. When targeting mannose-receptors at the surface of cells, the multivalency of the mannose presentation on the micelles will result in avidity phenomenon that will compensate the somewhat lower affinity observed in solution when comparing to free mannose. The next step of this study will focus on targeted drug delivery using these bioeliminable mannose-decorated micelles as carrier systems.

Acknowledgment. J.R., C.J., and R.J. are much indebted to the "Belgian Science Policy" for financial support in the frame of the "Interuniversity Attraction Poles Programme (IAP-PAI P6/27): Supramolecular Chemistry and Supramolecular Catalysis". J.R. was a recipient of a Marie Curie fellowship. C.J. is an Research Associate by the "Fonds National de la Recherche Scientifique" (FNRS). F.S., C.J., and R.J. also thank the FNRS for a grant to F.S. in the frame of the "SONS-EUROCORES" program. E.L. gratefully acknowledges the MENRT for her thesis grant at the CERMAV. Furthermore, the authors are very grateful to Gérard Chambat for technical support.

Supporting Information Available. SEC traces of α-*N*,*N*-dimethylaminoethyl PEO-*b*-PCL before and after glycosylation (copolymers **5** and **6**); ^{1}H NMR spectrum of α-*N*,*N*-dimethylaminoethyl PEO-*b*-PCL (copolymer **5**) and its discussion; optimization of the quaternization reaction: reaction conditions and yields. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683−720.
(2) Narain, R.; Armes, S. P. *Biomacromolecules* **2003**, *4*, 1746−1758.
(3) Bes, L.; Angot, S. Limer, A.; Haddleton, D. M. *Macromolecules* **2003**, *36*, 2493−2499.
(4) Mangold, S. L.; Cloninger, M. J. *Org. Biomol. Chem.* **2006**, *4*, 2458−2465.
(5) Corbell, J. B.; Lundquist, J. J.; Toone, E. J. *Tetrahedron: Asymmetry* **2000**, *11*, 95−111.
(6) Fernandez-Megia, E.; Correa, J.; Riguera, R. *Biomacromolecules* **2006**, *7*, 3104−3111.
(7) You, L.-C.; Lu, F.-Z.; Li, Z.-C.; Zhang, W.; Li, F.-M. *Macromolecules* **2003**, *36*, 1−4.
(8) Cade, D.; Ramus, E.; Rinaudo, M.; Auzely-Velty, R.; Delair, T.; Hamaide, T. *Biomacromolecules* **2004**, *5*, 922−927.
(9) Nishimura, S. I.; Matsuoka, K.; Kurita, K. *Macromolecules* **1990**, *23*, 4182−4184.
(10) Nishimura, S. I.; Matsuoka, K.; Furuike, T.; Ishii, S.; Kurita, K.; Nishimura, K. M. *Macromolecules* **1991**, *24*, 4236−4241.
(11) Roy, R.; Pon, R. A.; Tropper, F. D.; Andersson, F. O. *J. Chem. Soc., Chem. Commun.* **1993**, 264−265.
(12) Zeng, X.; Murata, T.; Kawagishi, H.; Usui, T.; Kobayashi, K. *Carbohydr. Res.* **1998**, *312*, 209−217.

