



HAL
open science

Caractérisation biochimique et structurale d'une lectine de graine de *Platypodium elegans* Vogel

Raquel Leite

► **To cite this version:**

Raquel Leite. Caractérisation biochimique et structurale d'une lectine de graine de *Platypodium elegans* Vogel. Autre [q-bio.OT]. Université de Grenoble, 2011. Français. NNT : 2011GRENV081 . tel-00683913

HAL Id: tel-00683913

<https://theses.hal.science/tel-00683913>

Submitted on 30 Mar 2012

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.



UNIVERSIDADE
FEDERAL DO CEARÁ

UNIVERSITÉ DE
GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

préparée dans le cadre d'une cotutelle entre l'*Université de Grenoble* et l'*Universidade Federal do Ceará*

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : le 6 janvier 2005 -7 août 2006

Présentée par

« **Raquel / BENEVIDES** »

Thèse dirigée par « **Anne/IMBERTY** » et « **Benildo/CAVADA** »
codirigée par « **Annabelle/VARROT** »

préparée au sein des **Centre de Recherche sur les Macromolécules Végétales (CERMAV)** et **Laboratório de Moléculas Biologicamente Ativas**

dans l'**École Doctorale Chimie / Sciences du Vivant** et **Programa de Doutorado em Bioquímica**

Caractérisation biochimique et structurale d'une lectine de graines de *Platypodium elegans* Vogel

Thèse soutenue publiquement le « **6 Décembre de 2011** »,
devant le jury composé de :

Mme Anne IMBERTY
DR, CNRS-UJF, Directeur de Thèse

Mme Annabelle VARROT
CR, CNRS-UJF, Membre

Mme Isabelle ANDRÉ
DR, CNRS-INSA, Rapporteur

M. Benildo S. CAVADA
Pr, UFC, Directeur de Thèse

M. Plínio DELATORRE
Pr, UFPB, Rapporteur

M. Alexandre H. SAMPAIO
Pr, UFC, Membre



Thèse en co-tutelle préparée
au Centre d'Étude et de Recherche des Macromolécules Végétales (CERMAV)
– Équipe de Glycobiologie Moléculaire
et à l'Université Fédéral du Ceará (UFC)
– Équipe de Macromolécules Biologiquement Actives (BioMol-Lab)

REMERCIEMENTS...

D'une manière très spéciale à mes deux superviseurs, le professeur Benildo Sousa Cavada et le Dr Anne Imberty, pour la façon dont j'ai été reçue, pour l'orientation et les enseignements. Merci pour leur bienveillante attention qui m'a constamment accompagnée tout au long de cette thèse. Leur encadrement et leurs conseils m'ont été très précieux. Je leur prie d'être assurés de ma sincère reconnaissance.

Je remercie le Dr. Plinio Delatorre et le Dr. Isabelle André pour avoir accepté de juger ce travail de thèse. Je remercie également le Dr. Alexandre Holanda Sampaio pour avoir participé à ce jury.

Je remercie le Dr Anabelle Varrot par sa co-orientation et le soutien dans une grande partie de la thèse, l'attention et la patience, et pour avoir participé à le jury de cette thèse. Je tiens également à remercier tout le soutien de Pr Christelle Breton pour ses enseignements et des conseils. Je remercie également le soutien et les discussions avec le Professeur Dr Celso Nagano Shiniti.

Je tiens à remercier la direction du Département de Biochimie et Biologie Moléculaire (UFC), actuellement représentée par le Prof. Dr Márcio Viana Ramos, pour l'acceptation et l'opportunité, ainsi que pour la direction du CERMAV (UJF), représentée par le Dr Redoane Borsali, pour l'accueil chaleureux. Je remercie également toutes les personnes des le deux institutions qui m'ont soutenue et encouragée tout au long de ces quatre années de thèse.

Ma reconnaissance va également à toute l'équipe BioMol-Lab (UFC): Profs. Bruno e Kyria, Talinhos, Júlia, Eduardo, Sâmia, Ranieri, Ito, Alysson, Rafael, Helton, Júnior, Camila, Suzete, Bruno Lopes, Mayron, Theodora, Thaiz, Joana, Raphael, Cíntia, Claudener, Rômulo, Arthur, Guilherme, Fernando, Kássia, Mayara Torquato, Ivandra, Ronniery, Mayara Queiroz, Vanir, Vinícius, Cláudia, Lia, Davi, Caio, Clareane, Alfa, Cibelle. Mon grand merci !

Je remercie également toute l'équipe de Glycobiologie Moleculaire (CERMAV-UJF) dirigée par Christelle Breton. Un grand merci à Catherine Gautier et Valérie Chazalet pour leurs conseils constructifs et pour leur disponibilité à chaque fois que j'ai eu à solliciter leurs compétences. Merci pour les soins. Je remercie Dr Olivie Lexourel et mes collègues Bertrand Blanchard, Sophie Mathieu, Michael Reynolds, Alessandra Nurisso, Sami Cecioni, Ondraj Sulak, Aymeric Audfray, Nicolas Sapay, Géraldine Ganne, Anita Sarkar et Sumaira Kousar.

Je remercie ma famille, spécialement et particulièrement, ma mère et mes sœurs, qui m'ont donné tout le soutien dans les soins, la compréhension et la patience durant cette phase: vous êtes ma plus grande motivation.

Pour finir, merci mon amour Elton Menezes Leite; tu as été d'une grande aide pendant toutes ces années. Merci d'avoir été toujours là pour moi, de m'avoir motivée quand il le fallait. Merci de tout de mon cœur!

PRÉSENTATION DU TRAVAIL DE THÈSE

Les sujets abordés dans cette thèse se situent dans le domaine de la glycobiochimie structurale, la branche de la biologie qui étudie les structures et les interactions des glycomolécules. Pendant les quatre années de cette thèse, passées pour moitié au Département de Biochimie et Biologie Moléculaire (UFC, Brésil) et au CERMAV (UJF, France) j'ai été confronté à des sujets différents mais complémentaires tels que l'utilisation de la biologie moléculaire, l'analyse structurale des complexes protéine-sucre et la caractérisation thermodynamique des interactions lectine-sucre. Cette thèse peut être divisée en trois parties. La première est dédiée au clonage et à la production de la lectine recombinante, la deuxième décrit les études d'affinité et de spécificité de cette lectine pour les glycannes et la troisième analyse la structure cristalline de la lectine complexée avec différents ligands glucides.

Le premier chapitre est une introduction à la glycobiochimie et à l'objet de notre étude : les lectines. Dans le deuxième chapitre, une attention particulière a été portée aux méthodes permettant l'étude structurale et l'analyse des interactions lectines-sucres. Une introduction aux méthodes expérimentales utilisées est incluse, notamment concernant la biologie moléculaire, les puces à sucres, la microcalorimétrie et la cristallographie. Le troisième chapitre est dédié à la production recombinante d'une lectine de légumineuses à partir du matériel génétique des graines. Le quatrième chapitre est dédié aux interactions lectines-sucres et décrit la spécificité de la lectine PELA envers les glucides et l'analyse de l'affinité pour quelques glycannes. Dans le cinquième chapitre, l'étude structurale des complexes de la lectine avec différents oligosaccharides est abordée. Quelques conclusions et perspectives terminent la présentation de ce travail. À la fin, il est fait un petit point sur les autres travaux développés au cours de cette thèse.

Cette thèse a été rédigée en langue portugaise, mais comprend deux annexes: le résumé du travail de thèse en langue française et le manuscrit en langue anglaise d'une publication qui sera soumise prochainement à *Journal of Biological Chemistry*.

Dans cette version, le résumé de la thèse en langue française est présenté au début, suivie par la version portugaise et par l'article pour publication.

**Caractérisation biochimique
et structurale d'une lectine de
graines de *Platypodium*
elegans Vogel**

(Résumé du Travail de Thèse en langue française)

Sommaire

PRÉSENTATION DU TRAVAIL DE THÈSE	4
LISTE DES PRINCIPALES ABRÉVIATIONS.....	8
I. Introduction :	10
1. Généralités sur la glycobiochimie et les lectines:	10
2. Les lectines de Légumineuses:.....	12
3. Clonage et expression hétérologue des lectines de légumineuses.....	17
4. Étude structurale et thermodynamique de l'interaction lectine-sucre:	19
5. Objectifs du travail de thèse:.....	20
II. Matériels et Méthodes :	22
1. Biologie Moléculaire: Clonage du gène de PELa	22
1.1. Purification de l'ARN	22
1.2. 3'-RACE	22
1.3. 5'-RACE	23
1.4. Clonage dans le vecteur pET-29a:	24
2. Expression et purification de PELa recombinante:	26
2.1. Expression de PELa dans <i>E. coli</i>	26
2.2. Renaturation et purification	26
3. Outils pour l'étude des interactions protéine-sucre.....	27
3.1. Tests d'hémagglutination.....	27
3.2. Glycan Array (puce à sucres).....	27
3.3. Microcalorimétrie isotherme de titration (ITC)	28
4. Cristallographie aux rayons X:	28
4.1. Cristallisation et collecte des données.....	28
4.2. Résolution et affinement des structures cristallographiques.....	29
III. Résultats et discussion :	31
1. Clonage, expression, renaturation et purification de la lectine PELa:	31
2. Caractérisation de la spécificité et de l'affinité de la lectine PELa:.....	32
2.1. Étude des propriétés d'hémagglutination de PELa:.....	32
2.2. Étude de la spécificité de PELa par Glycan Array	34
2.3. Étude de l'affinité de PELa pour les oligosaccharides complexes	36
3. Étude structurale de la lectine PELa en complexe avec différents oligosaccharides.....	40
3.1. Structure cristallographiques de PELa en complexe avec le trimanose	40
3.2. Structure cristallographique de PELa en complexe avec l'heptasaccharide symétrique	47
IV. Conclusion et Perspectives	53

V. Résumé sur les autres travaux réalisés au cours de la thèse.....	54
V. Références Bibliographiques.....	58
ANNEXES.....	67
ANNEXE 1: Caracterização bioquímica e estrutural de uma lectina recombinante de sementes de <i>Platypodium elegans</i> Vogel (Travail de thèse complète en portugais).....	68
ANNEXE 2: Sequence, structure and specificity of a lectin from <i>Platypodium elegans</i> seeds with unusual affinity for asymmetric complex <i>N</i> -glycans (Paper to Journal of Biological Chemistry)	247

LISTE DES PRINCIPALES ABRÉVIATIONS

AAP	<i>abridged anchor primer</i>
ADNc	acide désoxyribonucléique complémentaire
AP	<i>anchor(adapter) primer</i>
ARNm	Acide ribonucléique messenger
CFL	lectine de <i>Cratylia floribunda</i>
CGL	lectine de <i>Canavalia gladiata</i>
ConA	lectine de <i>Canavalia ensiformes</i>
ConBol	lectine de <i>Canavalia boliviana</i>
ConBr	lectine de <i>Canavalia brasiliensis</i>
ConM	lectine de <i>Canavalia maritima</i>
CRD	domaine de reconnaissance des glucides
DGL	lectine de <i>Dioclea grandiflora</i>
DGL	lectine de <i>Dioclea grandiflora</i>
Dgui	lectine de <i>Dioclea guianenses</i>
DMF	diméthylformamide
DO ₆₀₀	densité optique à 600 nm
Dvir	lectine de <i>Dioclea virgata</i>
ESRF	<i>European Synchrotron Radiation Facility</i>
GNA	agglutinine de <i>Galanthus nivalis</i>
IPTG	isopropyl- β -D thiogalactoside
ITC	microcalorimétrie de titration isotherme
K _a	constante d'association
K _d	constante de dissociation
LCA	agglutinine de <i>Lens culinaris</i>
LOL	lectine de <i>Lathyrus ochrus</i>
PAL	lectine de <i>Pterocarpus angolensis</i>
pb	paire de bases
PDB	<i>protein data bank</i>
PELa	lectine de <i>Platypodium elegans</i>

PHA	phytohemagglutinine
RACE	amplification rapide d'extrémités d'ADNc.
SDS-PAGE	<i>sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>
TFP	tetrafluoro-phenyl
Tm	Température de fusion
UH	unité d'hémagglutination
UTR	région terminale non codante

I. Introduction :

1. Généralités sur la glycobiologie et les lectines:

La glycobiologie est la science qui étudie les structures, les propriétés et les fonctions des glucides dans la matière vivante. Ces études ne concernent pas seulement les glucides et les glycoconjugués mais également toutes les autres classes de molécules biologiques (protéines, lipides, nucléotides) qui interagissent avec eux. Le but de la glycobiologie structurale est de déterminer la structure tridimensionnelle et d'établir des modèles pour les interactions impliquant ces molécules (TAYLOR & DRICKAMER, 2003).

Les glucides représentent la plus importante source de biomasse sur la planète Terre et leur complexité, surtout en terme de variations structurales, n'a pas d'égale dans les autres classes de polymères biologiques (NELSON & COX, 2009). L'énorme potentiel biologique codé par les glucides résulte de leurs capacités à créer des structures de grande complexité, incluant des points de ramification, à partir des simples monomères. Cette caractéristique des glucides est unique parmi tous les polymères biologiques. La conséquence directe de cette variabilité structurale est une grande diversité dans le rôle et dans la localisation des glucides dans les organismes vivants: renforcement structural, accumulation d'énergie, reconnaissance moléculaire, croissance et développement, défense, ancrage et ciblage (TAYLOR & DRICKAMER, 2003).

Des découvertes majeures ont marqué l'élucidation du rôle des oligosaccharides: (i) la structure glucidique des déterminants du système ABO sanguin a été établie ; (ii) l'interaction entre une hémagglutinine et l'acide sialique a été démontrée lors de l'agglutination des érythrocytes par le virus de la grippe; (iii) il a été observé que la défucosylation des lymphocytes empêche leur migration dans les tissus. En 1968, Ashwell et Morell identifièrent la première lectine animale, une protéine impliquée dans la clairance des glycoprotéines circulant dans le sang (ASHWELL & MORELL, 1974). Cette découverte fût à l'origine des nombreuses

recherches qui ont depuis été consacrées au rôle joué par les glucides dans les processus de communication et de reconnaissance moléculaire.

Le mot lectine dérive du verbe latin *legere* qui veut dire « sélectionner » ou « choisir », un nom bien approprié pour cette très importante classe de protéines. Les lectines ont été définies comme des protéines d'origine non-immune qui se lient spécifiquement et de façon réversible aux sucres et ne montrent aucune activité enzymatique pour ces substrats (KOCOUREK & HOREJSI, 1981). Les lectines sont des molécules ubiquitaires, car elles se retrouvent dans toutes les classes d'organismes, chez les microorganismes (virus, bactéries), chez les plantes, chez les insectes et les animaux. Elles sont aussi appelées agglutinines car elles sont généralement capable d'agglutiner les cellules (comme les érythrocytes) et les glycoconjugués. Cette caractéristique très importante des lectines est due au fait que ces protéines sont généralement multivalentes (VAN DAMME *et al.*, 1998). La découverte majeure que certains états physiologiques et pathologiques étaient associés à un changement de l'état de glycosylation des cellules fut possible grâce à l'utilisation des lectines. Le nombre de travaux publiés sur les lectines a vu une grande croissance principalement grâce à l'abondance des lectines dans tous les organismes vivants, accompagné d'une certaine facilité de purification.

L'intérêt majeur qui pousse aujourd'hui la recherche sur les lectines est lié sans doute à leur capacité unique de « lire » l'information biologique qui est codifiée dans la structure tridimensionnelle des sucres. Les lectines sont en fait les récepteurs spécifiques pour les interactions protéine-sucre qui jouent des rôles clé dans une multitude de processus de reconnaissance moléculaire et de signalisation cellulaire (AMBROSI *et al.*, 2005) (Figure 1).

Il est intéressant de noter que la plupart des lectines sont spécifiques pour un petit nombre de sucres et que, dans la majorité des cas, ces sucres sont présents dans et sur la surface des cellules, surtout sous la forme de glycoconjugués. On peut identifier deux classes de lectines par rapport à leur spécificité : celles qui reconnaissent un monosaccharide spécifique et celles qui reconnaissent exclusivement des oligosaccharides (SHARON, 2003).

Parmi les monosaccharides le plus souvent reconnus par les lectines on retrouve le mannose, le fucose, le galactose/*N*-acetylgalactosamine (GalNAc) et la *N*-acetylglucosamine (GlcNAc). Très peu de lectines reconnaissent l'acide sialique sous la forme de monosaccharide. En général la spécificité est forte pour un monosaccharide et peu de substitutions sont autorisées. La deuxième classe de lectines est formée par celles qui nécessitent un oligosaccharide

spécifique. L'affinité montrée pour les oligosaccharides est généralement beaucoup plus élevée que pour les monosaccharides par la présence de sites de liaison plus profonds formant plus de contacts avec le ligand (SHARON, 2003).

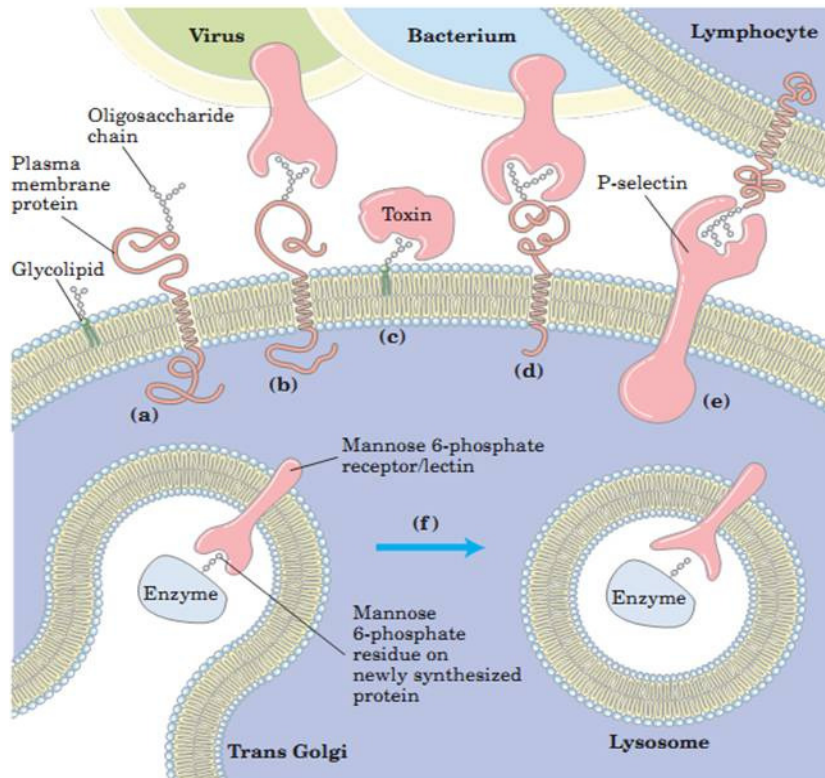


Figure 1. Rôle des oligosaccharides dans la reconnaissance et l'adhésion à la surface cellulaire. (a) Les oligosaccharides constitués d'un assemblage précis de monosaccharides (représentés par des hexagones) sont présentés par les glycoprotéines ou les glycolipides sur la surface externe de la membrane plasmique. (b) Certains virus qui infectent les cellules animales, tels que le virus de la grippe, se lient à des glycoprotéines de la surface cellulaire lors de la première étape de l'infection. (c) Des toxines bactériennes, comme celles impliquées dans le choléra et la coqueluche, se lient à des glycolipides avant de pénétrer dans les cellules. (d) Certaines bactéries comme *Helicobacter pylori* adhèrent à la cellule, puis colonisent ou infectent les cellules animales. (e) Les sélectines de la membrane plasmique de certaines cellules servent de médiateurs des interactions cellule-cellule, par exemple pour l'interaction entre les lymphocytes T avec les cellules endothéliales de la paroi des capillaires dans un site d'infection. (f) Le récepteur du mannose 6-phosphate (lectine du complexe de Golgi) se lie aux oligosaccharides des enzymes lysosomales, une étape nécessaire de leur transfert vers le lysosome. (Adapté de NELSON & COX, 2009).

2. Les lectines de Légumineuses:

Les lectines végétales sont des modèles de choix pour l'étude des bases moléculaires des événements de reconnaissance, car elles peuvent être facilement purifiées et montrent une grande variété de spécificité pour différents glucides (LORIS *et al*, 1998; LIS & SHARON,

1998, SHARON & LIS, 2002). Ces protéines ont été utilisés comme outils pour les études en immunologie et en glycobioologie, à cause de leur domaine de reconnaissance des glucides (CRD) qui interagit sélectivement avec certains résidus d'oligosaccharides (BONNEIL *et al*, 2004).

De nombreuses fonctions ont été proposées pour les lectines végétales, telles que la protection contre les pathogènes et les insectes, le transport et le stockage des glucides et la reconnaissance cellulaire (dans la cellule, entre les cellules ou entre organismes). Elles ont été également considérées comme des protéines de réserve ou comme des régulateurs de croissance (PUSTZAI, 1991) .

Chaque lectine a ses propres caractéristiques, notamment par rapport aux applications biologiques. Cela signifie que, en règle générale, chaque lectine a son application potentielle, ce qui explique pourquoi chacune d'entre elles, même si elle semble structurellement similaire à une autre lectine, mérite d'être étudiée séparément (PEUMANS & VAN DAMME, 1995). La découverte de nouvelles lectines ayant des spécificités différentes reste attrayante, car elles peuvent servir de nouveaux modèles pour la compréhension des processus dynamiques de la communication molécule-cellule, cellule-cellule et cellule-organisme.

Les lectines de légumineuses représentent le groupe le plus étudié parmi les lectines végétales (MANN *et al.*, 2001). Elles constituent une grande famille de protéines étroitement liées, présentes dans les espèces représentant la famille Leguminosae. Ces protéines peuvent être présentes dans les graines, feuilles, tiges et racines des végétaux (SHARON & LIS, 1990;. LORIS *et al*, 1998). Des domaine de la même famille existent dans de nombreuses autres plantes telles que le riz et *Arabidhopsis thaliana*, où ils sont généralement associés à des domaines kinases (Hervé et al., 1999).

Les structures cristallographiques d'un très grand nombre de lectines de légumineuses ont été résolues (210 structures disponibles de 35 lectines différentes) (VARROT *et al.*, 2011). Ces protéines s'assemblent sous différentes formes dimériques ou tétramériques mais la structure de tous les monomères de lectines de légumineuses est très conservée (Figure 2) (BANERJEE *et al* 1996;.. HAMELRYCK *et al*, 1998). Chaque sous-unité se compose d'environ 60 % des brins β antiparallèles disposés en deux ou trois feuillets reliés par des boucles. Ce repliement protéique, appartenant à la grande famille des β -sandwiches, est plus précisément connu sous le nom "jellyroll" ou "lectin fold " (SRINIVASAN *et al* 1996;. HAMELRYCK *et al*, 1998).

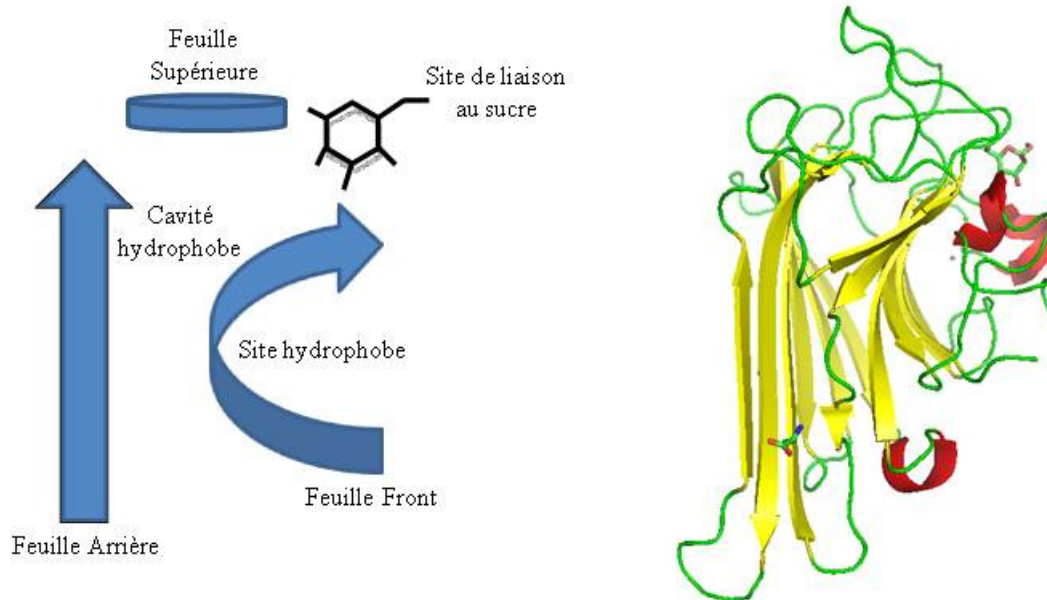


Figure 2. Schéma du principe de la structure tertiaire d'une monomère typique de lectine de légumineuse. À la droite, la structure tertiaire du monomère de la lectine de *Canavalia gladiata* (CGL) complexé avec α -méthylmannoside et Abu (l'acide α -aminobutyrique) (code PDB: 2D7F).

Le site de reconnaissance des glucides est constitué de plusieurs "boucles" présentant différents degrés de variabilité et formant une cavité sur la surface protéique. La reconnaissance spécifique des glucides se fait par des liaisons hydrogène entre la lectine et le sucre, des interactions de Van der Waals et l'établissement de contacts supplémentaires médiés par les molécules d'eau. (SHARMA & SUROLIA, 1997; WEIS & DRICKAMER, 1996; ELGAVISH & SHAANAN, 1998). Un ion calcium et un ion manganèse sont toujours présents dans le site de reconnaissance des lectines et les acides aminés assurant leur coordination sont très conservés. La présence de ces ions est nécessaire pour la reconnaissance des sucres. Le calcium et le manganèse n'interagissent pas directement avec le ligand glucidique mais ils maintiennent la conformation de certaines boucles du site, et en particulier la présence d'une liaison *cis* pour un aspartate très conservé (LORIS *et al* 1998;. BOUCKAERT *et al*, 2000). (BOUCKAERT *et al* 1995;. LORIS *et al*, 2004). Malgré le grand nombre de structures obtenues par cristallographie aux rayons X, relativement peu de lectines ont été résolues en complexe avec un oligosaccharide de grande taille (<http://www.cermav.cnrs.fr/lectines/>).

En plus du site de reconnaissance des glucides et du site de liaison aux métaux, des études récentes ont montré la capacité de ces protéines à interagir avec des composés

hydrophobes (HAMELRYCK *et al.*, 1998), qui peuvent être impliqués dans certains processus biologiques importants (DELATORRE *et al.*, 2007).

Les possibilités d'utilisation des lectines de légumineuses comme outils biotechnologiques sont considérables au vu du grand nombre d'études scientifiques démontrant les activités biologiques de ces molécules (KITADA *et al.*, 2010; KIMBLE *et al.*, 2010). Plusieurs activités biologiques ont été décrites pour les lectines extraites à partir d'espèces de légumineuses (Tableau 1).

Tableau 1. Exemples d'activités biologiques décrites pour les lectines extraites de légumineuses.

Lectine	Role	Référence
Dgui (<i>Dioclea guianensis</i>)	Activité antifongique	ARAUJO-FILHO, <i>et al.</i> , 2010
ConBr (<i>Canavalia brasiliensis</i>), CFL (<i>Cratylia floribunda</i>), Dgui, DGL (<i>D. grandiflora</i>), Dvir (<i>D. Virgata</i>)	Effet toxique sur les mollusques	SANTOS <i>et al.</i> , 2010
CGL (<i>Canavalia gladiata</i>)	Activité anti-inflammatoire et analgésique	NUNES <i>et al.</i> , 2009
ConBol (<i>Canavalia boliviana</i>)	Activité antinociceptive	FIGUEIREDO <i>et al.</i> , 2009
ConBr, CGL, ConM (<i>Canavalia maritima</i>)	Effet vasodilatateur	ASSREUY <i>et al.</i> , 2009
ConM	Relaxation de l'aorte et libération d'oxyde nitrique	GADELHA <i>et al.</i> , 2005
ConBr, DGL, DVL (<i>Dioclea violacea</i>)	Activation des lymphocytes	BARBOSA <i>et al.</i> , 2001
ConBr	Activité antidépresseur	BARAUNA <i>et al.</i> , 2006
ConA (<i>Canavalia ensiformes</i>), ConBr, CFL, DVL, DGL	Interférence dans le processus de formation des biofilms microbiens	TEIXEIRA <i>et al.</i> , 2006

La sous-tribu Dalbergiae, appartenant à la sous-famille Papilionacées de la famille Léguminosae, comprend 48 genres et 1200 espèces. La présence de lectines avec des spécificités variées a été rapportée dans 11 espèces de cette sous-tribu (Tableau 2).

Tableau 2. Spécificité pour les sucres des lectines de la sous-tribu Dalbergieae.

Lectine	Spécificité	Référence
<i>Andira fraxinifolia</i>	Man	RANGEL <i>et al.</i> , 2009
<i>Andira retusa</i>	Man	NASCIMENTO <i>et al.</i> , 2002
<i>Lonchocarpus araripensis</i>	GlcNAc	PIRES <i>et al.</i> , 2008
<i>Lonchocarpus carpassa</i>	Gal/GalNAc/Glc	JOUBERT <i>et al.</i> , 1986
<i>Lonchocarpus sericeus</i>	GlcNAc	ALENCAR <i>et al.</i> , 1999
<i>Machaerium acutifolium</i>	GlcNAc /Glc	BEZERRA <i>et al.</i> , 2003
<i>Platymiscium floribundum</i>	Man/ GlcNAc	PEREIRA-JÚNIOR, 2008
<i>Platypodium elegans</i>	Man/Glc	BENEVIDES <i>et al.</i> , 2008
<i>Pterocarpus angolensis</i>	Man/Glc	LORIS <i>et al.</i> , 2003
<i>Vatairea macrocarpa</i>	Gal/GalNAc	CAVADA <i>et al.</i> , 1998

Parmi les lectines de la sous-tribu Dalbergiae, seule la lectine à glucose/mannose de *Pterocarpus angolensis* (PAL) a été décrite par détermination de sa structure tridimensionnelle (LORIS *et al.*, 2003). Cette structure contient une boucle classique mannose spécifique, mais la boucle de reconnaissance des métaux a plus de similarité avec des lectines de spécificité différentes telles que celles de *Ulex europaeus* et *Maackia amurensis*. Les interactions avec la molécule de mannose dans le site principal de liaison sont conservées, mais les environs du site primaire diffèrent de ceux des autres lectines Glc/Man spécifiques telles que ConA. Ces observations expliquent les différences dans leur profil de spécificité fine par l'oligomanoses. (LORIS *et al.*, 2004). Les structures cristallines de PAL en complexe avec une série d'oligomanoses et de *N*-glycannes complexes ont été résolues (LORIS *et al.*, 2003, BUTS *et al.*, 2006, GARCIA-PINO *et al.*, 2007). Les oligomanoses se lient de façon préférentielle par la séquence tétrasaccharide au mannose de la branche 1–6. Les expériences de microcalorimétrie

isotherme de titration ont montré des affinités de 150 μM pour l'oligomannose M5, 16 μM pour le biantenné complexe GN2M3 et 12 μM pour la glycoprotéine agglutinine de soja.

Les études structurales des lectines de la sous-tribu Dalbergiae sont donc disponibles pour une seule lectine ce qui limite l'étude des mécanismes d'action basé sur l'affinité différentielle entre ligands spécifiques.

Les travaux préliminaires effectués lors de mon Master de Biochimie à l'Université Fédéral du Ceará (BENEVIDES *et al.*, 2008) ont mis en évidence une activité lectinique dans les extraits de graines de *Platypodium elegans*. Ces extraits sont capables d'agglutiner les érythrocytes de lapin, mais pas des érythrocytes du système ABO. Le potentiel fongicide et termiticide d'une fraction lectinique a également été décrit mais la présence d'isoformes et la difficulté à les séparer ont limité l'étude de cette lectine.

3. Clonage et expression hétérologue des lectines de légumineuses.

Le clonage des gènes codant pour les lectines végétales peut être considéré comme une étape cruciale pour l'obtention et la production de ces protéines à grande échelle (LUO *et al.*, 2005). L'expression des lectines végétales dans les bactéries se révèle être un moyen efficace de sélection pour l'étude isolée de chaque isoforme (LUO *et al.*, 2005). Le clonage et l'expression des lectines végétales recombinantes ont été utilisés pour différents buts: détermination de leurs structures primaires, analyse génétique et phylogénétique, évaluation du rôle des différents acides aminés dans la reconnaissance des glucides et l'association des monomères, production de lectines présentant de nouvelles spécificités, étude de la fonction des lectines dans les organismes végétaux (STREICHER & SHARON, 2003).

Escherichia coli est l'organisme le plus souvent utilisé pour la production des protéines recombinantes en raison de sa croissance rapide, sa génétique et sa protéomique bien caractérisée et la disponibilité des nombreux vecteurs de clonage (LUTOMSKI *et al.*, 2004). Dans le cas des lectines de légumineuses recombinantes, *E. coli* est également le système hétérologue le plus largement utilisé (BRANCO *et al.*, 2004; WANG *et al.*, 2004). Le clonage de ces gènes est facilité par la disponibilité de kits commerciaux. Ces dernières années, plusieurs lectines de légumineuses ont été clonées et exprimées dans des systèmes hétérologues *E. coli* (Tableau 3).

Tableau 3. Lectines de Legumineuses clonées et exprimées dans *E. coli*.

Espèce	Nom	Spécificité	Rendement	Référence
<i>Allium sativum</i>	ASAL	Man	5 mg/L	UPADHYAY <i>et al.</i> 2010
<i>Arachis hypogea</i>	PNA	GalNAc/Gal	12 µg/L	SHARMA & SUROLIA, 1994
<i>Artocarpus incisa</i>	Frutaline	Gal	16 mg/L	OLIVEIRA <i>et al.</i> 2009
<i>Bauhinia purpurea</i>	BPA	Gal	ND*	KUSUI <i>et al.</i> , 1991
<i>Canavalia brasiliensis</i>	Prepro ConBr	Man/Glc	30 mg/L	NOGUEIRA <i>et al.</i> , 2002
<i>C. ensiformis</i>	Prepro ConA	Man/Glc	1 mg/L	YAMAUCHI & MINAMIKAWA, 1990
<i>Castanea crenata</i>	CCA	Man/Glc	4,2 mg/L	NAKAMURA <i>et al.</i> , 2002
<i>Cratylia mollis</i>	CRAMOLL 1	Man/Glc	6 mg/L	VAREJÃO <i>et al.</i> , 2010
<i>Cycas revoluta</i>	CRLI	Man	18 mg/L	HARAGUCHI <i>et al.</i> , 2006
<i>Dolichus biflorus</i>	DBL	GalNAc	20 mg/L	CHAO <i>et al.</i> , 1994
<i>Erythrina crista-galli</i>	ECL	GalNAc/Gal	294 mg/L	STANCOMBE <i>et al.</i> , 2003
<i>E. corallodendron</i>	EcorL	GalNAc/Gal	4,7 mg/L	ARANGO <i>et al.</i> , 1992
<i>Glycine max</i>	SBA	GalNAc/Gal	100 µg/L	ADAR <i>et al.</i> , 1997
<i>Griffonia simplicifolia</i>	GS-II	Man/Glc	ND	ZHU <i>et al.</i> , 1996
<i>Maackia amurensis</i>	MAH	Neu5Ac2-3Gal1-3 (Neu5Ac2-6) GalNAc	ND	YIM <i>et al.</i> , 2001
<i>Phaseolus vulgaris</i>	PHA-L	Gal 1-4 GlcNAc 1-6(Gal1-4 GlcNAc1-2)	0,4-1 mg/L	HOFFMAN & DONALDSON, 1987

<i>P. lunatus</i>	LBL	GalNAc/Gal	5 mg/L	JORDAN & GOLDSTEIN, 1994
<i>Pisum sativum</i>	PSL	Man/Glc	2-5 mg/L 20 mg/L	STUBBS <i>et al.</i> , 1986; PRASTHOFER <i>et al.</i> , 1989
<i>Robinia pseudoacacia</i>	RBL	GalNAc/Gal	10 mg/L	NISHIGUSHI <i>et al.</i> , 1997

*ND : non déterminé

On peut également noter qu'un petit nombre de lectines des plantes ont été produites avec succès dans la levure *Pichia pastoris*, comme PHA (*Phaseolus vulgaris* agglutinin), GNA (*Galanthus nivalis* agglutinin) (RAEMAEKERS *et al.*, 1999) et ConBr (*Canavalia brasiliensis*) (BEZERRA *et al.*, 2006).

4. Étude structurale et thermodynamique de l'interaction lectine-sucre:

Une compréhension approfondie au niveau moléculaire du mécanisme de l'association protéine-sucre exige l'élucidation des aspects structuraux et dynamiques du processus. Une grande quantité de données structurales sont disponibles (3D-Lectine : <http://www.cermav.cnrs.fr/lectines/>), mais la corrélation entre les propriétés structurales et dynamiques n'est pas suffisamment éclaircie pour permettre une description de l'énergie d'interaction à partir des structures (TOONE *et al.*, 1994). La microcalorimétrie de titration fournit une méthode directe d'évaluation de la thermodynamique de l'interaction et ces données peuvent compléter l'information structurale (AMBROSI *et al.*, 2005).

Les interactions lectine-sucre sont généralement caractérisées par une faible affinité pour les ligands monovalents, une caractéristique contrebalancée par la multivalence des lectines pour les glycoconjugués présents en grande quantité à la surface des cellules. En général, une affinité millimolaire est observée pour les lectines en interaction avec des monosaccharides. Pour des oligosaccharides plus longs, correspondant à un site étendu de liaison sur la surface de la lectine, l'affinité augmente à un niveau micromolaire. La microcalorimétrie de titration est couramment utilisée pour caractériser la thermodynamique des associations protéine-sucre (AMBROSI *et al*

2005; DAM & BREWER, 2002). En général, ces associations sont caractérisées par un terme enthalpique favorable, en raison du nombre élevé de liaisons hydrogène et aux interactions de van der Waals, contrebalancé par une contribution entropique défavorable, qui a été attribuée au réarrangement des molécules de solvant et à la perte de flexibilité conformationnelle/mobilité du ligand (DAM & BREWER, 2002).

Dernièrement, il a été montré qu'une meilleure compréhension des interactions lectine-glucide est essentielle pour le développement d'applications biomédicales (JANAKIRAMAN *et al.*, 1994, MAMMEN *et al.*, 1998, MOLEMA & MEIJER, 1994). La connaissance profonde des aspects structuraux et dynamiques du processus est nécessaire et l'échec d'un certain nombre de médicaments basés sur les sucres peut être attribué à une pauvre compréhension de leurs mécanismes d'action présumés. (JANAKIRAMAN *et al.*, 1994 A l'opposé une bonne maîtrise des interactions a permis d'obtenir des résultats impressionnants. Un glucide multivalent a par exemple été conçu pour se lier au pentamère de lectines de la toxine de Shiga et une affinité subnanomolaire a été obtenue (KITOV *et al.*, 2000). De la même manière, un ligand pentavalent d'une entérotoxine de *E. coli* a montré une forte affinité avec un gain de 1000 par rapport au ligand monovalent (FAN *et al.*, 2000).

5. Objectifs du travail de thèse:

- Clonage et caractérisation moléculaire du gène de la lectine des graines de *Platypodium elegans.*;
- Développer des techniques d'expression chez *E. coli* et de purification de la lectine, avec l'obtention finale de la protéine recombinante (PELa) pure;
- Évaluer l'affinité et la spécificité de la lectine pour différents oligosaccharides par microcalorimétrie de titration isotherme (ITC) et *Glycan Array*.
- Cristalliser et déterminer la structure 3-D, en obtenant la première structure cristalline d'une lectine recombinante de la sous-tribu Dalbergiae complexée avec des oligosaccharides. Cela

fournira des nouvelles structures des lectines de la famille de légumineuses, dont la sous-tribu Dalbergiae dispose actuellement d'une seule structure décrite.

- Déterminer les similitudes et les différences dans la spécificité, la structure et les affinités entre PELa et les autres lectines de la même famille.