(13) Smart, J. D.; Nicholls, T. J.; Green, K. L.; Rogers, D. L.; Cook, J. D. *J. Pharm. Sci.* **2000**, *9*, 93−99.
(14) McGreal, E. P.; Miller, J. L.; Gordon, S. *Curr. Opin. Immunol.* **2005**, *17*, 18−24.
(15) Guermonprez, P.; Valladeau, J.; Zitvogel, L.; Thery, C.; Amigorena, S. *Annu. Rev. Immunol.* **2002**, *20*, 621−667.
(16) Stepanek, P. *Dynamic Light Scattering*; Brown, W., Ed.; Oxford University Press: New York, 1993; Chapter 4.
(17) Babak, V.; Boury, F. *Colloid Surf., A* **2004**, *243*, 33−42.
(18) Dubochet, J.; Adrian, M.; Chang, J. J.; Homo, J-C.; Lepault, J.; McDowall, A. W.; Schultz, P. *Q. Rev. Phys.* **1988**, *21*, 129−228.
(19) Harris, J. R. *Negative Staining and Cryoelectron Microscopy: The Thin Films Techniques*; Bios Scientific Publishers: Oxford, U.K., 1997.
(20) Negrete-Herrera, N.; Letoffe, J. M.; Putaux, J. L.; David, L.; Bourgeat-Lami, E. *Langmuir* **2004**, *20*, 1564−1571.
(21) Imberty, A.; Wimmerova, M.; Mitchell, E. P.; Gilboa-Garber, N. *Microb. Infect.* **2004**, *6*, 222−229.
(22) Lameignêre, E.; Imberty, A. Unpublished data, 2006.
(23) Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387−430.
(24) Thomson, J. A.; Ladbury, J. E. In *Biocalorimetry II. Applications of Calorimetry in the Biological Sciences*; Ladbury, J. E., Doyle, M. L., Eds.; J. Wiley & Sons, Ltd: Chichester, England, 2004; Chapter 2.
(25) Conchie, J.; Levvy, G. A. *Methods Carbohydr. Chem.* **1963**, *2*, 345−347.
(26) Duda, A.; Penczek, S. *Makromol. Chem. Macromol. Symp.* **1991**, *47*, 127−140.
(27) Vangeyte, P.; Jérôme, R. *J. Polym. Sci. Polym. Chem.* **2004**, *42*, 1132−1142.
(28) Detrembleur, C.; Mazza, M.; Lou, X.; Halleux, O.; Lecomte, P.; Mecerreyes, D.; Hedrick, J. L.; Jérôme, R. *Macromolecules* **2000**, *33*, 7751−7760.
(29) Detrembleur, C.; Mazza, M.; Halleux, O.; Lecomte, P.; Mecerreyes, D.; Hedrick, J. L.; Jérôme, R. *Macromolecules* **2000**, *33*, 14−18.
(30) Gautier, S.; D'Aloia, V.; Halleux, O.; Mazza, M.; Lecomte, P.; Jérôme, R. *J. Biomater. Sci., Polym. Ed.* **2003**, *14*, 63−85.
(31) Hazziza-Laskar, J.; Nurdin, N.; Helary, G.; Sauvet, G. *J. Appl. Polym. Sci.* **1993**, *50*, 651−662.
(32) Ybert, C.; Di Meglio, J.-M. *Langmuir* **1998**, *14*, 471−475.
(33) Beverung, C. J.; Radke, C. J.; Blanch, H. W. *Biophys. Chem.* **1998**, *70*, 121−132.
(34) Beverung, C. J.; Radke, C. J.; Blanch, H. W. *Biophys. Chem.* **1999**, *81*, 59−80.
(35) Yu, Y.; Eisenberg, A. *J. Am. Chem. Soc.* **1997**, *119*, 8383−8384.
(36) Vangeyte, P.; Gautier, S.; Jérôme, R. *Colloid Surf., A: Physicochem. Eng. Aspects* **2004**, *242*, 203−211.
(37) Jette, K. K.; Law, D.; Schmitt, E. A.; Kwon, G. S. *Pharm. Res.* **2004**, *21*, 1184−1191.
(38) Nagasaki, Y.; Okada, T.; Scholz, C.; Iijima, M.; Kato, M.; Kataoka, K. *Macromolecules* **1998**, *31*, 1473−1479.
(39) Mahmud, A.; Xiong, X.-B.; Lavasanifar, A. *Macromolecules* **2006**, *39*, 9419−9428.
(40) Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387−429.
(41) Pokorna, M.; Cioci, G.; Perret, S.; Rebuffet, E.; Kostlanova, N.; Adam, J.; Gilboa-Garber, N.; Mitchell, E. P.; Imberty, A.; Wimmerova, M. *Biochemistry* **2006**, *45*, 7501−7510.
(42) Charlot, A., Heyraud, A.; Guénot, P.; Rinaudo, M.; Auzély-Velty, R. *Biomacromolecules* **2006**, *7*, 907−913.
(43) Nagasaki, Y.; Yasugi, K.; Yamamoto, Y.; Harada, A.; Kataoka, K. *Biomacromolecules* **2001**, *2*, 1067−1070.
(44) Joralemon, M. J.; Murthy, K. S.; Remsen, E. E.; Becker, M. L.; Wooley, K. L. *Biomacromolecules* **2004**, *5*, 903−913.

BM070342Y

GLYCAN ARRAY OBTENUS AVEC DOMAIN-LECTIN DE BCLB

Chart #	Masterlist Name	RFU	STDEV	SEM	%CV
1	Neu5Aca2-8Neu5Acb-Sp17	1	16	8	1794
2	Neu5Aca2-8Neu5Aca2-8Neu5Acb-Sp8	-2	3	2	-177
3	Neu5Gcb2-6Galb1-4GlcNAc-Sp8	15	6	3	40
4	Galb1-3GlcNAcb1-2Mana1-3(Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp19	22	9	5	42
5	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	14	13	6	95
6	a-D-Gal–Sp8	20	17	8	84
7	a-D-Glc–Sp8	7	5	3	74
8	a-D-Man–Sp8	9	10	5	114
9	a-GalNAc–Sp8	8	8	4	106
10	a-L-Fuc–Sp8	11918	1488	744	12
11	a-L-Fuc–Sp9	11636	1723	861	15
12	a-L-Rha–Sp8	9	6	3	61
13	a-Neu5Ac–Sp8	4	15	7	350
14	a-Neu5Ac–Sp11	13	7	3	56
15	b-Neu5Ac-Sp8	19	17	8	88
16	b-D-Gal–Sp8	28	23	11	82
17	b-D-Glc–Sp8	0	10	5	-2079
18	b-D-Man–Sp8	86	48	24	56
19	b-GalNAc–Sp8	24	20	10	82
20	b-GlcNAc–Sp0	9	14	7	151
21	b-GlcNAc–Sp8	5	10	5	191
22	b-GlcN(Gc)-Sp8	16	13	7	82
23	(Galb1-4GlcNAcb)2-3,6-GalNAca–Sp8	-3	13	7	-434
24	GlcNAcb1-3(GlcNAcb1-4)(GlcNAcb1-6)GlcNAc-Sp8	22	44	22	196
25	[3OSO3][6OSO3]Galb1-4[6OSO3]GlcNAcb-Sp0	18	21	10	112
26	[3OSO3][6OSO3]Galb1-4GlcNAcb-Sp0	-2	14	7	-903
27	[3OSO3]Galb1-4Glcb-Sp8	-2	5	2	-231
28	[3OSO3]Galb1-4(6OSO3)Glcb–Sp0	12	16	8	137
29	[3OSO3]Galb1-4(6OSO3)Glcb–Sp8	996	96	48	10
30	[3OSO3]Galb1-3(Fuca1-4)GlcNAcb–Sp8	11653	2132	1066	18
31	[3OSO3]Galb1-3GalNAca–Sp8	102	33	16	32
32	[3OSO3]Galb1-3GlcNAcb–Sp8	6	9	5	152
33	[3OSO3]Galb1-4(Fuca1-3)GlcNAcb–Sp8	36	6	3	15
34	[3OSO3]Galb1-4[6OSO3]GlcNAcb-Sp8	9	8	4	98
35	[3OSO3]Galb1-4GlcNAcb–Sp0	-5	3	1	-58
36	[3OSO3]Galb1-4GlcNAcb-Sp8	12	10	5	82
37	[3OSO3]Galb–Sp8	19	11	5	58
38	[4OSO3][6OSO3]Galb1-4GlcNAcb-Sp0	4	19	9	519
39	[4OSO3]Galb1-4GlcNAcb-Sp8	13	6	3	47
40	6-H2PO3Mana–Sp8	17	17	9	104
41	[6OSO3]Galb1-4Glcb–Sp0	7	7	4	101
42	[6OSO3]Galb1-4Glcb–Sp8	1	10	5	1013
43	[6OSO3]Galb1-4GlcNAcb–Sp8	21	17	9	80
44	[6OSO3]Galb1-4[6OSO3]Glcb-Sp8	18	8	4	45
45	NeuAca2-3[6OSO3]Galb1-4GlcNAcb–Sp8	6	6	3	97
46	[6OSO3]GlcNAcb–Sp8	9	22	11	238
47	9NAcNeu5Aca-Sp8	17	21	10	118
48	9NAcNeu5Aca2-6Galb1-4GlcNAcb-Sp8	11	9	4	79
49	Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	2622	303	152	12
50	GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	19	7	4	38
51	Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	12	11	5	87
52	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp13	42	54	27	128
53	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp8	11	10	5	92
54	Fuca1-2Galb1-3GalNAcb1-3Gala-Sp9	2939	1039	520	35
55	Fuca1-2Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp9	23034	796	398	3
56	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb–Sp8	8314	307	153	4
57	Fuca1-2Galb1-3GalNAca–Sp8	24280	1722	861	7
58	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	21595	835	418	4
59	Fuca1-2Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp9	24316	1218	609	5

60	Fuca1-2Galb1-3GlcNAcb1-3Galb1-4Glcb–Sp10	4794	1372	686	29
61	Fuca1-2Galb1-3GlcNAcb1-3Galb1-4Glcb–Sp8	4242	400	200	9
62	Fuca1-2Galb1-3GlcNAcb–Sp0	6741	1300	650	19
63	Fuca1-2Galb1-3GlcNAcb–Sp8	7653	787	393	10
64	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	4613	723	361	16
65	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	7176	426	213	6
66	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb–Sp0	5278	598	299	11
67	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb–Sp8	5116	177	89	3
68	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAc–Sp0	27848	958	479	3
69	Fuca1-2Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	27240	394	197	1
70	Fuca1-2Galb1-4GlcNAcb–Sp0	26105	781	391	3
71	Fuca1-2Galb1-4GlcNAcb–Sp8	23386	1349	674	6
72	Fuca1-2Galb1-4Glcb–Sp0	4272	150	75	4
73	Fuca1-2Galb–Sp8	5302	1573	787	30
74	Fuca1-3GlcNAcb-Sp8	17186	3070	1535	18
75	Fuca1-4GlcNAcb–Sp8	16928	1677	838	10
76	Fucb1-3GlcNAcb-Sp8	10	19	10	192
77	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	20	10	5	51
78	GalNAca1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	3931	238	119	6
79	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp0	-6	9	5	-161
80	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb–Sp8	41	8	4	20
81	GalNAca1-3(Fuca1-2)Galb1-4Glcb-Sp0	13	8	4	63
82	GalNAca1-3(Fuca1-2)Galb–Sp8	22	7	4	34
83	GalNAca1-3GalNAcb–Sp8	4	6	3	151
84	GalNAca1-3Galb–Sp8	11	10	5	98
85	GalNAca1-4(Fuca1-2)Galb1-4GlcNAcb-Sp8	2287	290	145	13
86	GalNAcb1-3GalNAca–Sp8	6	8	4	147
87	GalNAcb1-3(Fuca1-2)Galb-Sp8	11	6	3	55
88	GalNAcb1-3Gala1-4Galb1-4GlcNAcb-Sp0	6	8	4	129
89	GalNAcb1-4(Fuca1-3)GlcNAcb-Sp0	30	11	5	36
90	GalNAcb1-4GlcNAcb–Sp0	10	4	2	42
91	GalNAcb1-4GlcNAcb–Sp8	8	13	6	163
92	Gala1-2Galb–Sp8	7	7	3	102
93	Gala1-3(Fuca1-2)Galb1-3GlcNAcb-Sp0	29	14	7	50
94	Gala1-3(Fuca1-2)Galb1-4(Fuca1-3)GlcNAcb-Sp0	29	34	17	117
95	Gala1-3(Fuca1-2)Galb1-4GlcNAc-Sp0	20	13	6	62
96	Gala1-3(Fuca1-2)Galb1-4Glcb-Sp0	35	23	11	65
97	Gala1-3(Fuca1-2)Galb–Sp8	163	40	20	24
98	Gala1-3(Gala1-4)Galb1-4GlcNAcb-Sp8	26	15	7	58
99	Gala1-3GalNAca-Sp8	12	17	8	146
100	Gala1-3GalNAcb–Sp8	23	16	8	72
101	Gala1-3Galb1-4(Fuca1-3)GlcNAcb–Sp8	33	36	18	111
102	Gala1-3Galb1-3GlcNAcb-Sp0	7	11	6	169
103	Gala1-3Galb1-4GlcNAcb–Sp8	13	9	5	68
104	Gala1-3Galb1-4Glcb–Sp0	25	15	7	58
105	Gala1-3Galb–Sp8	-2	16	8	-687
106	Gala1-4(Fuca1-2)Galb1-4GlcNAcb-Sp8	5197	632	316	12
107	Gala1-4Galb1-4GlcNAcb–Sp0	-7	2	1	-32
108	Gala1-4Galb1-4GlcNAcb–Sp8	1	5	3	432
109	Gala1-4Galb1-4Glcb–Sp0	15	24	12	163
110	Gala1-4GlcNAcb–Sp8	3	10	5	337
111	Gala1-6Glcb-Sp8	5	9	5	178
112	Galb1-2Galb–Sp8	17	10	5	58
113	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	24782	2762	1381	11
114	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4GlcNAcb-Sp0	9941	582	291	6
115	Galb1-3(Fuca1-4)GlcNAc–Sp0	18034	295	147	2
116	Galb1-3(Fuca1-4)GlcNAc–Sp8	21206	2514	1257	12
117	Galb1-3(Fuca1-4)GlcNAcb–Sp8	22215	2390	1195	11
118	Galb1-3(Galb1-4GlcNAcb1-6)GalNAca-Sp8	7	15	7	216
119	Galb1-3(GlcNAcb1-6)GalNAca-Sp8	-4	14	7	-305
120	Galb1-3(Neu5Aca2-6)GalNAca-Sp8	11	5	3	46