L'identification de la lectine présente dans les graines de *P. elegans* et sa caractérisation biochimique fourniront d'importantes informations structurales sur les interactions protéine-sucre nécessaires pour comprendre les mécanismes d'action, la reconnaissance, l'affinité et la spécificité des oligosaccharides complexes au niveau moléculaire. À long terme, cette lectine produite sous forme recombinante pourra être testée pour sa capacité à reconnaître les glycanes impliqués dans des événements biologiques intéressants. Ces études pourront être corrélées à la recherche de sa fonction dans l'organisme d'origine ou au développement d'applications biotechnologiques dans les domaines tels que la pharmacologie et l'immunologie.

II. Matériels et Méthodes :

1. Biologie Moléculaire: Clonage du gène de PELa

En l'absence d'information sur la séquence primaire de la lectine de *Platypodium elegans* (PELa), une stratégie a été mise en place pour l'obtention et le clonage de son gène à partir des ARNm extraits des graines. La production d'ADNc a été réalisée par des techniques telles que 3'-RACE et 5'-RACE, comme décrit ci-dessous. Ensuite, le gène a été cloné dans le vecteur d'expression pET-29a afin de permettre l'expression de la protéine PELa dans *Escherichia coli*.

1.1. Purification de l'ARN

Pour isoler l'ARN, 1 g de graines ont été congelées dans l'azote liquide puis réduites en poudre dans un mortier. L'ARN cellulaire total a été isolé en utilisant le kit *Concert Plant RNA reagent* (Invitrogen), selon le protocole du fabricant. La quantité des ARN totaux extraits a été déterminée par lecture au spectrophotomètre à 260 nm et 280 nm. Leur pureté et leur intégrité ont été évalués par électrophorèse en gel d'agarose 1 %.

1.2. 3'-RACE

Deux amorces dégénérées pour le gène de PELa ont été conçues avec une dégénérescence minimale (PLAT1 et PLAT2) à partir des séquences primaires de deux régions conservées de cette famille de lectines de légumineuses qui correspondent aux sites de liaison aux métaux (Tableau 4).

Tableau 4. Conception d'amorces dégénérées pour le gène de PELa.

Séquence d'acides aminés	Amorce	Séquence des amorces*	Température de fusion (T _m)
(V/L/I)AVEFDT	PLAT1	5'-TNGCNGTNGART TYGAYAC-3'	54,5
HIGIDVN	PLAT2	5'-CAYATHGGNATHGA YGTNAA-3'	52,8

* R, Y et N indiquent (A ou G), (C ou T), et (A, C, G ou T), respectivement.

A partir des ARN totaux, une transcription inverse a été effectuée en utilisant la transcriptase réverse (SuperscriptTM II) et l'amorce oligo(dT) contenant en son extrémité 5' un site supplémentaire (5'-GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTTTT-3'), encore appelée AP (*adapter primer*). Ainsi, après la transcription inverse de la population d'ARNm, la transcription du gène spécifique de PELa a été obtenue par PCR en utilisant une amorce dégénérée pour la région interne du gène (PLAT1 ou PLAT2) et l'amorce AP, conformément aux recommandations du fabricant 3'-RACE (Invitrogen). Le produit 3'-RACE a ensuite été purifié, cloné dans le vecteur pGEM-T easy (Promega) et envoyé pour séquençage.

1.3. 5'-RACE

Cette procédure a été réalisée en quatre étapes principales: la conception de l'amorce de la région 3' du gène, la synthèse et la purification de l'ADNc, l'ajout de l'étiquette poly d(C), et l'amplification du gène, selon le protocole du fabricant 5'-RACE (Invitrogen).

L'amorce dite « gene-spécifique » (PLAT3) pour l'expérience 5'-RACE a été conçue à partir de la séquence de l'extrémité 3' de l'ADNc (obtenue par l'approche 3'-RACE). Cette amorce a été choisie afin d'obtenir le gène complet jusqu'à son extrémité 5' (Tableau 5).

La stratégie générale de la technique 5'-RACE a consisté en la synthèse de l'ADNc d'une population de l'ARN total en utilisant une amorce antisens spécifique du gène (PLAT3) et la SuperscriptTM II. Après purification des ADNc sur colonne SNAP, l'enzyme TdT (*Terminal deoxynucleotidyl transferase*) a été utilisée pour ajouter une étiquette homopolymère poly d(C) à l'extrémité 3' de l'ADNc purifié. La combinaison d'une amorce adaptateur complémentaire à

l'homopolymère (5'-RACE anchor primer) et de l'amorce spécifique du gène de PELa (PLAT3) a permis l'amplification par PCR de la séquence inconnue à l'extrémité 5' de l'ARNm.

Tableau 5. Conception de l'amorce spécifique du gène pour le 5'-RACE à partir de la séquence nucléotidique du produit 3'RACE.

Séquence de référence	Séquence des nucléotides	Amorce	Séquence de l'amorce	Témpérature de fusion (Tm)
Extrémité 3'RACE	5'-TCCAGAGGGA TCTCTATATCA-3'	PLAT3	5'-TGATATAGAGATC CCTCTGGA-3'	55,9

Ensuite, le produit 5'-RACE a été purifié, cloné dans le vecteur pGEM-T easy (Promega) et envoyé pour séquençage.

1.4. Clonage dans le vecteur pET-29a:

La région codante du gène PELa a été clonée dans le vecteur pET-29a, à partir du produit 5'-RACE obtenu (Figure 3). Pour cela, deux amorces (PLAT4 et PLAT5) ont été définies, possédant en 5' les sites de clonage *NdeI* et *HindIII* pour permettre l'insertion orientée dans le vecteur (Tableau 6).

Tableau 6. Amorces utilisées dans le clonage du gène PELa dans le vecteur pET-29a.

Amorce	5'- Séquence – 3'*	Séquence d'acides aminés correspondante
PLAT4 (sens)	GGAATTCC <u>CATATG</u> CTACTGAATAGGGCAAAC	(<i>NdeI</i>)MLLNRA
PLAT5 (antisens)	GGCA <u>AAGCTT</u> TTCACATATCACGTGCAAGATAC	LARDM Stop (<i>HindIII</i>)

*Les séquences soulignées indiquent les positions des sites *NdeI* et *HindIII*, respectivement.

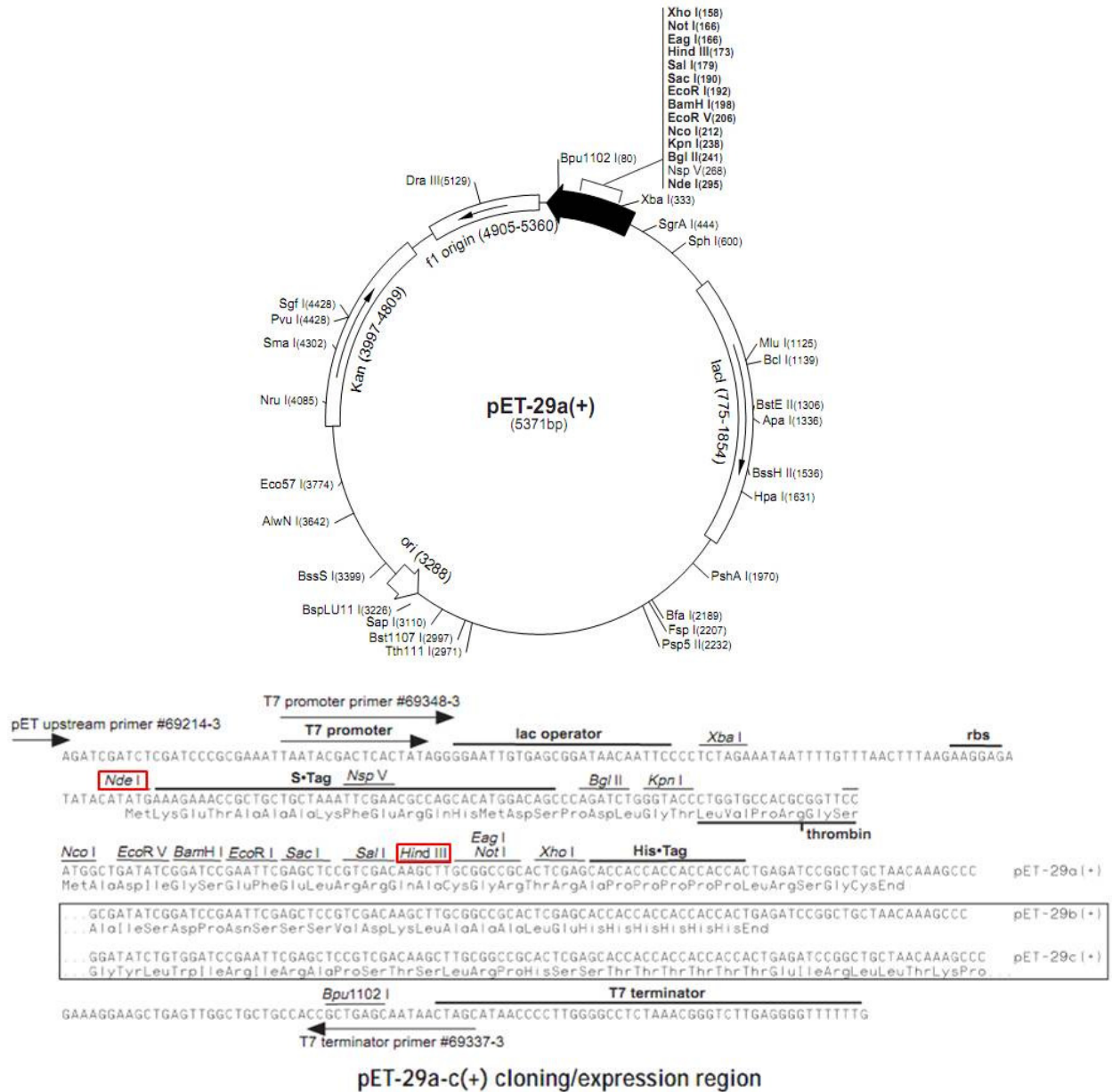


Figure 3. Carte du vecteur pET-29a (Novagen). Les sites pour les enzymes de restriction *NdeI* et *HindIII* sont indiqués en rouge.

Le produit 5'-RACE a donc été amplifié par PCR en utilisant les amorces PLAT4 et PLAT5. Le produit PCR obtenu a été purifié et digéré par les enzymes *NdeI* et *HindIII* tout comme le vecteur pET-29a vide. Après purification, le vecteur et l'insert digérés ont été ligués. Le plasmide résultant (pET-29a_PELa) a été utilisé pour transformer des cellules compétentes de *E. coli* BL21.

2. Expression et purification de PELa recombinante:

2.1. Expression de PELa dans *E. coli*

Les cellules de *E. coli* BL21 contenant le plasmide pET-29a_rPELa ont été cultivées dans un milieu LB contenant l'antibiotique kanamycine (30 µg/mL) à 37°C. Lorsque la culture atteint une DO_{600nm} de 0,6, l'expression des protéines est induite par l'ajout de l'IPTG (isopropyl-β-D thiogalactoside) à une concentration finale de 0,5 mM pendant 4 heures à 30 ° C.

Les cellules ont été centrifugées et le culot bactérien a été suspendu dans le tampon de lyse 50 mM Tris-HCl pH 7,5, 150 mM NaCl, 5 mM CaCl₂, 5 mM MnCl₂ (5 mL pour 1g de culot cellulaire). Les cellules ont été lysées par à une forte pression en utilisant le *one shot cell disrupter* (Constant Systems) à 1700 bars. Après centrifugation à 20000xg pendant 30 min à 4° C, le culot a été lavé deux fois avec 50 mM de Tris-Cl pH 8, 4 M urée, 0,5 M NaCl, 1 mM EDTA, 0,5% de Triton X-100 (5 mL de tampon pour 1g de culot bactérien). Les corps d'inclusion du culot final ont été dénaturés et solubilisés dans 50 mM de Tris-HCl pH 8, 8 M urée, 10 mM DTT (2 mL du tampon pour 1g de culot bactérien). Après une centrifugation à 20000xg pendant 30 min à 4 ° C, la concentration de la solution de protéine dénaturée a été ajustée à 1 mg/mL.

2.2. Renaturation et purification

La solution de protéine dénaturée (1mg/mL) a été progressivement diluée 20 fois dans 50 mM Tris-HCl pH 7,5, 240 mM NaCl, 2 mM MnCl₂, 2 mM de CaCl₂, 10 mM KCl, 1 mM de DTT. Après 1 heure à 10 °C suivie d'une filtration (papier filtre), les protéines renaturées ont été soumises à l'étape de purification.

La purification de PELa a été effectuée sur une colonne de D-mannose-Sepharose (Sigma-Aldrich). PELa a été déposée sur la colonne dans le tampon d'équilibration (50 mM Tris-HCl pH 7,5, NaCl 150 mM, CaCl₂ 5 mM, 5 mM MnCl₂). Après un lavage avec le tampon d'équilibration, PELa a été éluée par le même tampon additionné de 200 mM de D-mannose. La

protéine purifiée a été concentrée à l'aide d'un dispositif d'ultrafiltration Vivaspin (coupure de 10 kDa; GE Healthcare) et conservée à 4 °C.

Les fractions des protéines obtenues ont été analysées par SDS-PAGE 12% (LAEMMLI, 1970) avec coloration au bleu de Coomassie (Sigma-Aldrich).

3. Outils pour l'étude des interactions protéine-sucre

3.1. Tests d'hémagglutination

L'activité hémagglutinante de PELa a été mesurée dans des plaques de microtitration 96 puits avec fond en V (Nunc). La protéine a été préparée dans un tampon 50 mM Tris-HCl pH 7,5 NaCl 150 mM CaCl₂ 5 mM 5 mM MnCl₂ à une concentration de 3 mg/mL. Dans une première étape, 25 µL de tampon seul ont été déposés dans chaque puits. Ensuite, un volume de 25 µL de protéine a été déposé dans le premier puits, puis une gamme de concentration a été réalisée par dilution en cascade de deux en deux dans les puits suivants. Les hématies de lapin (BioMérieux) ont été suspendues dans une solution saline de 150 mM NaCl à une concentration finale à 2 %. Un volume de 25 µL de ces érythrocytes a été ajouté aux 25 µL de protéine dans chaque puits. La lecture d'activité hémagglutinante a été réalisée après 30 minutes à 37 °C. La plus basse concentration en lectine pour laquelle une hémagglutination est encore visible est appelée point d'équivalence et la concentration en lectine de ce puits correspond à une unité d'hémagglutination (UH).

3.2. *Glycan Array* (puce à sucres)

La lectine PELa a été marquée par le tétrafluoro-phenyl ester grâce au réactif Alexa-fluor 488-TFP (Invitrogen). La protéine a été préparée à une concentration de 1mg/mL dans un tampon bicarbonate de sodium pH 9,0. Le réactif TFP a été préparé extemporanément dans du DMF (diméthylformamide) à une concentration de 10 mg/mL. Le mélange a été incubé une heure à température ambiante sous agitation. La protéine a été finalement séparée du réactif

résiduel par colonne de dessalage sur colonne D-salt polyacrylamide (Pierce) dans un tampon 50 mM Tris-HCl pH 7,5, NaCl 150 mM, CaCl₂ 5 mM, 5 mM MnCl₂. L'échantillon marqué a été ensuite envoyé au *Consortium for Functional Glycomics* (<http://www.functionalglycomics.org/>) pour être testé sur les glycan arrays.

3.3. Microcalorimétrie isotherme de titration (ITC)

L'affinité entre PELa et les oligosaccharides définis à partir des résultats de *Glycan array* a été mesurée par microcalorimétrie de titration en injectant des aliquots du ligand dans la cellule contenant la protéine et en mesurant la chaleur dégagée.

Les solutions de protéine et d'oligosaccharides ont été préparées dans le même tampon : 50 mM Tris-HCl pH 7,5, NaCl 150 mM, CaCl₂ 5 mM, 5 mM MnCl₂ de manière à avoir une concentration en sucre 10 fois supérieure à celle de la protéine. Des aliquots de 10 µL de sucre ont été injectés toutes les 5 min dans la cellule de mesure contenant la protéine thermostatée à 25 °C. Chaque expérience a comporté une série de 30 injections.

La courbe intégrée de la chaleur dégagée a été analysée par régression non linéaire en utilisant un modèle à un ou deux sites de liaison (origine 7.0). La courbe théorique obtenue à partir des données expérimentales donne accès à la constante d'association (K_a), l'enthalpie de liaison (ΔH) et la stochiométrie (n). Les autres paramètres thermodynamiques tels que les variations d'énergie libre (ΔG) et d'entropie d'interaction (ΔS) ont été calculées à partir de l'équation $\Delta G = \Delta H - T\Delta S = -RTLnK_a$, où T est la température absolue et R est la constante molaire des gaz (8,314 J.mol⁻¹.K⁻¹).

4. Cristallographie aux rayons X:

4.1. Cristallisation et collecte des données

La méthode de cristallisation sélectionnée a été la diffusion de vapeur en goutte pendante. Une goutte de solution concentrée en protéine (1µL) a été déposée sur une lamelle de verre

siliconée et diluée au demi par une solution de précipitant (1 μL). La lamelle a été ensuite retournée et scellée avec de la graisse sur un puits contenant ce même agent précipitant (500 μL).

Les expériences de co-cristallisation de PELa avec le ligand trisaccharidique et heptasaccharidique ont été effectuées en préparant des solutions de protéine à 10 mg/mL et des solutions de ligand à 1 mM. Le criblage des conditions de cristallisation a été réalisé avec les kits *Hampton Research Screens I and II* (Hampton Research, Riverside, CA, USA).

Les cristaux obtenus ont été congelés dans l'azote liquide (100K) après trempage pendant une durée aussi courte que possible dans 25 ou 30% (v/v) de glycérol dans la solution de précipitation. Les données de diffraction ont été collectées à l'ESRF (*European Synchrotron Radiation Facility*), Grenoble, France, sur les stations ID14-2 et BM30A en utilisant un détecteur CCD (charge-coupled device) ADSC Q4R (Quantum Corp.)

Un cristal de PELa cocrystallisé avec le trimannose a été utilisé pour obtenir un jeu de données de diffraction collectées à 2,1 Å de résolution. Un ensemble de 140 images (oscillation de 0,75 °) a été enregistré. Les données de diffraction ont été indexées, intégrées et mises à l'échelle en utilisant MOSFLM (LESLIE, 1992) et SCALA (EVANS, 2006), et converties en facteurs de structure en utilisant la suite du programme CCP4 (COLLABORATIVE COMPUTATIONAL PROJECT, 1994). Le cristal appartient au groupe d'espace orthorhombique $p2_12_12_1$ et le coefficient de Matthews de 2,19 Å³.Da⁻¹ a indiqué la présence de deux monomères dans l'unité asymétrique (MATTHEWS, 1968).

Un cristal de PELa cocrystallisé avec l'heptasaccharide a permis de collecter une jeu de données à 1,6 Å de résolution. Un ensemble de 91 images (oscillation de 1 °) a été enregistré. Les données ont été traitées comme indiqué ci-dessus. Le cristal appartient au groupe d'espace orthorhombique $p2_12_12_1$ et le coefficient de Matthews de 2,29 Å³.Da⁻¹ a indiqué la présence de deux monomères dans l'unité asymétrique (MATTHEWS, 1968).

4.2. Résolution et affinement des structures cristallographiques

La structure du complexe PELa/trimannose a été résolue par remplacement moléculaire en utilisant un monomère de lectine *Pterocarpus angolensis* (PDB : 1N3P) comme modèle pour le programme PHASER (McCOY *et al.*, 2004). L'affinement de la structure a été effectué avec

le programme REFMAC (MURSHUDOV, 1997) et la construction du modèle avec le programme COOT (EMSLEY et COWTAN, 2004). La structure du complexe PELa/heptasaccharide a été résolue en utilisant cette première structure de PELa comme modèle. Les coordonnées ont été déposées à la *Protein Data Bank* sous les codes 3ZVX et 3ZYR. Les dessins moléculaires ont été préparés en utilisant PyMOL (DELANO, 2002) et LigPlot (WALLACE *et al.*, 1995).

III. Résultats et discussion :

1. Clonage, expression, renaturation et purification de la lectine PELa:

L'ARN total extrait d'un gramme de graines matures de *P. elegans* par le kit *Plant RNA Reagent* (Invitrogen) été obtenu avec un rendement de 238,6 ng/mL, et un profil électrophorétique montrant un faible degré de dégradation. La Figure 4 montre la structure de l'ARNm codant de la lectine PELa.

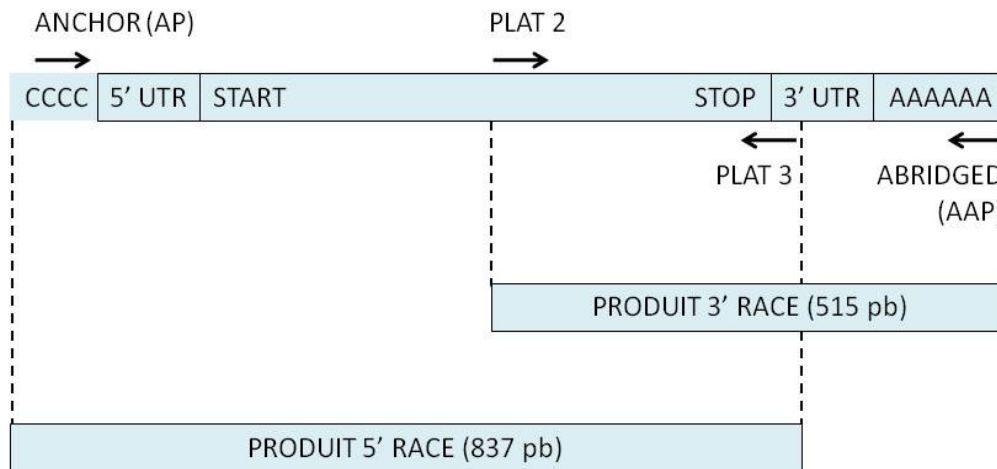


Figure 4. Structure complète de l'ARNm codant pour la lectine de *P. elegans* (région codante : 786 nucléotides) dérivés à partir des produits de 3' et 5' RACE. UTR: région terminal non codante.

Le 3' RACE a été réalisé en utilisant les amorces PLAT2 et AP, mais la bande obtenue (environ 500 pb) était très faible. Ce fragment a été réamplifié avant d'être cloné pour le séquençage. L'analyse des séquences des 500 paires de bases de l'extrémité du produit 3' RACE a montré une région codante de 321 pb, terminée par un codon *stop*, une séquence de 165 pb de la région 3' non codante et d'une queue poly-(A).

Le 5' RACE a été effectué en utilisant des fragments d'ADNc. Le produit amplifié par les amorces PLAT3 et AAP a donné une séquence de 800 pb de l'ARNm de PELa, avec un codon d'initiation et une séquence de 15 pb de la région 5'-non codante.

La séquence d'acides aminés a été déduite du produit 5' RACE. L'analyse de la séquence nucléotidique indique que la protéine a 261 acides aminés. La séquence protéique présente des similarités avec d'autres lectines de légumineuses spécifiques pour le Glc/Man telles que *Pterocarpus angolensis* (72%), *Pterocarpus rotundifolius* (82%) et *Arachis hypogaea* (70%) (numéros d'identification: CAD19804, AAT57665 et AAA74572, respectivement). La Figure 5 montre l'alignement de la séquence de PELa avec celles d'autres lectines de légumineuses de structure cristalline connue.

Le gène de la lectine PELa a été cloné dans un vecteur pET-29a pour produire la protéine recombinante dans *E. coli* BL21. PELa a été exprimée et accumulée dans des corps d'inclusion et un protocole de renaturation a été utilisé pour obtenir la protéine soluble et fonctionnelle. Les corps d'inclusion ont été lavés dans 4 M d'urée et ont été solubilisés dans 8 M l'urée puis ont été renaturés par dilution lente. La protéine PELa purifiée et active a été obtenue par passage sur une colonne de mannose-agarose et élution par 200 mM de D-mannose. L'analyse par SDS-PAGE indique une masse moléculaire de 29 kDa (Figure 6), correspondant bien à la MM théorique calculée d'après la séquence primaire. 5 mg de PELa sont obtenus à partir d'un litre de culture cellulaire. La protéine a été caractérisée par son activité d'hémagglutination avant chaque analyse réalisée afin de confirmer la fonctionnalité de la lectine.

2. Caractérisation de la spécificité et de l'affinité de la lectine PELa:

2.1. Étude des propriétés d'hémagglutination de PELa:

Afin de tester l'activité lectinique de PELa recombinante, des expériences d'hémagglutination sur des érythrocytes de lapin ont été réalisées (Figure 7). A partir d'une concentration de départ de 1,5 mg/mL, l'unité d'hémagglutination (UH) obtenue est de 23 µg/mL. Les résultats montrent que cette protéine possède des propriétés d'hémagglutination, ce qui confirme le repliement correct et l'activité de PELa.

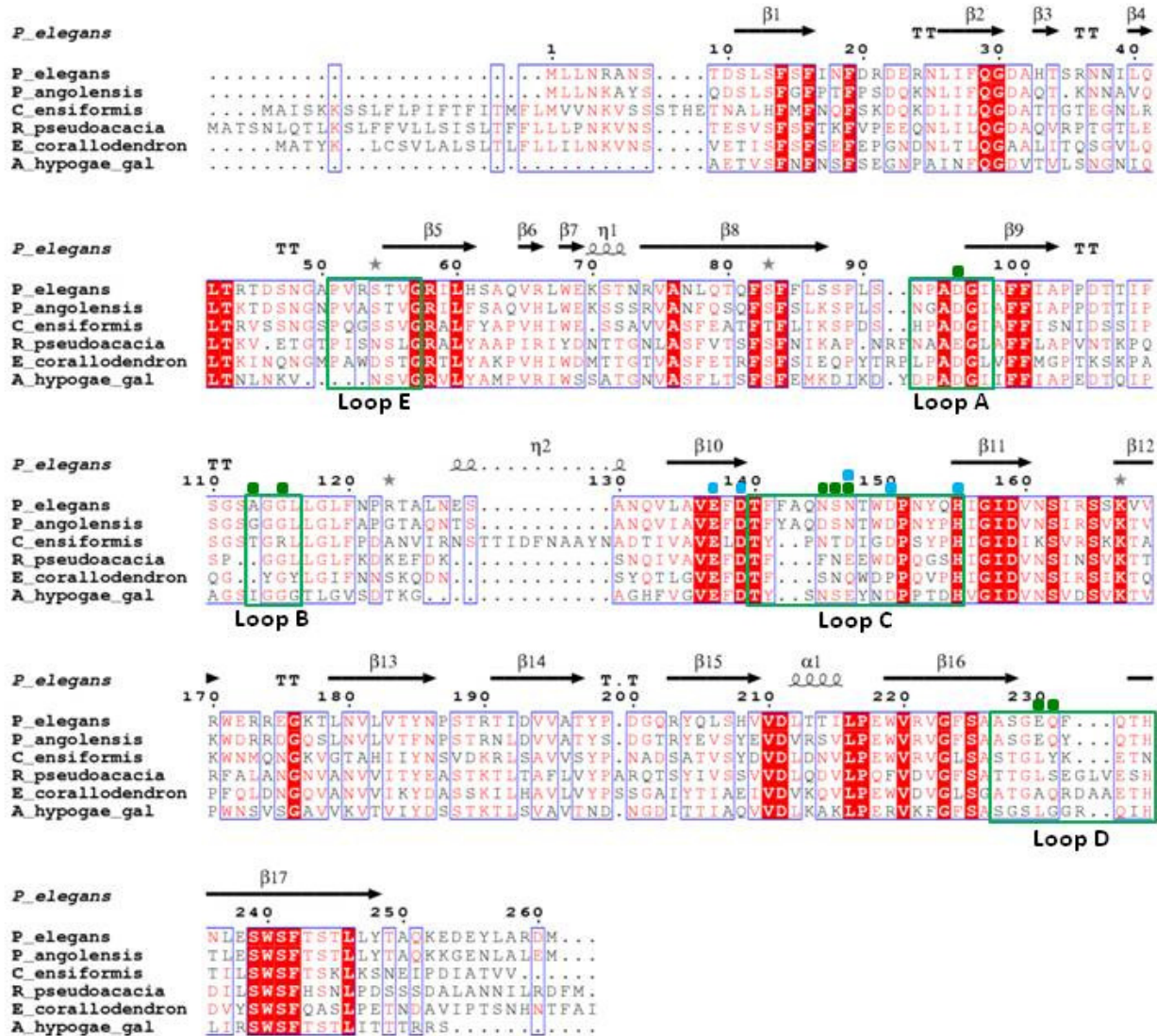


Figure 5. Alignment des séquences primaires des lectines de légumineuses. L'alignement multiple et la prédiction des structures secondaires ont été obtenus en utilisant le programme webESPrIPT (GOUET, *et al.*, 1999). Les séquences ont été obtenues à partir de la base de données *Entrez protein sequence database*. Numéros d'identification: AEK69351 (*Platypodium elegans* lectin), CAD19804 (*Pterocarpus angolensis* lectin), 2PEL_A (*Arachis hypogaea* lectin), CAA25787 (*Canavalia ensiformis* lectin), BAA36415 (*Robinia pseudoacacia* lectin) et CAA36986 (*Erythrina collarodendrom* lectin). Symboles: hélice α (α), hélice η (η), brin- β (β), coude- α (TTT), coude- β (TT), acides aminés du site de reconnaissance du monosaccharide (carré vert), acides aminés de liaison des métaux (carré bleu).

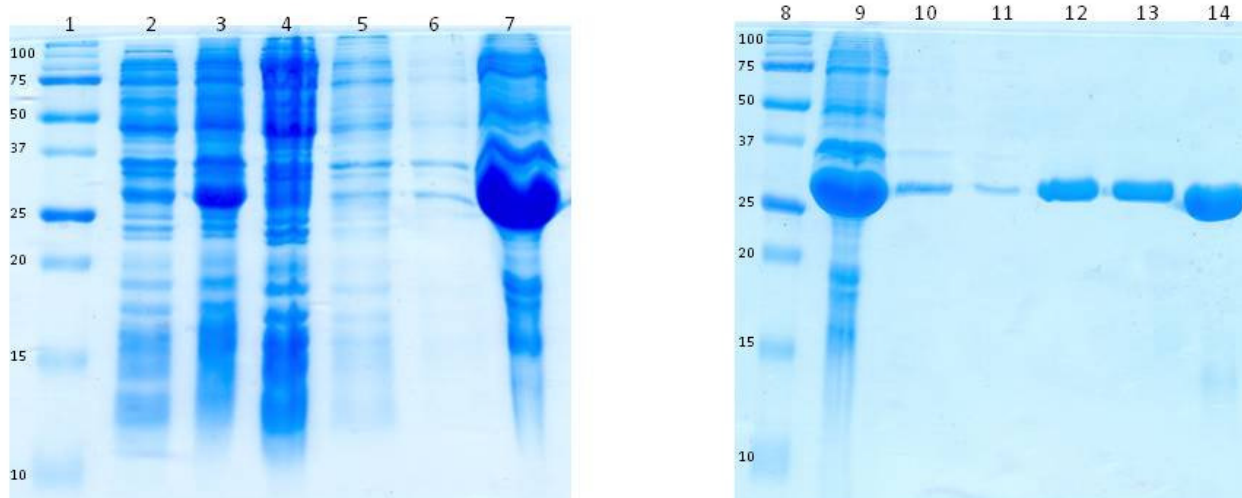


Figure 6. SDS-PAGE 12%: renaturation et purification de PELA. 1, 8. Marqueurs moléculaires (kDa). 2. Extrait cellulaire avant induction par l'IPTG. 3. Extrait cellulaire après 4 heures d'induction par l'IPTG. 4. Fraction soluble après la lyse. 5, 6. Lavage des corps d'inclusion dans 4 M l'urée. 7, 9. Fraction insoluble (corps d'inclusion) solubilisée dans 8 M d'urée. 10. Après renaturation. 11. Fraction non retenue sur la colonne de Mannose-agarose. 12, 13. Fractions retenues et éluées par 200 mM de mannose. 14. PELA purifiée et concentrée à 3mg/mL.

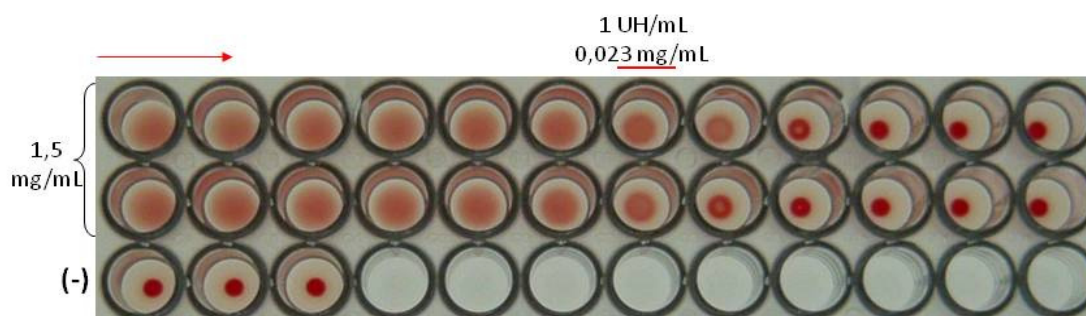


Figure 7. Hémagglutination des érythrocytes de lapin par la lectine PELA. (-) contrôle sans lectine.

2.2. Étude de la spécificité de PELA par Glycan Array

Les méthodes d'analyse de spécificité disponibles auprès du *Consortium for Functional Genomics* permettent de tester un très grand nombre d'oligosaccharides. La lectine PELA a été marquée avec le fluorochrome Alexa Fluor 488 et envoyée pour l'analyse. Plus de 450 oligosaccharides naturels ont été testés sur les puces à sucres (version 4.1). Trois concentrations différentes de lectine ont été testées : 50, 10 et 1 $\mu\text{g/mL}$. Les résultats obtenus pour la protéine PELA sont de très bonne qualité à une faible concentration (10 $\mu\text{g/mL}$). La comparaison entre les

résultats obtenus à trois concentrations différentes permet une analyse semi-quantitative pour les oligosaccharides les plus intéressants (Figure 8).

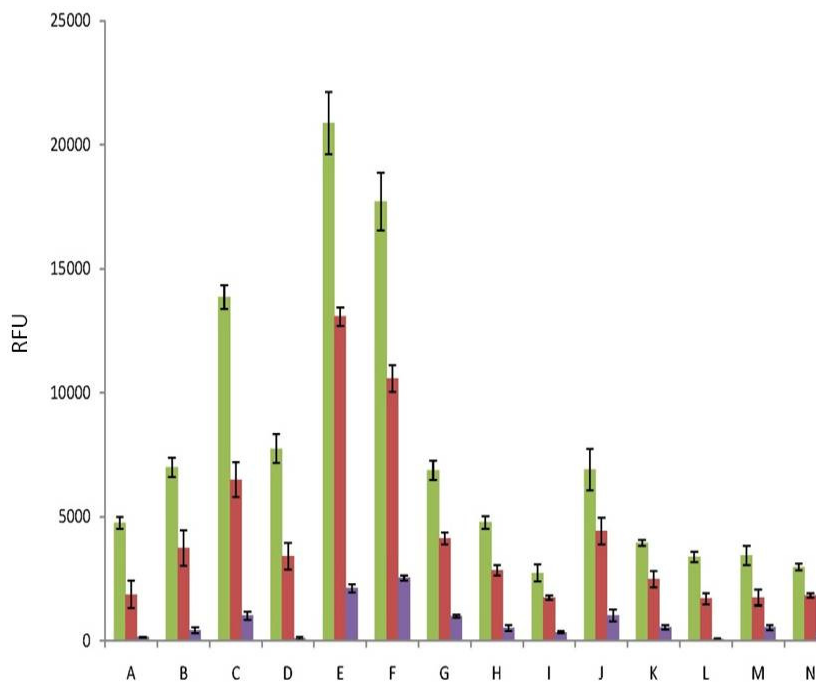
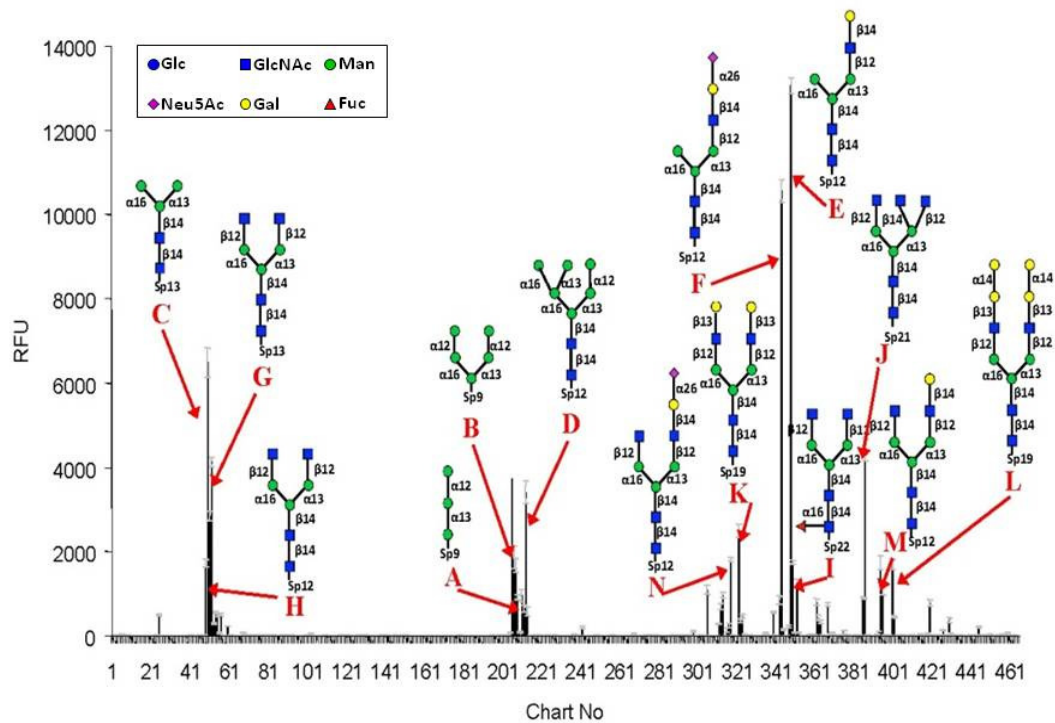


Figure 8. Glycan Array. Haut: liaison de PELA marquée (10 µg/ml) à tous les glucides disponibles de glycan array (version 4.1). Bas: Comparaison de la liaison des 14 principaux glucides (A-N) à trois concentrations de lectine (vert 50 µg/mL, orange 10 µg/mL, violet 1 µg/mL). RFU: relative fluorescence unit.

A partir des données obtenues par *Glycan Array*, la lectine PELa a montré une spécificité remarquable pour les glycanes biantennés, correspondant à la grande majorité des 14 principaux glucides (A-N) mis en évidence dans la figure 8. Les glycanes de types oligomannose (composés A à D) et les glycanes complexes biantennés et symétriques (composés G à L) sont très reconnus, mais une préférence nette apparaît pour des glycanes dissymétriques présentant une branche 1-6 courte ne comportant que le premier mannose et une branche 1-3 longue avec au moins un mannose et un GlcNAc (composés E et F). La préférence pour les glycanes dissymétriques est conservée pour les trois concentrations testées de la lectine (50, 10 à 1 $\mu\text{g.ml}^{-1}$). L'influence de l'acide sialique terminal sur la branche 1-6 semble être modérée puisque les oligosaccharides E et F ont des réponses équivalentes à des concentrations faibles de protéines. La comparaison de la fluorescence obtenue pour les oligosaccharides E et M, et aussi F et N, démontre clairement que l'ajout d'un GlcNAc sur la branche 1-6 diminue significativement l'affinité pour la lectine.