121	Galb1-3(Neu5Acb2-6)GalNAca-Sp8	11	7	4	67
122	Galb1-3(Neu5Aca2-6)GlcNAcb1-4Galb1-4Glcb-Sp10	20	8	4	37
123	Galb1-3GalNAca-Sp8	-3	9	5	-291
124	Galb1-3GalNAcb–Sp8	7	10	5	152
125	Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	10	3	2	33
126	Galb1-3GalNAcb1-4(Neu5Aca2-3)Galb1-4Glcb-Sp0	25	8	4	34
127	Galb1-3GalNAcb1-4Galb1-4Glcb–Sp8	5	11	5	205
128	Galb1-3Galb–Sp8	16	19	10	120
129	Galb1-3GlcNAcb1-3Galb1-4GlcNAcb-Sp0	6	16	8	242
130	Galb1-3GlcNAcb1-3Galb1-4Glcb–Sp10	45	17	8	37
131	Galb1-3GlcNAcb–Sp0	5	10	5	192
132	Galb1-3GlcNAcb–Sp8	3	19	10	657
133	Galb1-4(Fuca1-3)GlcNAcb–Sp0	-3	9	4	-284
134	Galb1-4(Fuca1-3)GlcNAcb–Sp8	1364	316	158	23
135	Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb-Sp0	118	113	56	95
136	Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb1-4Galb1-4(Fuca1-3)GlcNAcb–Sp0	2666	166	83	6
137	Galb1-4[6OSO3]Glcb–Sp0	3	6	3	222
138	Galb1-4[6OSO3]Glcb–Sp8	7	11	6	169
139	Galb1-4GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	28	28	14	102
140	Galb1-4GalNAcb1-3(Fuca1-2)Galb1-4GlcNAcb-Sp8	23	14	7	58
141	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	14	10	5	71
142	Galb1-4GlcNAcb1-3GalNAca–Sp8	2	7	4	300
143	Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	1054	310	155	29
144	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb–Sp0	0	9	4	2400
145	Galb1-4GlcNAcb1-3Galb1-4GlcNAcb–Sp0	10	13	6	122
146	Galb1-4GlcNAcb1-3Galb1-4Glcb–Sp0	14	4	2	31
147	Galb1-4GlcNAcb1-3Galb1-4Glcb–Sp8	4	10	5	256
148	Galb1-4GlcNAcb1-6(Galb1-3)GalNAca–Sp8	9	5	2	51
149	Galb1-4GlcNAcb1-6GalNAca–Sp8	-4	2	1	-57
150	Galb1-4GlcNAcb–Sp0	4	2	1	46
151	Galb1-4GlcNAcb–Sp8	0	13	7	4145
152	Galb1-4Glcb–Sp0	17	12	6	73
153	Galb1-4Glcb–Sp8	6	8	4	134
154	GlcNAca1-3Galb1-4GlcNAcb-Sp8	3	6	3	197
155	GlcNAca1-6Galb1-4GlcNAcb-Sp8	21	60	30	282
156	GlcNAcb1-2Galb1-3GalNAca–Sp8	15	20	10	135
157	GlcNAcb1-3(GlcNAcb1-6)GalNAca–Sp8	2	4	2	173
158	GlcNAcb1-3(GlcNAcb1-6)Galb1-4GlcNAcb–Sp8	1	14	7	1283
159	GlcNAcb1-3GalNAca–Sp8	16	8	4	47
160	GlcNAcb1-3Galb-Sp8	-2	12	6	-630
161	GlcNAcb1-3Galb1-3GalNAca-Sp8	24	17	8	68
162	GlcNAcb1-3Galb1-4GlcNAcb–Sp0	13	4	2	33
163	GlcNAcb1-3Galb1-4GlcNAcb-Sp8	-1	6	3	-563
164	GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	8	12	6	146
165	GlcNAcb1-3Galb1-4Glcb–Sp0	5	14	7	254
166	GlcNAcb1-4MDPLys	6	10	5	171
167	GlcNAcb1-4(GlcNAcb1-6)GalNAca-Sp8	1	10	5	670
168	GlcNAcb1-4Galb1-4GlcNAcb-Sp8	12	3	2	26
169	(GlcNAcb1-4)6b-Sp8	5	12	6	263
170	(GlcNAcb1-4)5b-Sp8	15	9	4	56
171	GlcNAcb1-4GlcNAcb1-4GlcNAcb–Sp8	-9	10	5	-112
172	GlcNAcb1-6(Galb1-3)GalNAca–Sp8	6	5	2	82
173	GlcNAcb1-6GalNAca–Sp8	8	4	2	45
174	GlcNAcb1-6Galb1-4GlcNAcb-Sp8	11	1	1	10
175	Glca1-4Glcb–Sp8	-6	4	2	-67
176	Glca1-4Glca–Sp8	4	12	6	302
177	Glca1-6Glca1-6Glcb-Sp8	8	11	6	139
178	Glcb1-4Glcb-Sp8	23	14	7	64
179	Glcb1-6Glcb-Sp8	17	6	3	37
180	G-ol-Sp8	10	15	7	143
181	GlcAa-Sp8	5	9	4	189