D'autres lectines de légumineuses ont été analysées par *glycan array* par différents groupes et les résultats sont disponibles sur le site du *Consortium for Functional Glycomics* (<http://www.functionalglycomics.org/>). On peut ainsi voir que ConA, PSA, DGL et LCA, toutes spécifiques pour Glc/Man, ont une préférence plus marquée pour les *N*-glycanes de type oligomannose, suivie par une spécificité moindre pour les glycanes complexes biantennés, puis plus faible pour les glycanes dissymétriques.

Ces données ont permis de décrire PELa comme une lectine présentant une spécificité rare pour des glycanes dissymétriques, et clairement différente des lectines de légumineuses Glc/Man spécifique de la famille de la ConA.

2.3. Étude de l'affinité de PELa pour les oligosaccharides complexes

L'affinité de PELa pour les oligosaccharides a été étudiée par microcalorimétrie isotherme de titration (ITC). L'ITC a été utilisé afin de mieux caractériser les interactions de PELa avec deux oligosaccharides biantennés liés à un résidu asparagine. Deux heptasaccharides-Asn ont été préparés par notre collaborateur Carlo Unverzagt (Univ. Bayruth, Allemagne), un symétrique

présentant deux branches se terminant par un GlcNAc, et un dissymétrique présentant un mannose isolé sur la branche 1-6 et un trisaccharide β Gal1-4 β GlcNAc1-2 α Man sur la branche 1-3 (Figure 9). La partie oligosaccharide de ces composés correspond aux oligosaccharides H et E de la figure 8.

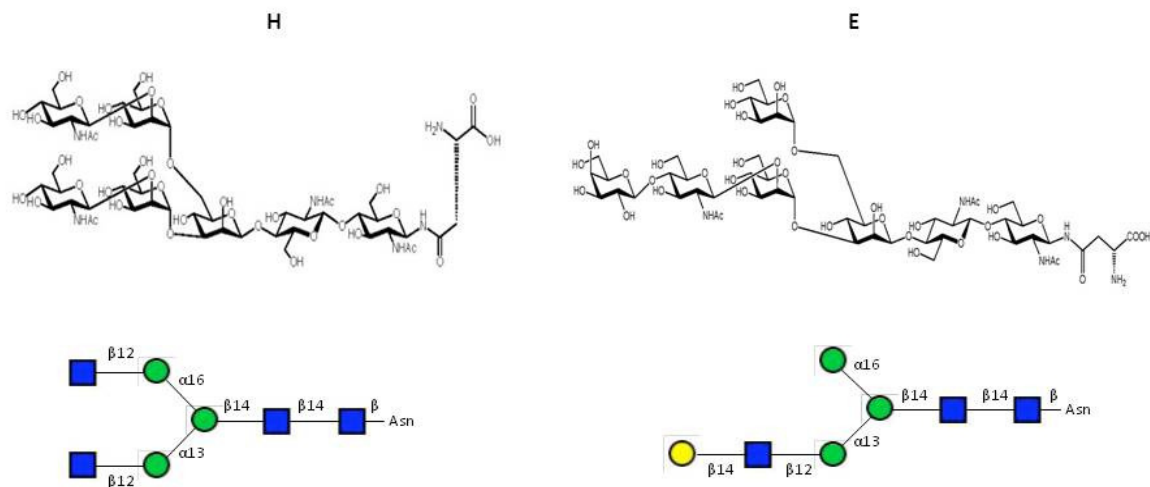


Figure 9. Oligosaccharides H à gauche (heptasaccharide symétrique) et E à droite (heptasaccharide dissymétrique) sous leur forme liée à l'Asn.

En utilisant un modèle simple d'un oligosaccharide par monomère, les deux oligosaccharides se lient à PELa avec des affinités similaires. Une constante de dissociation, K_d , proche de $5 \mu\text{M}$ est mesurée, ce qui correspond à une affinité forte pour une interaction protéine-sucre. Pour le composé H, un meilleur fit est obtenu en utilisant un modèle à deux sites ce qui indique que chacune des branches de l'heptasaccharide symétrique peut se lier à la lectine, mais avec des affinités différentes.(Figure 10). Il n'est cependant pas possible de déterminer de manière quantitative les affinités des deux sites.

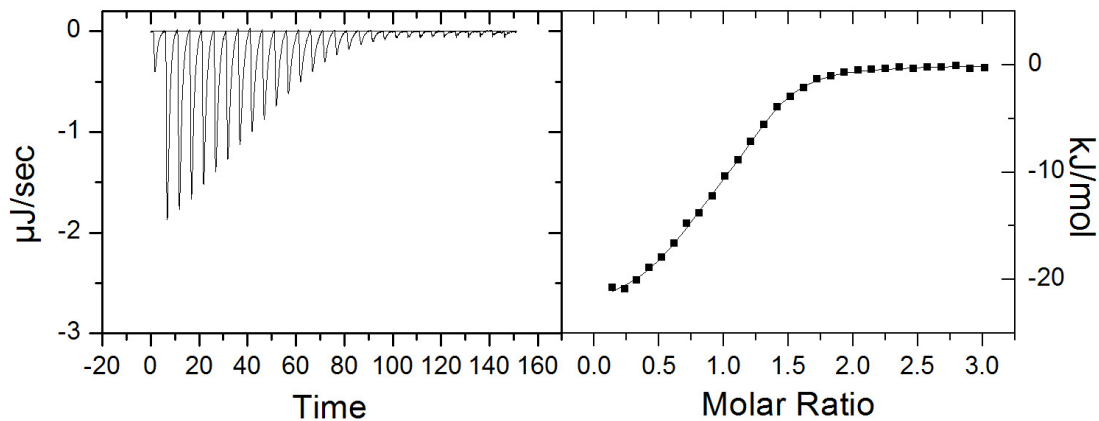


Figure 10. Courbe de microcalorimétrie obtenue pour PELa (0,06 mM) avec l’heptasaccharide symétrique H (0,775 mM). Thermogramme à gauche et intégration et ajustement au modèle théorique à deux sites à droite.

L’étude cristallographique de l’interaction entre PELa et l’heptasaccharide symétrique, ci-dessous, montre que chacun des mannoses (position 3 ou 6) peut venir se placer dans le site principal, ce qui peut expliquer le modèle à deux sites.

Les courbes de titration ont été obtenues pour l’interaction de PELa et les oligosaccharides E et H (figure 11) et les données thermodynamiques pour chaque interaction sont rassemblées dans le tableau 7 en utilisant un modèle à un site.

Tableau 7. Données thermodynamiques obtenues par ITC à 25°C pour la titration de la lectine PELa par les oligosaccharides H et E. Les déviations standards ont été calculées à partir de deux expériences indépendantes.

Ligand	K_a ($10^6 M^{-1}$)	K_d (μM)	n	ΔG (kJ/mol)	$-\Delta H$ (kJ/mol)	$-T\Delta S$ (kJ/mol)
H	0,22	4,4 (0,3)	1,06 (0,6)	-30,6	23,0 (0,3)	-7,6 (0,4)
E	0,16	6,0 (0,2)	1	-29,8	59,3 (1,8)	29,4 (1,8)

L’affinité pour les deux oligosaccharides testés (K_a) est de l’ordre du micromolaire (μM), et représente une affinité très forte par rapport aux autres lectines de légumineuses. Les seules activités lectiniques d’affinité comparable ont été obtenues avec des oligosaccharides. Les lectines PAL et ConA (GARCIA-PINO *et al* 2007; MANDAL *et al.*, 1994, BUTS *et al.*, 2006), ont montré des affinité de l’ordre du micromolaire pour les oligosaccharides (ConA-trimannose : $K_d = 30 \mu M$; PAL-trimannose : $K_d = 16 \mu M$; PAL-biantenne : $K_d = 16 \mu M$).

L'affinité entre l'oligosaccharide H et la lectine est encore 1,36 x supérieure à l'affinité montrée par l'oligosaccharide E.

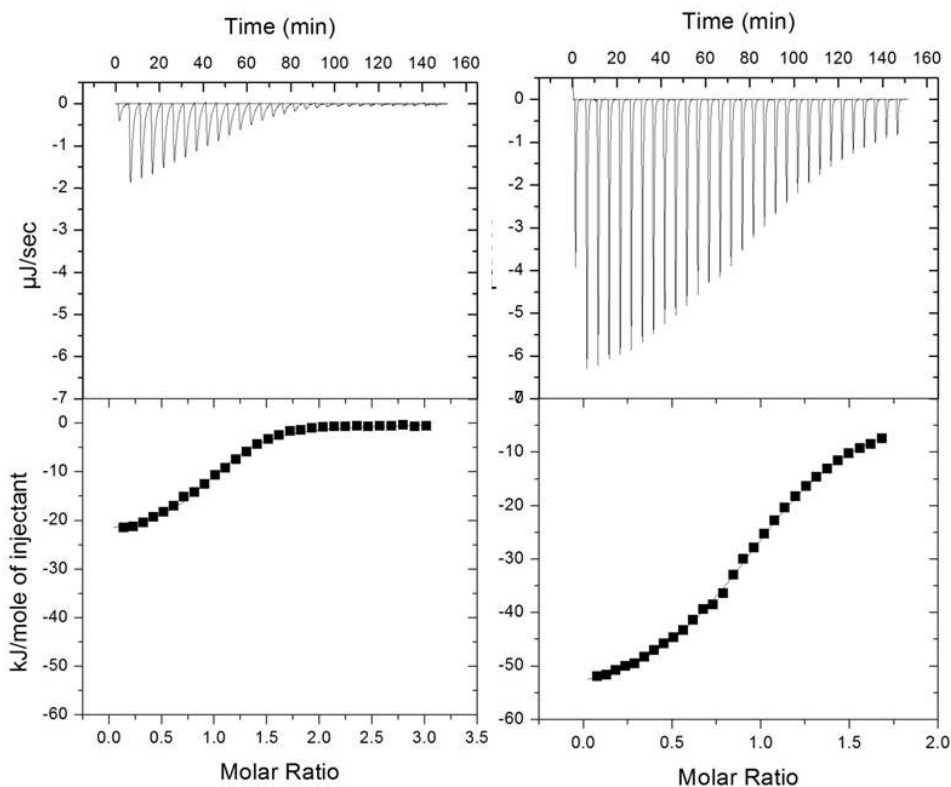


Figure 11. Courbes de microcalorimétrie obtenues pour PELA (0,6 mM) et les oligosaccharides H (0,775 mM) à gauche, et E (0,45 mM) à droite, dans un tampon Tris (pH 7,6) à 25 ° C. Haut, thermogrammes obtenus pour 30 injections automatiques de 10 µl de chaque sucre dans la cellule de mesure contenant la protéine. Bas, intégration de la chaleur totale libérée en fonction de la titration de ligand. La ligne continue représente le meilleur ajustement aux données obtenues par une courbe théorique à un site.

L'affinité de PELA pour les deux oligosaccharides est proche, et donc la variation d'énergie libre (ΔG) également. Cependant l'analyse des contributions thermodynamiques indique que les modes d'interactions sont différents. L'interaction de PELA avec l'oligosaccharide H est dictée par une contribution enthalpique forte ($\Delta H = -23$ kJ/mol), avec une faible contribution entropique favorable ($-T\Delta S = -7,6$ kJ/mol), environ 20% de l'énergie de liaison, peu habituelle pour les interactions impliquant les oligosaccharides. L'interaction de PELA avec l'oligosaccharide E a une très forte contribution enthalpique ($\Delta H = -59,3$ kJ/mol), compensée par une contribution entropique très défavorable ($-T\Delta S = 29,4$ kJ/mol). Généralement, les associations protéine/sucre sont caractérisées par un terme enthalpique favorable du au nombre élevé de liaisons hydrogène et contacts van der Waals entre les faces hydrophobes du sucre et des acides aminés aromatiques. La contribution entropique est généralement défavorable, ce qui

a été attribué à la fois au réarrangement des molécules de solvant et à la perte de flexibilité conformationnelle/mobilité du ligand (DAM et BREWER, 2002).

3. Étude structurale de la lectine PELa en complexe avec différents oligosaccharides.

Les structures cristallines de la lectine PELa en complexe avec les oligosaccharides A (trimannose) et H (heptasaccharide symétrique), (figure 12) ont été obtenues.

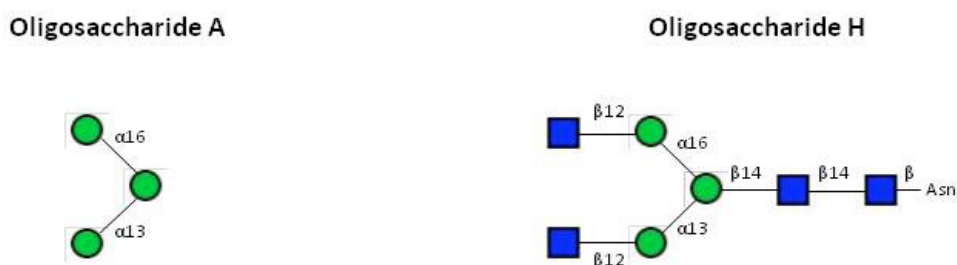


Figure 12. Représentation des oligosaccharides A (trimannose) et H (heptasaccharide symétrique), utilisés dans des expériences de cristallographie.

Le premier jeu de données (PELa_tri) a été collecté à une résolution de 2,1 Å et le deuxième (PELa_hept) à 1,6 Å. La première structure résolue est celle du complexe PELa/trimannose et elle est utilisée ci-dessous pour l'analyse structurale de la lectine.

3.1. Structure cristallographiques de PELa en complexe avec le trimannose

A partir d'une solution de protéine à environ 10 mg/ml et de trimannoside à 1mM, les cristaux ont été obtenus avec une solution contenant 12 % glycérol, 1,5 M $(\text{NH}_4)_2\text{SO}_4$ et 100 mM Tris-HCl pH 8,5 (Hampton screen 2 solution 42). Des données ont été collectées à l'ESRF sur la ligne de lumière ID14-1 à une résolution de 2,1 Å dans le groupe d'espace $P2_12_12_1$ ($a= 42,62$, $b= 80,68$ et $c= 148,76$ Å; $\alpha=\beta=\gamma=90,0^\circ$) après trempage des cristaux dans une solution où la concentration en glycérol a été augmentée à 25 % afin de les congeler dans l'azote liquide.

La structure a été résolue par remplacement moléculaire en utilisant les coordonnées de la lectine de *Pterocarpus angolensis* (PDB : 1N3O) comme modèle pour le programme PHASER.

Les données statistiques détaillées de la collecte, du traitement des données, du remplacement moléculaire et de l'affinement sont présentées dans le tableau 8.

Deux molécules de protéine et une molécule de trimannose sont observées dans l'unité asymétrique. Chaque monomère de PELa présente le repliement classique des lectines de légumineuses et inclut un ion calcium et un ion manganèse proche du site de fixation du sucre (Figure 13). Les deux monomères dans l'unité asymétrique sont associés par un arrangement parallèle d'un des feuillets- β , formant ainsi le «dimère canonique des lectine de légumineuses » qui est classiquement observé dans les arrangements dimériques et tétramériques de ces protéines (MANOJ & SUGUNA, 2001; SRINIVAS et al., 2001). Les deux sites potentiels de glycosylation (Asn7 et Asn127) sont situés respectivement dans la région N-terminale désordonnée sur la petite hélice- α à la surface de la protéine, et sont donc susceptibles d'être glycosylés dans les plantes.

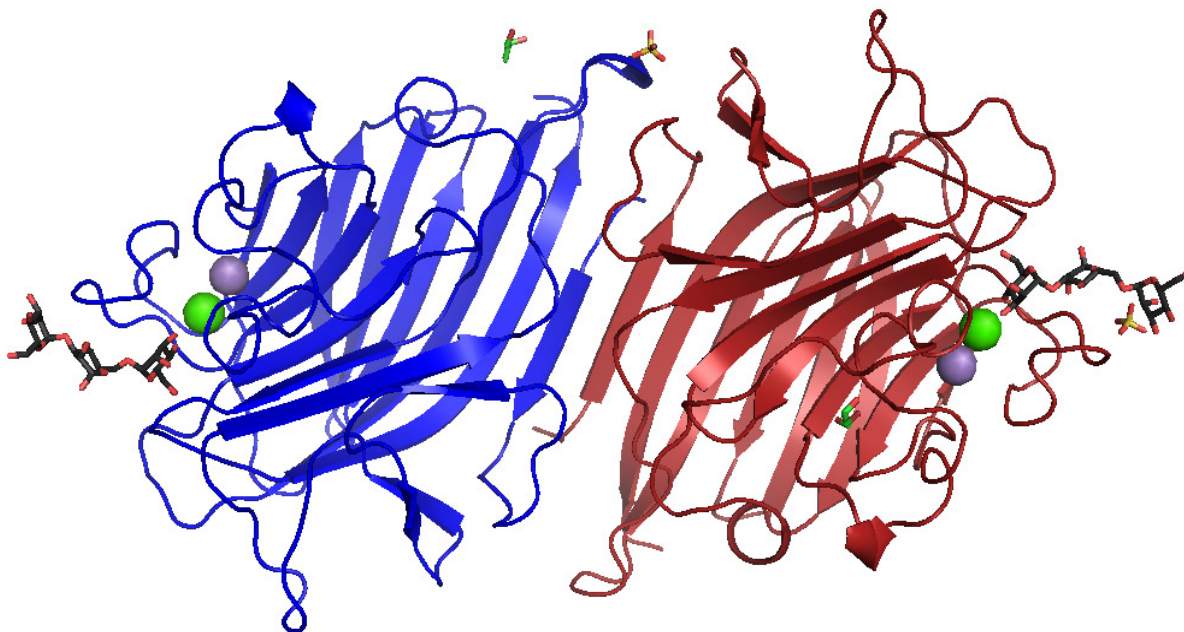


Figure 13. Présentation de la structure dimérique de PELa en complexe avec le trimannose formé par le monomère A en rouge et le monomère B en bleu. On note la présence des métaux Ca^{2+} et Mn^{2+} (sphères vertes et violettes, respectivement), du ligand ($\alpha\text{Man1-3}(\alpha\text{Man1-6})\text{Man}$) (en noir), de deux molécules de glycérol (en jaune) et de deux ions sulfates (en vert).

Tableau 8. Données statistiques de la collecte et de l'affinement des structures cristallines de PELa en complexe avec le trimannose et l'heptasaccharide.

	Trimannose
Ligne de lumière (ESRF) (Longueur d'onde, Å)	Id14EH1 (0,933)
Groupe d'espace	P2 ₁ 2 ₁ 2 ₁
Maille <i>a, b, c</i> (Å) <i>α, β, γ</i> (deg,)	42,62 80,68 148,76 90,0 90,0 90,0
Reflections mesurées/unique	99535/28894
Résolution (outer shell)*, Å	42,6 - 2,10 (2,21-2,10)
Multiplicité moyenne	3,4 (2,4)
R _{merge}	0,082 (0,330)
Completude (%)	93,7 (70,1)
Moyenne I/σI	9,3 (2,0)
Facteur B de Wilson	24,7
<i>R</i> _{cryst} / <i>R</i> _{free}	19,5/25,2
R _{msd} bonds, Å	0,015
R _{msd} angles, °	1,54
Rmsd chiral, Å ³	0,106
Nombre d'atomes: Protéine (A/B) Bfac Å ²	1935/1926 23,5/23,7
Nombre d'atomes: Solvent (A/B) Bfac, Å ²	177/149 29,0/28,3
Nombre d'atomes: Ligand (A/B) Bfac, Å ²	34 22,6
Nombre d'atomes: Hetero (A/B) Bfac, Å ²	18/8 32,4/33,5
Code PDB deposition	3ZVX

* Les nombres entre parenthèses font référence à la zone de plus haute résolution.

L'examen de la première carte de densité électronique a révélé sans ambiguïté la position du trisaccharide qui établit un pont entre les dimères dans le cristal (Figure 14). De par la position du trimannose, le site de liaison de chaque monomère partage des portions du même ligand. La figure 15 donne une liste des distances des liaisons hydrogène présentes dans le site de liaison des monomères A et B de la lectine. Les monomères A et B fixent respectivement les α Man liés en 1-3 et en 1-6 sur le mannose central. Dans chacun des sites, les acides aminés impliqués sont ceux qu'on retrouve habituellement dans les interactions entre lectine de légumineuse et mannose (Asp95, Ala113, Gly115, Asn145, Ser146, Asn147, Gln231). Des contacts protéine-protéine sont créés par le pontage du trimannose avec la participation de liaisons hydrogène entre la chaîne latérale de Arg123B et les chaînes principales de Thr148A et Asp150A.

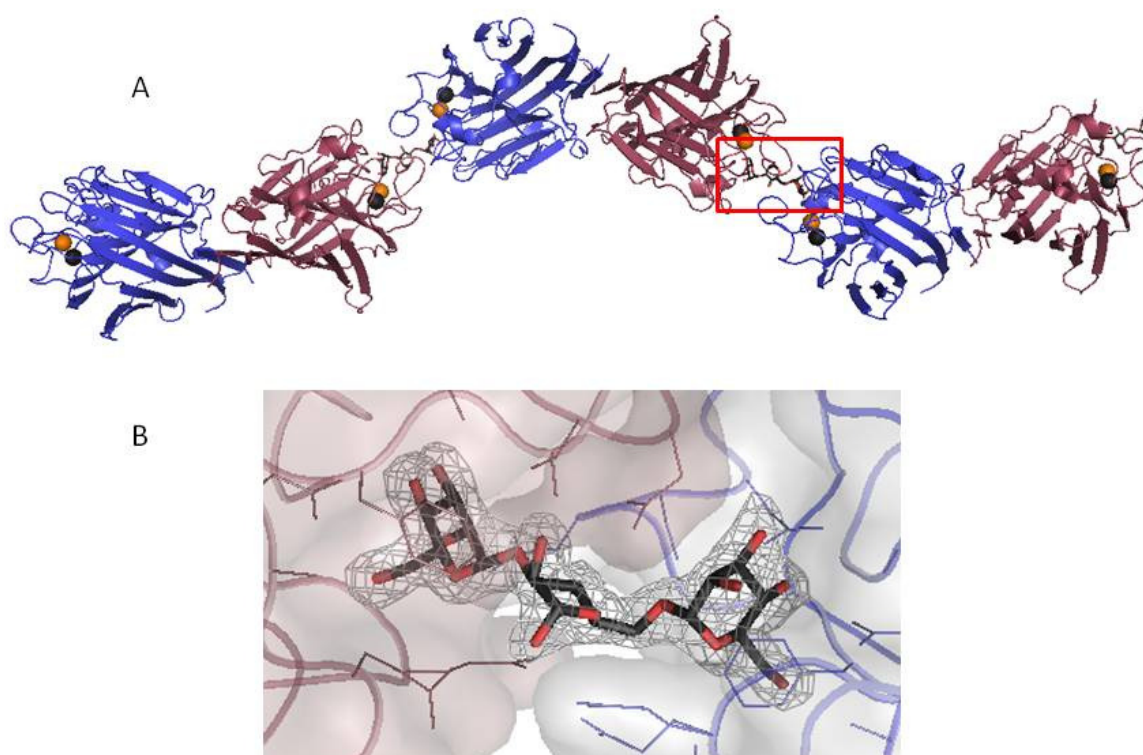


Figure 14. Structure cristalline de PELA en complexe avec trimannose. A. Le pont entre les dimères liés par une molécule de trimannose. B. Densité électronique expérimentale du ligand trimannose entre les monomères A et B de dimères différents.

Les deux atomes calcium et manganèse sont liés de façon réversible à la protéine dans une poche proche du site de reconnaissance. Ces métaux sont situés à une distance de 4,2 Å de l'autre et sont coordonnés par les chaînes latérales de six résidus (Glu137, Asp139, Phe141,

Asn147, Asp150, His155) et par quatre molécules d'eau. Le calcium maintient la conformation cis de l'Asp95 du site de liaison du sucre qui est conservée dans toutes les lectines de légumineuses.

Le site de liaison de glucides dans toutes les lectines de légumineuses partage des résidus qui composent jusqu'à cinq boucles, A à E (SHARMA & SUROLIA, 1997). La taille et la séquence de ces boucles varient à des degrés divers entre les lectines ayant des spécificités différentes. La figure 16 donne une représentation des boucles de PELa et leurs implications dans le complexe avec le trimannose.

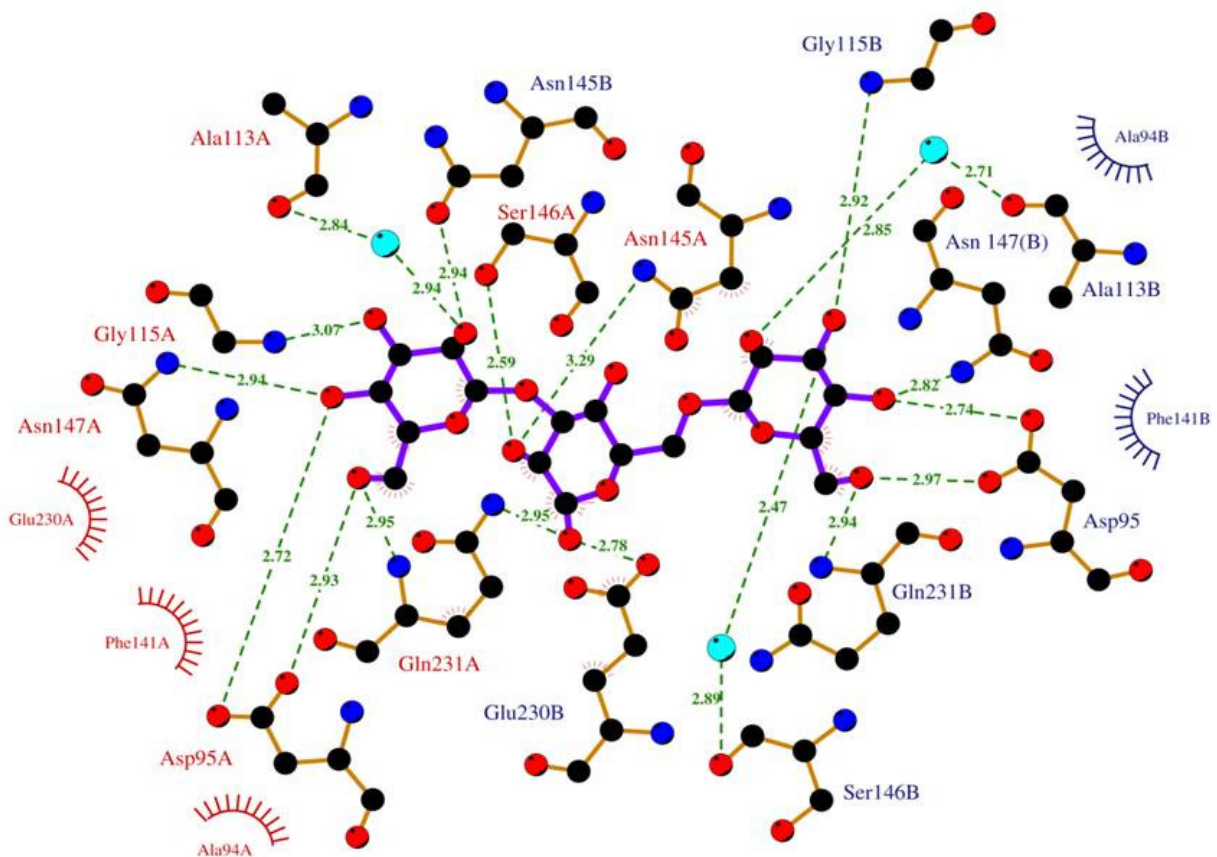


Figure 15. Représentation schématique bidimensionnelle (Ligplot) du réseau de liaisons hydrogène entre les résidus clés impliqués dans l'interaction entre PelA et le trimannose, partagé entre les monomères A (résidus en rouge) et B (résidus en bleu) de dimères différents. Longueur des liaisons en Å.

Les boucles A et B de PELa contiennent respectivement les résidus Asp95 et Gly115 qui sont impliqués dans la reconnaissance du mannose. La boucle C est appelée boucle de liaison

aux métaux . Dans PELa, cette boucle a une longueur de 16 résidus et une conformation semblable à celle de PAL et des lectines de *Maackia amurensis* et *Ulex europaeus* spécifiques pour le sialylactose et le chitobiose, respectivement (LORIS *et al*, 2004, IMBERTY *et al* 2000; LORIS *et al*, 2000). Les autres lectines spécifiques pour Glc/Man ont une boucle C avec 14 (ConA) ou 15 résidus (LOL) (LORIS *et al.*, 2004). Concernant la boucle D, elle est similaire à celles présentes dans les structures de lectines spécifiques pour Glc/Man (LORIS *et al* 2004; HAMELRYCK *et al*, 2000, SANZ-APARICIO *et al* 1997; WAH *et al* . 2001; DEREWENDA *et al*, 1989, BOURNE *et al* 1990; RINI *et al*, 1993). La boucle E n'est pas impliquée dans la liaison de PELa au trimannose. Cette boucle peut intervenir dans l'interaction de certaines lectines avec des oligosaccharides telles que la leucoagglutinine (MAL) en complexe avec le sialylactose (LORIS *et al.*, 2000), la lectine IV de *Griffonia simplicifolia* (GS-IV) en complexe avec un tetrasaccharide (DELBAERE *et al.*, 1993) et ConA et PAL en complexe avec le Man(α 1-2)Man (MOOTHOO *et al.*, 1999, LORIS *et al.*, 2004).

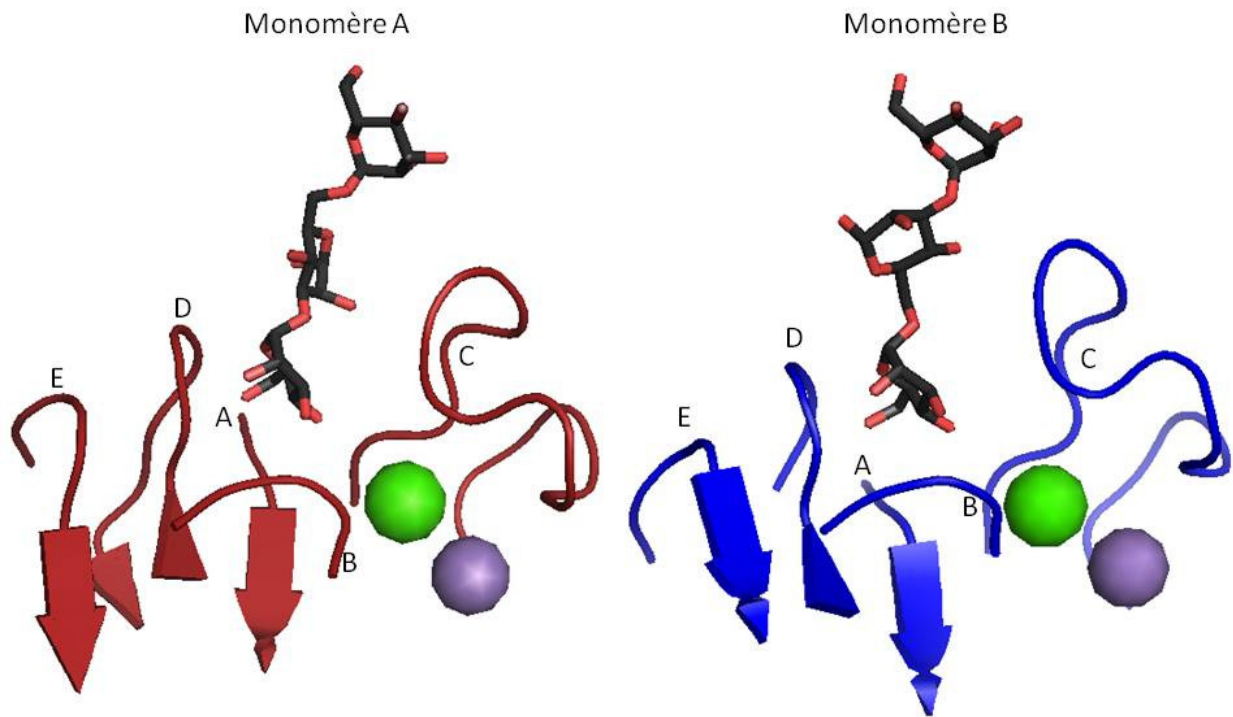


Figure 16. Représentation de l'architecture du site de liaison de PELa (monomères A et B) en complexe avec le trimannose, montrant les cinq boucles (A-E) typiques des lectines de légumineuses. Dans cette interaction, les résidus appartenant aux boucles A, B, C et D participent de l'interaction lectine-sucres.

En raison de son interaction avec les deux monomères, la forme du trimannoside est très étendue et différente de ce qui a été observé pour le même ligand dans d'autres structures de lectines de légumineuses telles que PAL (Loris, 2004), ConA (Loris, 1996) et DGL (Rozwarski, 1998). La comparaison entre PELA et PAL confirme la similarité des sites de liaison avec des conformations identiques pour les cinq boucles nommées de A à E (figure 17). PAL lie le trimannoside avec le bras man1-6 dans le site de liaison principal, comme le monomère B de PELA. Néanmoins, la conformation de l'oligosaccharide observée dans PAL est très différente puisque le trimannoside est recourbé autour du résidu phényle du site de liaison et le 3-mannose établit une liaison hydrogène avec un résidu asparagine dans la boucle A.

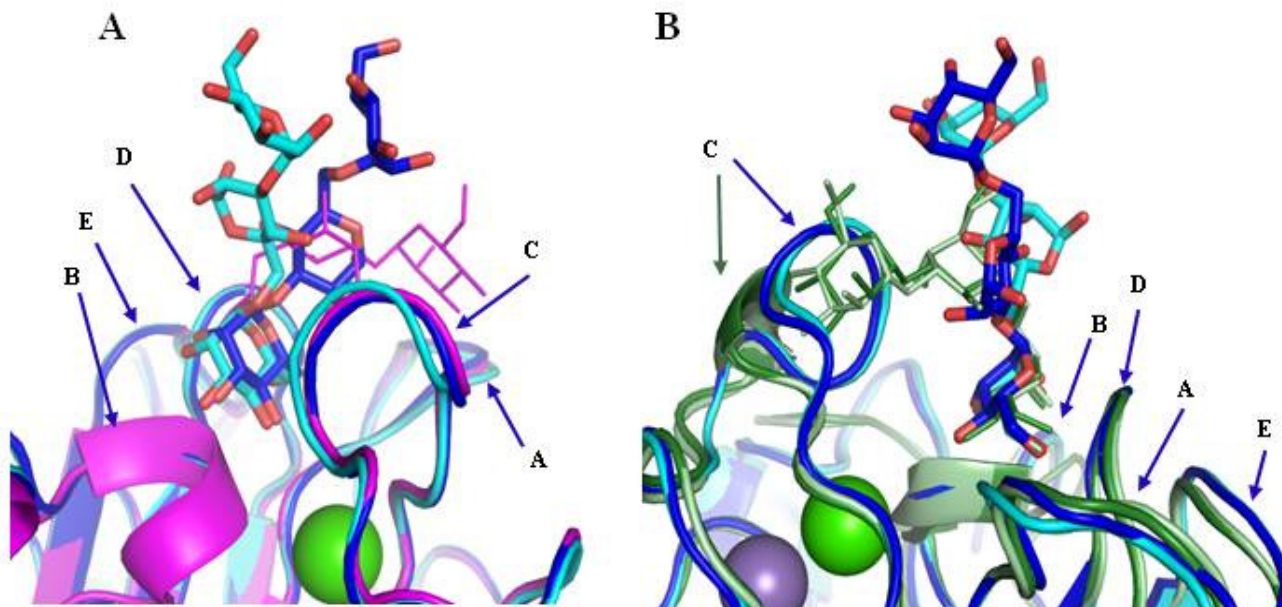


Figure 17. Comparaison de la reconnaissance du trisaccharide pour PELA et autres lectines de légumineuse. A. Les sites de liaison superposés des monomères A (bleu) et B (cyan) du complexe PELA:trimannose avec le complexe PAL:trimannose (magenta, code PDB : 1QV8). B. Vue des monomères A et B de PELA avec les lectines ConA (vert foncé, code PDB : 1ONA) et DGL (vert clair, code PDB : 1DGL). Les boucles des sites de liaison sont montrées.

Une comparaison avec les lectines de la tribu Dioclae est encore affichée dans la figure 17. Les boucles des sites de liaison sont bien conservées entre Dioclae (ConA et DGL) et Dalbergiae (PAL et PELA), à l'exception de la boucle C qui est plus courte dans ConA / DGL que dans PAL / PELA (14 acides aminés au lieu de 16) résultant en un site de liaison plus ouvert tel que décrit précédemment (LORIS *et al.*, 2004). En conséquence, la conformation du trimannose est encore différente dans ConA et DGL: il est replié sur la surface de la protéine et le mannose

1-3 est en contact avec des acides aminés supplémentaires de la boucle C. La conformation observée dans ConA ne peut pas être possible dans PELa en raison de la conformation différente de la boucle C qui entraînerait un conflit stérique.

Même si le trimannose a été observé dans les structures cristallines de plusieurs lectines, cette étude rapporte le premier cas de chaîne infinie par liaison au trimannose observé en cristallographie pour des lectines de légumineuses.

3.2. Structure cristallographique de PELa en complexe avec l'heptasaccharide symétrique

Pour le complexe de PELa avec l'heptasaccharide symétrique GlcNAc β 1-2Man α 1-3(GlcNAc β 1-2Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAcB-Asn, des cristaux ont été obtenus avec une solution contenant 10 % PEG 20K, 2 % dioxane et 100 mM bicine pH 9 (Hampton screen 2 solution 48). Des données à 1,6 Å ont été collectées sur BM30A à l'ESRF avec 30 % glycérol comme cryoprotectant. Les cristaux appartiennent au groupe d'espace P2₁2₁2₁ (a= 51,40, b= 76,98 et c=125,59 Å, $\alpha=\beta=\gamma= 90.0^\circ$). La structure a été résolue par remplacement moléculaire en utilisant les coordonnées de la structure de PELa en complexe avec le trimannoside comme modèle. Les données statistiques détaillées de la collecte, du traitement des données, du remplacement moléculaire et du raffinement sont présentées dans le tableau 9.

L'analyse de la carte de densité électronique a clairement révélé la position de l'heptasaccharide présent dans le site de liaison de chaque monomère. La figure 18 montre la densité électronique expérimentale sur l'heptasaccharide dans le site de liaison de PELa, proche des ions Ca²⁺ et Mn²⁺. Les contacts de l'oligosaccharide avec la surface protéique sont beaucoup plus étendus que dans le complexe avec le trimannose. La plupart des interactions se font sur le core trimannose et la branche 1-3 (GlcNAc β 1-4Man α 1-3)Man, confirmant ainsi les préférences observées sur le *glycan array*.

Tableau 9. Données statistiques de la collecte et de l'affinement des structures cristallines de PELa en complexe avec le trimannose et l'heptasaccharide.

	Heptasaccharide
Ligne de lumière (ESRF) (Longueur d'onde, Å)	BM30A (0,9797)
Groupe d'espace	P2 ₁ 2 ₁ 2 ₁
Maille <i>a, b, c</i> (Å) <i>α, β, γ</i> (deg,)	51,40 76,98 125,59 90,0 90,0 90,0
Reflections mesurées/unique	186416/55022
Résolution (outer shell)*, Å	47,57 - 1,65 (1,74-1,65)
Multiplicité moyenne	3,4 (2,6)
R _{merge}	0,054 (0,250)
Completude (%)	91,4 (65,8)
Moyenne I/σI	13,6 (3,9)
Facteur B de Wilson	14,00
<i>R</i> _{cryst} / <i>R</i> _{free}	15,5/18,8
R _{msd} bonds, Å	0,015
R _{msd} angles, °	1,59
Rmsd chiral, Å ³	0,098
Nombre d'atomes: Protéine (A/B) Bfac Å ²	2004/1986 12,4/14,0
Nombre d'atomes: Solvent (A/B) Bfac, Å ²	295/223 25,7/26,6
Nombre d'atomes: Ligand (A/B) Bfac, Å ²	89/92 14,2/16,6
Nombre d'atomes: Hetero (A/B) Bfac, Å ²	30/12 24,7/25,7
Code PDB deposition	3ZYZ

* Les nombres entre parenthèses font référence à la zone de plus haute résolution.