182	GlcAb-Sp8	3	21	10	697
183	GlcAb1-3Galb-Sp8	1	10	5	931
184	GlcAb1-6Galb-Sp8	7	9	5	127
185	KDNa2-3Galb1-3GlcNAcb–Sp0	3	8	4	254
186	KDNa2-3Galb1-4GlcNAcb–Sp0	13	14	7	114
187	Mana1-2Mana1-2Mana1-3Mana-Sp9	22	17	8	76
188	Mana1-2Mana1-3(Mana1-2Mana1-6)Mana-Sp9	1209	318	159	26
189	Mana1-2Mana1-3Mana-Sp9	247	39	20	16
190	Mana1-6(Mana1-2Mana1-3)Mana1-6(Mana2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	1905	174	87	9
191	Mana1-2Mana1-6(Mana1-3)Mana1-6(Mana2Mana2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	2003	299	150	15
192	Mana1-2Mana1-2Mana1-3(Mana1-2Mana1-3(Mana1-2Mana1-6)Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	78	29	14	37
193	Mana1-3(Mana1-6)Mana–Sp9	831	402	201	48
194	Mana1-3(Mana1-2Mana1-2Mana1-6)Mana-Sp9	1270	389	195	31
195	Mana1-6(Mana1-3)Mana1-6(Mana2Mana1-3)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	2549	436	218	17
196	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb1-4GlcNAcb1-4 GlcNAcb-Sp12	1672	242	121	14
197	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	10	19	9	183
198	Manb1-4GlcNAcb-Sp0	12	8	4	67
199	Fuca1-3(Galb1-4)GlcNAcb1-2Mana1-3(Fuca1-3(Galb1-4)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	1431	686	343	48
200	Neu5Aca2-3Galb1-3GalNAca-Sp8	10	3	2	30
201	NeuAca2-8NeuAca2-8NeuAca2-8NeuAca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	5	11	6	236
202	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb-Sp0	9	9	4	97
203	Neu5Aca2-8Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb–Sp0	12	6	3	51
204	Neu5Aca2-8Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb–Sp0	17	9	5	56
205	Neu5Aca2-8Neu5Aca2-8Neu5Aca-Sp8	16	3	2	22
206	Neu5Aca2-3(6-O-Su)Galb1-4(Fuca1-3)GlcNAcb–Sp8	39	45	23	116
207	Neu5Aca2-3(GalNAcb1-4)Galb1-4GlcNAcb-Sp0	4	9	5	244
208	Neu5Aca2-3(GalNAcb1-4)Galb1-4GlcNAcb-Sp8	-5	12	6	-224
209	Neu5Aca2-3(GalNAcb1-4)Galb1-4Glcb–Sp0	7	5	3	78
210	NeuAca2-3(NeuAca2-3Galb1-3GalNAcb1-4)Galb1-4Glcb-Sp0	15	6	3	39
211	Neu5Aca2-3(Neu5Aca2-6)GalNAca–Sp8	-7	9	4	-119
212	Neu5Aca2-3GalNAca–Sp8	22	16	8	71
213	Neu5Aca2-3GalNAcb1-4GlcNAcb-Sp0	11	9	5	84
214	Neu5Aca2-3Galb1-3(6OSO3)GlcNAc-Sp8	8	3	2	44
215	Neu5Aca2-3Galb1-3(Fuca1-4)GlcNAcb–Sp8	27411	3107	1553	11
216	NeuAca2-3Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb Sp0	24258	1388	694	6
217	Neu5Aca2-3Galb1-3(Neu5Aca2-3Galb1-4)GlcNAcb-Sp8	38	35	17	92
218	Neu5Aca2-3Galb1-3[6OSO3]GalNAca-Sp8	0	14	7	6612
219	Neu5Aca2-3Galb1-3(Neu5Aca2-6)GalNAca–Sp8	18	7	3	39
220	Neu5Aca2-3Galb-Sp8	1	2	1	182
221	NeuAca2-3Galb1-3GalNAcb1-3Gala1-4Galb1-4Glcb-Sp0	21	19	9	87
222	NeuAca2-3Galb1-3GlcNAcb1-3Galb1-4GlcNAcb-Sp0	-8	12	6	-154
223	Neu5Aca2-3Galb1-3GlcNAcb–Sp0	-5	15	7	-270
224	Neu5Aca2-3Galb1-3GlcNAcb–Sp8	12	11	5	86
225	Neu5Aca2-3Galb1-4[6OSO3]GlcNAcb-Sp8	4	15	8	373
226	Neu5Aca2-3Galb1-4(Fuca1-3)(6OSO3)GlcNAcb–Sp8	0	7	4	-2294
227	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb–Sp0	943	74	37	8
228	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb–Sp0	45	27	13	60
229	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb–Sp8	10	7	3	72
230	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb-Sp8	37	22	11	59
231	Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4GlcNAcb-Sp8	66	7	3	11
232	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAc-Sp0	11	16	8	142
233	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb–Sp0	5	10	5	192
234	Neu5Aca2-3Galb1-4GlcNAcb–Sp0	8	3	2	40
235	Neu5Aca2-3Galb1-4GlcNAcb–Sp8	10	18	9	187
236	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	-7	14	7	-193
237	Neu5Aca2-3Galb1-4Glcb–Sp0	6	3	1	46
238	Neu5Aca2-3Galb1-4Glcb–Sp8	3	10	5	291
239	Galb1-4GlcNAcb1-2Mana1-3(Fuca1-3(Galb1-4)GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	5	12	6	225
240	Neu5Aca2-6GalNAca–Sp8	3	17	9	672
241	Neu5Aca2-6GalNAcb1-4GlcNAcb-Sp0	89	124	62	140
242	Neu5Aca2-6Galb1-4[6OSO3]GlcNAcb-Sp8	10	13	6	123