Le site de liaison de PELa en complexe avec l'heptasaccharide implique principalement 13 résidus d'acides aminés qui interagissent avec la partie biantennée du ligand, en particulier le

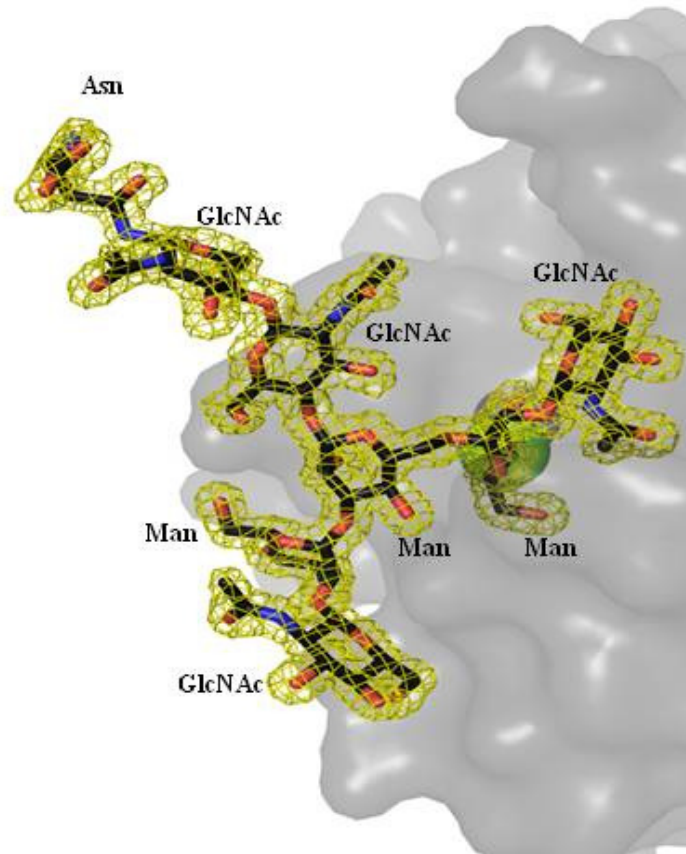


Figure 18. Représentation de la surface du site de liaison de PELa en complexe avec l'heptasaccharide. Densité électronique expérimentale du ligand sur la surface du site de liaison de PELa.

core mannose et la branche 1-3. La figure 19 décrit les liaisons hydrogène présentes dans le site de liaison de la lectine PELa avec l'heptasaccharide.

Parmi les résidus d'acides aminés impliqués dans le site, huit résidus correspondent à ceux déjà décrits dans le site de reconnaissance du trimannose (Asp95, Ala113, Gly115, Asn145, Ser146, Asn147, Glu230, Gln231). Les cinq autres résidus (Ser54, Ser91, Asn92, Gly111, Gly229) établissent des contacts supplémentaires. De plus, cinq molécules d'eau sont représentées dans l'interaction.

Le nombre de liaisons hydrogène impliquant le bras 1-3 est plus élevé que celles impliquant le bras 1-6. Plus de 15 liaisons hydrogène directes et 6 liaisons médiées par des molécules d'eau sont observées, rationalisant ainsi la grande affinité mesurée par microcalorimétrie de titration.

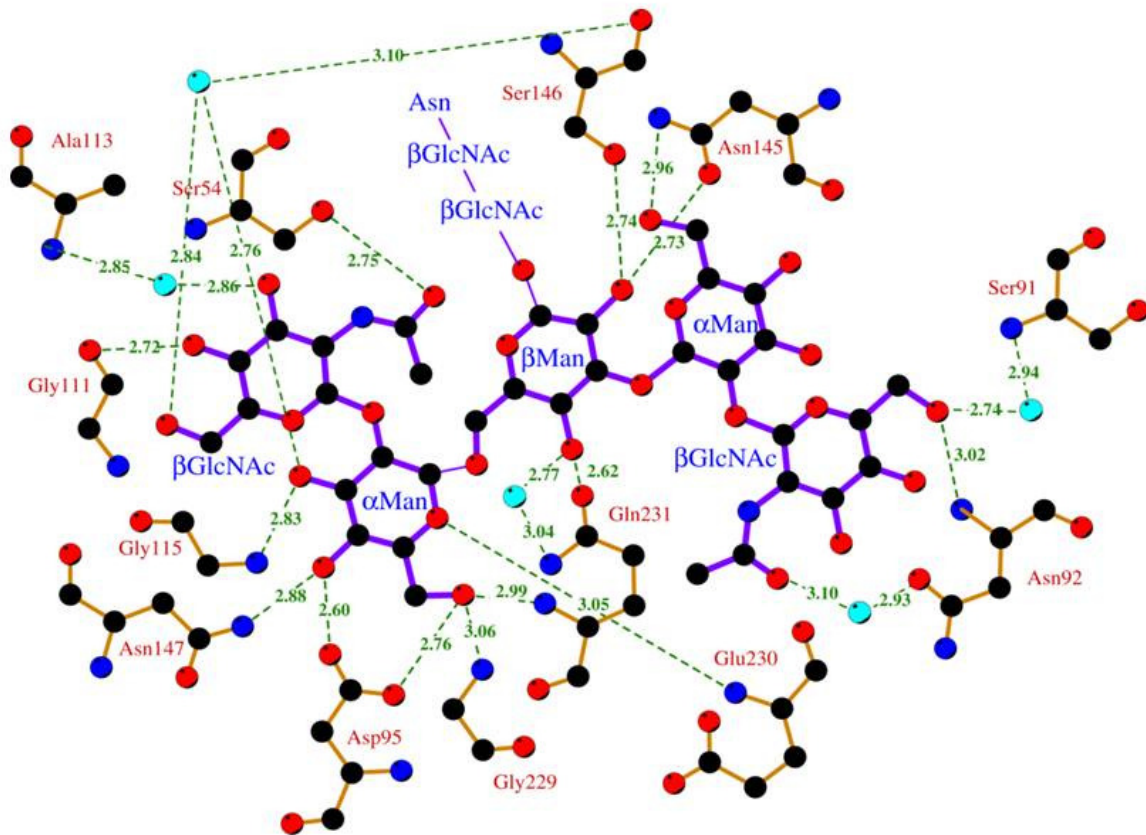


Figure 19. Représentation schématique bidimensionnelle (Ligplot) du réseau de liaisons hydrogène entre les résidus clés impliqués dans l'interaction du site de liaison de PELA en complexe avec le ligand heptasaccharide. Longueur des liaisons en Å.

L'analyse de la participation des cinq boucles (A-E) de l'architecture des sites de liaison des lectines de légumineuses (SHARMA et SUROLIA, 1997) présente des différences par rapport à l'interaction avec le trimannose. La figure 20 donne une représentation des boucles de PELA et de leur implication dans le complexe avec l'heptasaccharide. La participation du résidu Ser54 dans cette interaction implique la participation de la boucle E, qui n'était pas présente dans l'interaction avec le trimannose.

La boucle E interagit avec les glucides dans de rares cas, en général en agissant comme un sous-site. Parmi les lectines de légumineuses à Glc/Man, la boucle E joue un rôle dans les sites de liaison de ConA et PAL avec Man (α 1-2) Man (MOOTHOO *et al.*, 1999, LORIS *et al.*, 2004). Dans PELA, Ser54 stabilise l'extrémité du bras 1-3, ce qui peut expliquer une partie de la faible entropie observée dans l'interaction de la lectine avec ce sucre.

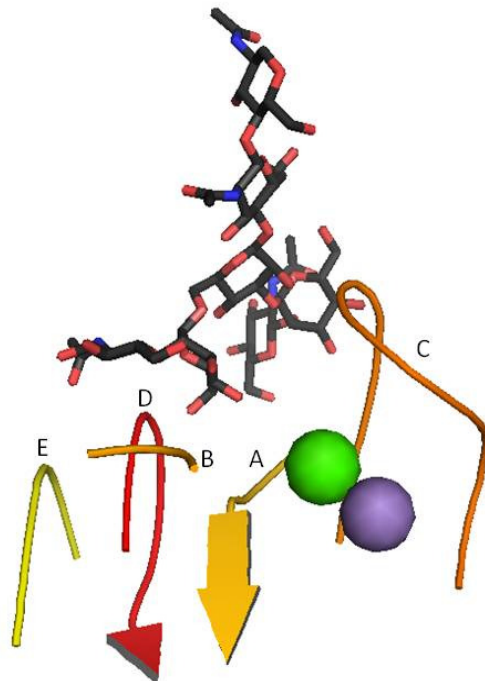


Figure 20. Représentation de l'architecture du site de liaison de glucides de PELA en complexe avec l'heptasaccharide, montrant les cinq boucles (A-E) typiques des lectines de légumineuses. Dans cette interaction, des résidus appartenant aux cinq boucles A, B, C, D et E sont impliqués dans la liaison aux glucides.

Enfin une comparaison a été faite avec les structures des complexes avec le trimannose et l'heptasaccharide, en superposant les deux structures (PDB: 3ZVX et 3ZYR, respectivement) (figure 21). Dans cette comparaison, il apparaît que le mannose du bras 1-6 de l'heptasaccharide occupe le site de liaison primaire de sucres de PELA dans une manière similaire à celle observée pour les trimannosides. La position du second mannose (core mannose des *N*-glycannes) est assez similaire dans les trois complexes. La différence principale provient de l'orientation du deuxième bras qui pointe vers l'extérieur (vers l'autre monomère) pour les trisaccharides et qui se recourbe vers la surface de la protéine pour l'heptasaccharide en établissant de nouveaux contacts pour le fragment $\alpha\text{GlcNAc1-2Man}$.

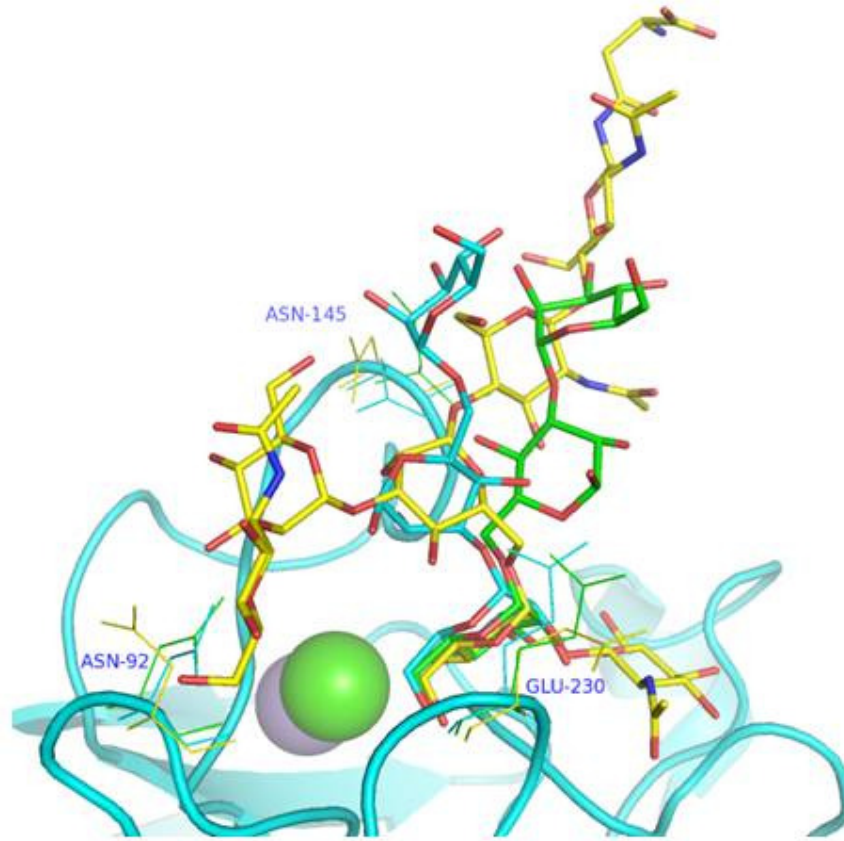


Figure 21. Comparaison de la reconnaissance des ligands trimannose et heptasaccharide par la lectine PELa. Comparaison entre les trois modes de liaison étudiés: Trimannose lié par les résidus α Man1-3 (cyan) et par le α Man1-6 (vert), et l'heptasaccharide (jaune).

IV. Conclusion et Perspectives

La conception et l'exécution de ce travail de thèse a permis le clonage du gène d'une lectine (dénommée PELa) des graines de la légumineuse *Platypodium elegans*, et la production en système hétérologue de la protéine. L'analyse de la spécificité de cette lectine recombinante a pu être effectuée, ainsi que la caractérisation des aspects structuraux et thermodynamiques de l'interaction avec plusieurs oligosaccharides. Ce travail contribue à la compréhension de la façon dont la reconnaissance des glucides est impliquée dans différents processus biologiques par les lectines.

La lectine recombinante A de *P. elegans* (PELa) spécifique pour Glc/Man, a montré une spécificité inhabituelle pour les oligosaccharides du type complexe biantenné, avec une préférence particulière pour les formes asymétriques, avec un bras 1-3 court. La protéine a montré une affinité considérable pour ce ligand, par rapport à d'autres lectines de la même famille, ce qui permet de faire un point sur la relation possible entre cette classe d'oligosaccharides et le ligand naturel de la lectine.

L'étude de la lectine des graines de *P. elegans* et sa caractérisation biochimique ont donc fourni des informations structurales importantes au sujet des interactions protéines-sucres nécessaires à la compréhension des mécanismes d'action, de la reconnaissance, de la spécificité et de l'affinité aux oligosaccharides complexes au niveau moléculaire. Une fois en possession de ces informations, l'analyse de la capacité de la lectine à reconnaître les glycanes impliqués dans des événements biologiques intéressants, ou la recherche de sa fonction dans l'organisme d'origine, pourront être entreprises.

Le plus grand défi maintenant sera de tenter de répondre à la question sur l'importance biologique de cette interaction pour la plante d'origine: quels sont les événements liés, quelles sont les étapes de la vie de l'organisme impliquées et comment ces processus peuvent-ils être explorés. La connaissance du gène et de la forme recombinante de la lectine, combinée avec la connaissance de son interaction avec des oligosaccharides complexes, sont des outils importants de départ pour cette approche.

V. Résumé sur les autres travaux réalisés au cours de la thèse.

Les travaux impliquant l'analyse structurale, la spécificité et l'affinité de PELa sous sa forme recombinante ont été utilisés pour la rédaction de cette thèse, car ils rassemblent une grande quantité de résultats obtenus par la collaboration des équipes françaises et brésiliennes. Deux autres projets ont été débutés dans le cadre de cette collaboration mais les difficultés techniques rencontrées dans la purification des protéines et/ou d'obtention du matériel génétique à partir des échantillons biologiques n'ont pas permis d'aller aussi loin que pour PELa.

Protéine	Justification	Objectif	Situation/problèmes
Lectine de l'algue <i>Codium isthmocladum</i>	Très peu de lectines ont pu être isolées et caractérisées à partir d'algues. Le laboratoire brésilien a pu mettre en évidence une activité lectine dans les extraits de l'algue verte <i>C. isthmocladum</i> .	<ul style="list-style-type: none"> - Clonage et expression hétérologue - Cristallisation en complexe avec les métaux lourds 	Après plusieurs tentatives, nous avons préparé l'ADN génomique de l'algue. Les tentatives pour identifier le gène de la lectine à partir d'amorces correspondant à la séquence N-terminale se sont révélées infructueuses.
Lectine de graine de la mimosaïdé <i>Albizia lebeck</i>	Les mimosaïdés contiennent des lectines de la famille de la Jacaline. Les études préliminaires menées au Brésil ont montré que les extraits de graine d' <i>A. lebeck</i> (ALL) ont une activité lectine spécifique du galactose.	<ul style="list-style-type: none"> - Etude de la spécificité - Clonage et expression hétérologue 	<p>La lectine a été envoyée au CFC pour être testée sur les puces à sucre après marquage. Aucune interaction significative n'a été observée, probablement due à une détérioration de la lectine au cours du marquage ou du transport.</p> <p>Le clonage a été tenté à partir des ARNs de graines. Cependant il n'a pas été possible d'identifier le gène de la lectine à partir des amorces préparées avec plusieurs stratégies.</p>

Pendant la période du doctorat, d'autres projets ont été réalisés dans chacune des deux universités.

Des projets concernant des lectines végétales ont été exécutés à l'Université Fédérale de Ceará, résultant en sept publications comme co-auteur. Les travaux ont porté sur l'étude de la spécificité et des effets biologiques de la lectine de *Canavalia brasiliensis* (ConBr), des lectines spécifiques pour le manose et le lactose de *Cymboseman roseum* (CRLI e CRLII), de la lectine de *Dioclea rostrata* (DrosL) et de la lectine de *Dioclea wilsonii* (DwiiL). Ces études ont été publiées dans les revues suivantes (plus six travaux qui ont été publiés pendant les stages de license et master).

1. PIRES, A. F., ASSREUY A. M., LOPES, E. A., CELEDÔNIO, N. R., SOARES, C. E., RODRIGUES, N. V., SOUSA, P. L., BENEVIDES, R. G. , NAGANO, C. S., CAVADA, B. S., LEAL-CARDOSO, J. H., COELHO-DE-SOUZA, A. N., SANTOS, C. F. Opioid-like antinociceptive effects of oral administration of a lectin purified from the seeds of <i>Canavalia brasiliensis</i> . <i>Fundam. Clin. Pharmacol.</i> (Print), 2011
2. BEZERRA, E. H. SA., ROCHA, B. A. M., NAGANO, C. S., BEZERRA, G. A., MOURA, T. R., BEZERRA, M. J. B., BENEVIDES, R. G. , SAMPAIO, A. H., DELATORRE, P., CAVADA, B. S. Structural analysis of ConBr reveals molecular correlation between the carbohydrate recognition domain and nitric oxide release from endothelial cells. <i>Biochemical and Biophysical Research Communications</i> , 408 (4): 566 - 570, 2011.
3. DAM, T. K., CAVADA, B. S., NAGANO, C. S., ROCHA, B. A., BENEVIDES, R. G. , NASCIMENTO, K. S., DE SOUSA, L. A., OSCARSON, S., BREWER, C. F. Fine specificities of two lectins from <i>Cymbosema roseum</i> seeds: a lectin specific for high-mannose oligosaccharides and a lectin specific for blood group H type II trisaccharide. <i>Glycobiology</i> . 2011 Jul;21(7):925-33.
4. ROCHA, B. A. M., DELATORRE, P., OLIVEIRA, T. M., BENEVIDES, R. G. , PIRES, A. F., SOUSA, A. A. S., SOUZA, L. A. G., ASSREUY, A. M. S., DEBRAY, H., AZEVEDO, W. F., SAMPAIO, A. H., CAVADA, B. S. Structural basis for both pro- and anti-inflammatory response induced by mannose-specific legume lectin from <i>Cymbosema roseum</i> . <i>Biochimie (Paris)</i> , 93: 806 - 816, 2011.
5. FIGUEIREDO, JOZI G. B., FLÁVIO S., MOTA, M. R. L., SILVESTRE, P. P., AGUIAR, C. N., BENEVIDES, R. G. , NASCIMENTO, K. S., MOURA, T. R., DAL-SECCO, D., ASSREUY, A. M. S., CAVADA, B. S. Pharmacological analysis of the neutrophil migration induced by <i>D. rostrata</i> lectin: Involvement of cytokines and nitric oxide. <i>Toxicon (Oxford)</i> , 54: 736 - 744, 2009.
6. ROCHA, B. A. M., MORENO, F. B. M. B., DELATORRE, P., SOUZA, E. P., MARINHO, E. S., BENEVIDES, R. G. , RUSTIGUEL, J. K. R., SOUZA, L. A. G., NAGANO, C. S., DEBRAY, H., SAMPAIO, A. H., AZEVEDO, W. F., CAVADA, B. S. Purification, Characterization, and Preliminary X-Ray Diffraction Analysis of a Lactose-Specific Lectin from <i>Cymbosema roseum</i> Seeds. <i>Applied Biochemistry and Biotechnology</i> , 152: 383 - 393, 2009.
7. OLIVEIRA, T. M., DELATORRE, P., ROCHA, B. A. M., SOUZA, E. P., NASCIMENTO, K. S., BEZERRA, G. A., MOURA, T. R., BENEVIDES, R. G. , BEZERRA, E. H. S., MORENO, F. B., CAVADA, B. S. Crystal structure of <i>Dioclea rostrata</i> lectin: Insights into understanding the pH-dependent dimer-tetramer equilibrium and the structural basis for carbohydrate recognition in <i>Diocleinae</i> lectins. <i>Journal of Structural Biology</i> , 164: 177 - 182, 2008.
8. BEZERRA, G. A., OLIVEIRA, T. M., MORENO, F. B. M. B., SOUZA, E. P., ROCHA, B. A. M., BENEVIDES, R. G. , DELATORRE, P., AZEVEDO, W. F., CAVADA, B. S. Structural analysis of <i>Canavalia maritima</i> and <i>Canavalia gladiata</i> lectins complexed with different dimannosides: new insights into the understanding of the structure-biological activity relationship in legume lectins.. <i>Journal of Structural Biology</i> . , 160: 168 - 176, 2007.