243	Neu5Aca2-6Galb1-4GlcNAcb–Sp0	3	9	5	331
244	Neu5Aca2-6Galb1-4GlcNAcb–Sp8	6	8	4	122
245	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	606	85	43	14
246	Neu5Aca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	10	13	7	136
247	Neu5Aca2-6Galb1-4Glcb–Sp0	19	25	13	132
248	Neu5Aca2-6Galb1-4Glcb–Sp8	2	24	12	1460
249	Neu5Aca2-6Galb–Sp8	7	10	5	149
250	Neu5Aca2-8Neu5Aca-Sp8	12	13	6	106
251	Neu5Aca2-8Neu5Aca2-3Galb1-4Glcb–Sp0	18	15	7	83
252	Neu5Acb2-6GalNAca–Sp8	16	13	6	80
253	Neu5Acb2-6Galb1-4GlcNAcb-Sp8	15	7	3	43
254	Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp21	13	9	5	72
255	Neu5Gca2-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	8010	1196	598	15
256	Neu5Gca2-3Galb1-3GlcNAcb-Sp0	11	10	5	94
257	Neu5Gca2-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	3	6	3	207
258	Neu5Gca2-3Galb1-4GlcNAcb–Sp0	3	16	8	598
259	Neu5Gca2-3Galb1-4Glcb–Sp0	111	91	45	82
260	Neu5Gca2-6GalNAca–Sp0	8	7	3	86
261	Neu5Gca2-6Galb1-4GlcNAcb–Sp0	9	15	8	166
262	Neu5Gca–Sp8	10	12	6	119
263	[3OSO3]Galb1-4(Fuca1-3)(6OSO3)Glc-Sp0	15262	924	462	6
264	[3OSO3]Galb1-4(Fuca1-3)Glc-Sp0	12378	312	156	3
265	[3OSO3]Galb1-4[Fuca1-3][6OSO3]GlcNAc-Sp8	10	16	8	163
266	[3OSO3]Galb1-4[Fuca1-3]GlcNAc-Sp0	0	5	3	-2488
267	Fuca1-2[6OSO3]Galb1-4GlcNAc-Sp0	19777	1187	593	6
268	Fuca1-2Galb1-4[6OSO3]GlcNAc-Sp8	19966	732	366	4
269	Fuca1-2[6OSO3]Galb1-4[6OSO3]Glc-Sp0	17682	785	392	4
270	Fuca1-2-(6OSO3)-Galb1-4Glc-Sp0	15256	909	454	6
271	Fuca1-2-Galb1-4[6OSO3]Glc-Sp0	15514	972	486	6
272	Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	20407	703	352	3
273	Galb1-3-(Galb1-4GlcNacb1-6)GalNAc-Sp14	-5	10	5	-221
274	Galb1-3(GlcNacb1-6)GalNAc-Sp14	8	9	4	115
275	Galb1-3-(Neu5Aa2-3Galb1-4GlcNacb1-6)GalNAc-Sp14	3	3	1	92
276	Galb1-3GalNAc-Sp14	-9	8	4	-94
277	Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	-6	7	3	-111
278	Galb1-4[Fuca1-3][6OSO3]GlcNAc-Sp0	25	5	2	20
279	Galb1-4[Fuca1-3][6OSO3]Glc-Sp0	9121	535	267	6
280	Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	19329	2301	1150	12
281	Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	8	9	4	115
282	Neu5Aca2-3Galb1-3GlcNAcb1-3Galb1-3GlcNAcb-Sp0	13	15	7	111
283	Neu5Aca2-3Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	2	6	3	260
284	[3OSO3]Galb1-4[6OSO3]GlcNAcb-Sp0	20	12	6	58
285	[3OSO3][4OSO3]Galb1-4GlcNacb-SpSp0	21	20	10	95
286	[6OSO3]Galb1-4[6OSO3]GlcNacb-Sp0	4	6	3	140
287	6-H2PO3Glcb-Sp10	11	17	9	165
288	Gala1-3(Fuca1-2)Galb–Sp18	61	19	10	31
289	Gala1-3GalNAca-Sp16	4	12	6	310
290	Galb1-3GalNAca-Sp16	19	10	5	51
291	Galb1-3(Neu5Aca2-3Galb1-4(Fuca1-3)GlcNAcb1-6)GalNAc–Sp14	28	19	9	66
292	Galb1-3Galb1-4GlcNAcb-Sp8	4	16	8	394
293	Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	19	18	9	99
294	Galb1-4GlcNAcb1-3(Galb1-4GlcNAcb1-6)Galb1-4GlcNAc-Sp0	-8	14	7	-180
295	Galb1-4GlcNAcb1-3(GlcNAcb1-6)Galb1-4GlcNAc-Sp0	24	16	8	66
296	Galb1-4GlcNAca1-6Galb1-4GlcNAcb-Sp0	14	4	2	29
297	Galb1-4GlcNAcb1-6Galb1-4GlcNAcb-Sp0	10	5	3	55
298	GalNAca-Sp15	42	48	24	113
299	GalNAca1-3(Fuca1-2)Galb–Sp18	984	311	156	32
300	GalNAcb1-3Galb-Sp8	11	12	6	107
301	GlcAb1-3GlcNAcb-Sp8	14	3	1	20
302	GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	20	12	6	56
303	GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	7	5	2	74

304	GlcNAcb1-3Man-Sp10	-2	9	4	-383
305	GlcNAcb1-4GlcNAcb-Sp10	7	19	9	256
306	GlcNAcb1-4GlcNAcb-Sp12	30	45	22	150
307	HOOC(CH3)CH-3-O-GlcNAcb1-4GlcNAcb-Sp10	7	4	2	57
308	Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	342	97	48	28
309	Mana1-6Manb-Sp10	17	10	5	59
310	Mana1-6(Mana1-3)Mana1-6(Mana1-3)Manb-Sp10	1362	383	192	28
311	Mana1-2Mana1-2Mana1-3(Mana1-2Mana1-6(Mana1-3)Mana1-6)Mana-Sp9	1713	313	156	18
312	Mana1-2Mana1-2Mana1-3(Mana1-2Mana1-6(Mana1-2Mana1-3)Mana1-6)Mana-Sp9	1636	469	235	29
313	Neu5Aca2-3Galb1-3(Neu5Aca2-3Galb1-4GlcNAcb1-6)GalNAc–Sp14	3	6	3	246
314	Neu5Aca2-3Galb1-3(Neu5Aca2-6)GalNAc-Sp14	31	11	5	35
315	Neu5Aca2-3Galb1-3GalNAc–Sp14	3	6	3	210
316	Neu5Aca2-3Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	12	6	3	50
317	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	27	12	6	46
318	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	16	6	3	38
319	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-N(LT)AVL	16	11	5	69
320	Fuca1-2Galb1-3GalNAca-Sp14	18248	1128	564	6
321	Galb1-3(Neu5Aca2-6)GalNAca-Sp14	11	6	3	52
322	Galb1-4GlcNAcb1-3GalNAc-Sp14	2	6	3	371
323	NeuAc(9Ac)a2-3Galb1-4GlcNAcb-Sp0	9	13	7	144
324	NeuAc(9Ac)a2-3Galb1-3GlcNAcb-Sp0	5	8	4	170
325	NeuAca2-6Galb1-4GlcNAcb1-3Galb1-3GlcNAcb-Sp0	7	4	2	62
326	NeuAca2-3Galb1-3(Fuca1-4)GlcNAcb1-3Galb1-3(Fuca1-4)GlcNAcb-Sp0	27133	2062	1031	8
327	NeuAca2-6Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	41	28	14	70
328	Gala1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	12	11	5	89
329	GalNAcb1-3Gala1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	10	14	7	136
330	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	1344	222	111	17
331	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	111	55	28	50
332	(Neu5Aca2-3-Galb1-3)(((Neu5Aca2-3-Galb1-4(Fuca1-3))GlcNAcb1-6)GalNAc–Sp14	20	28	14	139
333	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	82	59	30	72
334	GlcNAca1-4Galb1-4GlcNAcb-Sp0	20	9	5	46
335	GlcNAca1-4Galb1-3GlcNAcb-Sp0	21	13	7	63
336	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4Glcb-Sp0	11	15	7	137
337	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb1-3Galb1-4(Fuca1-3)GlcNAcb-Sp0	821	60	30	7
338	GlcNAca1-4Galb1-4GlcNAcb1-3Galb1-4GlcNAcb-Sp0	54	7	3	12
339	GlcNAca1-4Galb1-3GalNAc-Sp14	12	26	13	225
340	Mana1-3(Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	1078	217	109	20
341	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	20	7	3	35
342	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-6Manb1-4GlcNAcb1-4GlcNAc-Sp12	14	19	10	140
343	Neu5Aca2-6Galb1-4GlcNAcb1-2Mana1-3Manb1-4GlcNAcb1-4GlcNAc-Sp12	-9	15	7	-166
344	Galb1-4GlcNAcb1-2Mana1-3Manb1-4GlcNAcb1-4GlcNAc-Sp12	14	8	4	60
345	Galb1-4GlcNAcb1-2Mana1-6Manb1-4GlcNAcb1-4GlcNAc-Sp12	25	5	2	20
346	Galb1-4GlcNAcb1-2Mana1-3(Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	18	27	14	155
347	GlcNAcb1-2Mana1-3(GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	1040	200	100	19
348	Galb1-4GlcNAcb1-2Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	-3	10	5	-278
349	Galb1-3GlcNAcb1-2Mana1-3(Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4(Fuca1-6)GlcNacb-Sp22	27	43	22	160
350	Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3[Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Sp19	4316	319	160	7
351	(6SO3)GlcNAcb1-3Gal b1-4GlcNAc-b-Sp0	11	5	2	41
352	KDNa2-3Galb1-4(Fuca1-3)GlcNAc-Sp0	19	18	9	95
353	KDNa2-6Galb1-4GlcNAc-Sp0	21	15	8	75
354	KDNa2-3Galb1-4Glc-Sp0	14	6	3	44
355	KDNa2-3Galb1-3GalNAca-Sp14	31	22	11	73
356	Fuca1-2Galb1-3GlcNAcb1-2Mana1-3(Fuca1-2Galb1-3GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	14457	456	228	3
357	Fuca1-2Galb1-4GlcNAcb1-2Mana1-3(Fuca1-2Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	14637	2781	1391	19
358	Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3[Fuca1-2Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAb-Sp20	996	49	24	5
359	Gala1-3Galb1-4GlcNAcb1-2Mana1-3(Gala1-3Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp20	2	11	6	675
360	Mana1-3(Galb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAcb-Sp12	532	186	93	35
361	Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3[Galb1-3(Fuca1-4)GlcNAcb1-2Man a1-6]Manb1-4GlcNAcb1-4(Fuca1-6)GlcNAcb-Sp22	3588	268	134	7
362	Neu5Aca2-6GlcNAcb1-4GlcNAc-Sp21	3	5	2	183
363	Neu5Aca2-6GlcNAcb1-4GlcNAcb1-4GlcNAc-Sp21	1	6	3	431
364	Fuca1-2Galb1-3GlcNAcb1-3[Galb1-4(Fuca1-3)GlcNAcb1-6]Galb1-4Glc-Sp21	10386	491	245	5

365	Galb1-4GlcNAcb1-2(Galb1-4GlcNAcb1-4)Mana1-3[Galb1-4GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAc-Sp21	447	138	69	31
366	GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3[GalNAca1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Sp20	16	15	8	94
367	Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-3[Gala1-3(Fuca1-2)Galb1-4GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Sp20	27	4	2	15
368	Gala1-3Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-3[Gala1-3Galb1-4(Fuca1-3)GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Sp20	1689	45	22	3
369	GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3[GalNAca1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Sp20	23	23	11	98
370	Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-3[Gala1-3(Fuca1-2)Galb1-3GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Sp20	110	58	29	53
371	Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-3[Fuca1-2Galb1-3(Fuca1-4)GlcNAcb1-2Mana1-6]Manb1-4GlcNAcb1-4GlcNAcb-Sp19	812	56	28	7
372	NeuAca2-3Galb1-4GlcNAcb1-3GalNAc-Sp14	35	22	11	64
373	NeuAca2-6Galb1-4GlcNAcb1-3GalNAc-Sp14	15	8	4	51
374	NeuAca2-3[Fuca1-3]Galb1-4GlcNAcb1-3GalNAc-Sp14	56	23	11	41
375	GalNAcb1-4GlcNAcb1-2Mana1-6(GalNAcb1-4GlcNAcb1-2Mana1-6)Manb1-4GlcNAcb1-4GlcNAc-Sp12	27	8	4	29
376	Galb1-3GalNAca1-3(Fuca1-2)Galb1-4Glc-Sp14	2	2	1	144
377	Galb1-3GalNAca1-3(Fuca1-2)Galb1-4GlcNAc-Sp14	81	16	8	19

RESUME

La colonisation des poumons par des germes comme *Pseudomonas aeruginosa* et *Burkholderia cenocepacia* représente la première cause de maladie et de mortalité chez les patients atteints de la mucoviscidose. Quatre gènes de séquence apparentée à celui codant pour la lectine soluble PA-IIL (*LecB*) de *P. aeruginosa* ont été identifiés dans le génome de *B. cenocepaci*a. Ces lectines pourraient être impliquées dans la reconnaissance des cellules épithéliales de l'hôte au stade précoce de l'infection, ou encore dans la formation du biofilm responsable de la résistance aux antibiotiques. Les travaux développés dans cette thèse portent sur l'étude structurale et fonctionnelle de deux de ces lectines : BclA et BclB, ainsi que sur la localisation et la fonction des protéines chez la bactérie. L'utilisation de glycan arrays associée à la microcalorimétrie de titration ont permis de déterminer la spécificité des lectines et leur affinité pour les meilleurs ligands. Les études de localisation ont montré ces lectines cytoplasmiques sont aussi détectées à la surface de la membrane externe. Enfin, l'obtention de cristaux de lectine complexée à son ligand permet de mieux comprendre l'interaction à l'échelle moléculaire.

SUMMARY

Opportunistic infection by pathogens such as *Pseudomonas aeruginosa* and *Burkholderia cenocepacia* is the first cause of morbidity and mortality in cystic fibrosis patients. The opportunistic pathogen *Burkholderia cenocepacia* contains three soluble carbohydrate-binding proteins, related to the fucose-binding lectin PA-IIL from *Pseudomanas aeruginosa*. These lectins could play a role in early stage of infection through specific binding to epithelial cells of hosts. They could also be involved in the building of biofilm that is responsible for resistance to antibiotics. The thesis is focused on structure-function studies of two *B. cenocepacia* lectins, BclA and BclB with the aim to correlate the data to the localisation and function in the bacteria. Glycan array data associated with titration microcalorimetry allowed to determine the specificity of the lectins and the affinity towards the best ligands. Localisation studies demonstrate the presence of the lectins in the bacteria cytoplasm but also on the outer membrane. Finally, crystal structures of lectin complexed with carbohydrate give the molecular basis of the interaction.