9. ALENCAR, N. M., ASSREUY, A. M., HAVT, A., **BENEVIDES, R. G.**, MOURA, T. R., SOUSA, R. B., RIBEIRO, R. A., CUNHA, F. Q., CAVADA, B. S. Vatairea macrocarpa (Leguminosae) lectin activates cultured macrophages to release chemotactic mediators. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 374: 275 - 282, 2007.
10. DELATORRE, P., NASCIMENTO, K. S., MELO, L. M., SOUZA, E. P., ROCHA, B. A. M., **BENEVIDES, R. G.**, OLIVEIRA, T. M., BEZERRA, G. A., BEZERRA, M. J. B., CUNHA, R. M. S., CUNHA, F. A. B., FREIRE, V. N., CAVADA, B. S. Crystallization and preliminary X-ray diffraction analysis of the lectin from Dioclea rostrata Benth seeds. *Acta Crystallographica. Series F*, 62: 166/62 - 168, 2006.
11. MOTA, M. R., CRIDDLE, D. N., ALENCAR, N. M., GOMES, R. C., MEIRELES, A. V., SANTI-GADELHA, T., GADELHA, C. A., OLIVEIRA, C. C., **BENEVIDES, R. G.**, CAVADA, B. S., ASSREUY, A. M. Modulation of acute inflammation by a chitin-binding lectin from Araucaria angustifolia seeds via mast cells. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 374: 1 - 10, 2006.
12. MORENO, F. B. M. B., BEZERRA, G. A., OLIVEIRA, T. M., SOUZA, E. P., ROCHA, B. A. M., **BENEVIDES, R. G.**, DELATORRE, P., CAVADA, B. S., AZEVEDO, W. F. New crystal forms of Diocleinae lectins in the presence of different dimannosides. *Acta Crystallographica. Series F*, 62: 1100 - 1103, 2006.
13. CAVADA, B. S., MARINHO, E. S., SOUZA, E. P., **BENEVIDES, R. G.**, DELATORRE, P., SOUZA, L. A. G., NASCIMENTO, K. S., SAMPAIO, A. H., MORENO, F. B. M. B., RUSTIGUEL, J. K. R., CANDURI, F., AZEVEDO, W. F., DEBRAY, H. Purification, partial characterization and preliminary X-ray diffraction analysis of a mannose-specific lectin from Cymbosema roseum seeds. *Acta Crystallographica. Section F*, 62: 235 - 237, 2006

Au cours du stage en co-tutelle à l'étranger, un autre projet majeur a également été entrepris au CERMAV, portant sur l'étude structurale et l'analyse de la spécificité de la lectine B de *Burkholderia cenocepacia* (BclB), impliquant la production et l'analyse des deux domaines de cette lectine. *B. cepacia* est un ensemble de bactéries à gram négatif qui comprend plusieurs espèces très similaires sur le plan phénotypique, et qui sont responsables d'épidémies, en particulier chez les patients atteints de la mucoviscidose. Les glycoconjugués ont un rôle important dans les phénomènes de reconnaissance moléculaire hôte-pathogène. L'arsenal que les organismes pathogènes utilisent pour reconnaître, adhérer et infecter leur hôte comprend divers types de protéines, dont un certain nombre de lectines et adhésines. Certaines bactéries du complexe BCC possèdent dans leur génome quatre séquences homologues au gène *lecA* codant pour la lectine PA-III chez *Pseudomonas aeruginosa*. L'analyse de la séquence protéique traduite du génome de *B. cenocepacia* J2314 montre que la protéine BclB 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-III. Un travail de thèse effectué avant cette date (Emilie Lameignère), décrit le clonage du domaine C-terminal et décrit sa spécificité pour le fucose pour les oligosaccharides fucosylés Lewis a, Lewis X, et H type. Dans ce contexte, le domaine BclB C-terminal a été cristallisé en présence des ligands Lewis a, H type 1 et 2. Les données de diffraction des rayons X ont été mesurées pour les cristaux en complexes avec les trisaccharides H type 1 et 2, et sont actuellement en train d'être traitées et analysées. Le domaine N-terminal a également été cloné et exprimé et la cristallisation

a pu être possible après optimisation de la taille du gène. En l'absence de similarité avec des protéines connues, le remplacement moléculaire ne peut être utilisé et la protéine a donc été exprimée dans un milieu riche en séléniométhionine. La structure cristallographique a pu être ainsi résolue. Ce domaine semble constituer un motif similaire à celui présent dans la chaîne légère des complexes dynéine, une protéine motrice des microtubules, pili ou flagelles. Une publication est en cours de rédaction.

V. Références Bibliographiques

ADAR R., STREICHER H., ROZENBLATT S., SHARON N. Synthesis of soybean agglutinin in bacterial and mammalian cells. *Eur J Biochem.* 249 (3), 684-689, 1997.

ALENCAR, N. M., TEIXEIRA, E. H., ASSREUY, A. M. S., CAVADA, B. S., FLORES, C. A., RIBEIRO, R. A. Leguminous lectins as tools for studying the role of sugar residues in leukocyte recruitment. *Mediators. Inflamm.* 8, 107–13, 1999.

AMBROSI, M., CAMERON, N. R., AND DAVIS, B. G. Lectins: tools for the molecular understanding of the glycocode. *Org. Biomol. Chem.*, 3, 1593-1608, 2005.

AMBROSI, M., CAMERON, N. R., AND DAVIS, B. G. Lectins: tools for the molecular understanding of the glycocode. *Org. Biomol. Chem.*, 3, 1593-1608, 2005.

ARANGO, R., ADAR, R., ROZENBLATT, S., SHARON, N. Expression of Erythrina corallodendron lectin in Escherichia coli. *Eur. J. Biochem.* 205, 575–581, 1992.

ARAÚJO-FILHO, J. H., VASCONCELOS, I. M., MARTINS-MIRANDA, A. S., GONDIM, D. M., OLIVEIRA, J. T. A ConA-like lectin from Dioclea guianensis Benth. has antifungal activity against Colletotrichum gloeosporioides, unlike its homologues, ConM and ConA. *J Agric Food Chem.* 58(7), 4090-4096, 2010.

ASHWELL, G., AND MORELL, A.G. The role of surface carbohydrates in the hepatic recognition and transport of circulating glycoproteins. *Adv. Enzymol.*, 41, 99-128, 1974.

ASSREUY, A. M., FONTENELE, S. R., PIRES, A. F., FERNANDES, D. C., RODRIGUES, N. V., BEZERRA, E. H., MOURA. T. R., DO NASCIMENTO, K. S., CAVADA, B. S. Vasodilator effects of Diocleinae lectins from the Canavalia genus. *Naunyn Schmiedebergs Arch Pharmacol.* 380 (6), 509-21, 2009.

BANERJEE, R., DAS, K., RAVISHANKAR, R., SUGUNA, K., SUROLIA, A., VIJAYAN, M. Conformation, Protein-Carbohydrate Interactions and a Novel Subunit Association in the Refined Structure of Peanut Lectin-Lactose Complex. *J. Mol. Biol.*, 259, 281-296, 1996.

BARAUNA, S. C. KASTER, M. P, HECKERT, B. T., DO NASCIMENTO, K. S., ROSSI, F. M., TEIXEIRA, E. H., CAVADA, B. S., RODRIGUES, A. L., LEAL, R. B. Antidepressant-like effect of lectin from Canavalia brasiliensis (ConBr) administered centrally in mice. *Pharmacol. Biochem. Behavior*, 85, 160-169, 2006.

BARBOSA, T., ARRUDA, S., CAVADA, B., GRANGEIRO, T.B., FREITAS, L.A.R., BARRALNETTO, M. *In Vivo* Lymphocyte Activation and Apoptosis by lectins of the Diocleinae Subtribe. *Memórias do Instituto Oswaldo Cruz.* 96(5), 673-678, 2001.

BENEVIDES, R.G. Avaliação do Potencial Fungicida e Termiticida de uma Fração Protéica Lectínica de Sementes de Platypodium elegans Vogel e Obtenção da Lectina Purificada. Dissertação (Mestrado em Bioquímica) – Centro de Ciências, Universidade Federal do Ceará, Fortaleza, 151 p., 2008.

BEZERRA, D. P.; LEITE, K. B.; NUNES, E. P.; MARINHO, E. S.; CAVADA, B. S. Purificação de uma Lectina Presente em Sementes de *Machaerium acutifolium*. *Anais da XXVI Reunião Nordestina de Botânica*, In: XXVI Reunião Nordestina de Botânica, Fortaleza. 2003.

BEZERRA, W. M.; CARVALHO, C. P. S.; MOREIRA, R. A.; GRANGEIRO, T. B. Establishment of a heterologous system for the expression of *Canavalia brasiliensis* lectin: a model for the study of protein splicing. *Genetics and Molecular Research*, 5(1), 216-223, 2006.

BONNEIL, E.; YOUNG, N. M.; LIS, H.; SHARON, N.; THIBAUT, P. Probing genetic variation and glycoform distribution in lectins of the *Erythrina* genus by mass spectrometry. *Archives of Biochemistry and Biophysics*, 426, 241-249, 2004.

BOUCKAERT J, LORIS R, POORTMANS F, WYNS L. Crystallographic structure of metal-free concanavalin A at 2.5 Å resolution. *Proteins Struct Funct Genet.*, 23, 510-524, 1995.

BOUCKAERT, J., DEWALLEF, Y., POORTMANS, F., WYNS, L., LORIS, R. The structural features of concanavalin A governing non-proline peptide isomerization. *Journal of Biological Chemistry*, 275, 19778–19787, 2000.

BOURNE, Y., ROUSSEL, A., FREY, M., ROUGE, P., FONTECILLA-CAMPS, J. C. & CABBILLAU, C. Three-dimensional structures of complexes of *Lathyrus ochrus* isolectin I with glucose and mannose: fine specificity of the monosaccharide-binding site. *Proteins: Struct. Funct. Genet.* 8, 365–376, 1990.

BRANCO, A. T.; BERNABÉ, R. B.; FERREIRA, B. S.; OLIVEIRA, M. V. V.; GARCIA, A. B.; SOUZA-FILHO, G. A. Expression and purification recombinant lectin from rice (*Oryza sativa* L.). *Protein Expression & Purification*, 33, 34-38, 2004.

BUTS, L., GARCIA-PINO, A., IMBERTY, A., AMIOT, N., BOONS, G. J., BEECKMANS, S., VERSÉES, W., WYNS, L., LORIS, R. Structural basis for the recognition of complex-type biantennary oligosaccharides by *Pterocarpus angolensis* lectin. *FEBS J.*, 273, 2407–2420, 2006.

CAVADA, B. S.; SANTOS, C. F.; GRANGEIRO, T. B.; NUNES, E. P.; SALES, P. V. P.; RAMOS, R. L. Purification and characterization of a lectin from seeds of *Vatairea macrocarpa* Ducke. *Phytochemistry*, 49, 675-680, 1998.

CHAO, Q.; CASALONGUE, C., QUINN, J.M.; ETZLER, M.E. Expression and partial characterization of *Dolichos biflorus* seed lectin in *Escherichia coli*. *Archives of Biochemistry and Biophysics*, 313, 346-350, 1994.

COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4. The CCP4 suite: programs for protein crystallography. *Acta Cryst D.*, 50(4), 760-763, 1994.

DAM, T. K., AND BREWER, C. F. Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.*, 102, 387-429, 2002.

DELANO, W.L. The PyMOL Molecular Graphics System. *DeLano Scientific*, San Carlos, CA, USA, 2002.

DELATORRE, P., ROCHA, B. A. M., SOUZA, E. P., OLIVEIRA, T.M., BEZERRA, G.A., MORENO, F.B., FREITAS, B.T., SANTI-GADELHA, T., SAMPAIO, A. H., AZEVEDO JR., W. F., CAVADA, B. S. Structure of a lectin from *Canavalia gladiata* seeds: new structural insights for old molecules, *BMC Struct. Biol.*, 7, 52, 2007.

- DELBAERE, L. T., VANDONSELAAR, M., PRASAD, L., QUAIL, J. W., WILSON, K. S. & DAUTER, Z. Structures of the lectin IV of *Griffonia simplicifolia* and its complex with the Lewis b human blood group determinant at 2.0 Å resolution. *J. Mol. Biol.* 230, 950–965, 1993.
- DEREWENDA, Z., YARIV, J., HELLIWELL, J. R., KALB, A. J., DODSON, E. J., PAPIZ, M. Z. The structure of the saccharide-binding site of concanavalin A. *EMBO J.* 8, 2189–2193, 1989.
- ELGAVISH, S., SHAANAN, B., Structures of the *Erythrina corallodendron* lectin and of its complexes with mono- and disaccharides. *Journal of Molecular Biology*, 277(4), 917-932, 1998.
- EMSLEY, P., AND COWTAN, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, 60, 2126-2132, 2004.
- EVANS, P.R. "Data reduction", Proceedings of CCP4 Study Weekend, *Data Collection & Processing*, 114-122, 1993.
- FAN, E., ZHANG, Z., MINKE, W. E., HOU, Z., VERLINDE C. L. M. J., HOL, W. G. J. High-affinity pentavalent ligands of *Escherichia coli* heat-labile enterotoxin by modular structure-based design. *J. Am. Chem. Soc.*, 122, 2663-2664, 2000.
- FIGUEIREDO, J. G., BITENCOURT, F., BESERRA, I. G., TEIXEIRA, C. S., LUZ, P. B., BEZERRA, E. H., MOTA, M. R., ASSREUY, A. M., DE QUEIROZ CUNHA, F., CAVADA, B. S., DE ALENCAR, N. M. Antinociceptive activity and toxicology of the lectin from *Canavalia boliviana* seeds in mice. *Naunyn Schmiedebergs Arch Pharmacol.* 380(5), 407-414, 2009.
- GADELHA, C. A., MORENO, F. B., SANTI-GADELHA, T., CAJAZEIRAS, J. B., ROCHA, B. A., ASSREUY, A. M. Native crystal structure of a nitric oxide-releasing lectin from the seeds of *Canavalia maritima*. *Journal Structural Biology*, 152, 185–194, 2005.
- GARCIA-PINO, A., BUTS, L., WYNS, L., IMBERTY, A., LORIS, R. How a plant lectin recognizes high mannose oligosaccharides. *Plant Physiology.* 144(4),1733-1741, 2007.
- GOUET, P., COURCELLE, E., STUART, D.I. AND METOZ, F. ESPript: multiple sequence alignments in PostScript. *Bioinformatics.* 15, 305-308, 1999.
- HAMELRYCK, T. W., MOORE, J. G., CHRISPEELS, M. J., LORIS, R. & WYNS, L. The role of weak protein–protein interactions in multivalent lectin– carbohydrate binding: crystal structure of cross- linked FRIL. *J. Mol. Biol.* 299, 875–883, 2000.
- HAMELRYCK, T.W., LORIS, R., BOUCKAERT, J., WYNS, L. Properties and structure of the legume lectin family. *Trends in Glycoscience and Glycobiology*, 10, 349-404, 1998.
- HARAGUCHI, T., NOMURA K., YAGI, F. “Cloning and Expression of a Mannose-Binding Jacalin-Related Lectin from Leaves of Japanese Cycad (*Cycas revoluta* Thunb.)”, *Biosci. Biotechnol. Biochem.*, 70, 2222-2229, 2006.
- HERVÉ C, SERRES J, DABOS P, CANUT H, BARRE A, ROUGÉ P, LESCURE B. Characterization of the *Arabidopsis* lecRK-a genes: members of a superfamily encoding putative receptors with an extracellular domain homologous to legume lectins. *Plant Mol Biol.*, 39(4), 671-82, 1999.
- HOFFMAN, L. M., DONALDSON; D. D. Synthesis of mitogenic phytohemagglutinin-L in *Escherichia coli*. *Biotechnology.* 5: 157-160, 1987.

- IMBERTY, A., GAUTIER, C., LESCAR, J., PEREZ, S., WYNS, L. & LORIS, R. An unusual carbohydrate binding site revealed by the structures of two *Maackia amurensis* lectins complexed with sialic acid-containing oligosaccharides. *J. Biol. Chem.* 275, 17541–17548, 2000.
- JANAKIRAMAN, M. N., WHITE, C. L., LAVER, W. G., AIR G. M. AND LUO, M. Structure of influenza virus neuraminidase B/Lee/40 complexed with sialic acid and a dehydro analog at 1.8-Å resolution: implications for the catalytic mechanism. *Biochemistry*, 33, 8172-8179, 1994.
- JORDAN, E. T., GOLDSTEIN, I. J.. The sequence of a second member of the lima bean lectin gene family and the expression and characterization of recombinant lectin in *Escherichia coli*. *J. Biol. Chem.*, 269(10), 7674-7681, 1994.
- JOUBERT, F. J., SHARON, N., MERRIFIELD, E. H. Purification and properties of a lectin from *Lonchocarpus capassa* (apple-leaf) seed. *Phytochemistry*, 25(2), 323-327, 1986.
- KIMBLE, B.; RAMIREZ NIETO, G.; PEREZ, D. R. Characterization of influenza virus sialic acid receptors in minor poultry species. *Virol. J.*, 7, 365, 2010.
- KITADA, M.; KURODA, Y.; DEZAWA, M. Lectins as a Tool for Detecting Neural Stem/ Progenitor Cells in the Adult Mouse Brain. *Anat. Rec. (Hoboken)*, 1, 1-19, 2010.
- KITOV, P. I., SADOWSKA, J. M., MULVEY, G., ARMSTRONG, G. D., LING, H., PANNUS, N. S., READ, R. J. AND BUNDLE, D. R., Shiga-like toxins neutralized by tailored multivalent carbohydrate ligands. *Nature*, 403, 669-672, 2000.
- KOCOUREK, J., AND HOREJSI, V. Defining a lectin. *Nature*, 290, 188, 1981.
- KUSUI, K.; YAMAMOTO, K.; KONAMI, Y.; OSAWA, T. cDNA cloning and expression of *Bauhinia purpurea* lectin. *Journal Biochemistry*. 109, 899-903, 1991.
- LAEMMLI, U. K., Cleavage of structural proteun during the assembly of the bactetiophage t4, *Nature*, 227, 680-685, 1970.
- LESLIE, A. G. W., Recent changes to the MOSFLM package for processing film and image plate data, *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, 26, 1992.
- LIS, H., AND SHARON, N. Lectins: Carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.*, 98, 637-674, 1998.
- LORIS, R., DE GREVE, H., DAO-THI, M.-H., MESSENS, J., IMBERTY, A. & WYNS, L. Structural basis of carbohydrate recognition by lectin II from *Ulex europaeus*, a protein with a promiscuous carbohydrate-binding site. *J. Mol. Biol.* 301, 987–1002, 2000.
- LORIS, R., MAES, D., POORTMANS, F., WYNS, L. BOUCKAERT, J. A structure of the complex between concanavalin A and methyl-3,6-di-O-(alpha-D-mannopyranosyl)-alpha-D-mannopyranoside reveals two binding modes. *J. Biol. Chem.* 271, 30614–30618, 1996.
- LORIS, R., IMBERTY, A., BEECKMANS, S., VAN DRIESSCHE, E., READ, J. S., BOUCKAERT, J. Crystal structure of *Pterocarpus angolensis* lectin in complex with glucose, sucrose and turanose. *J. Biol. Chem.*, 278, 16297–16303, 2003.

- LORIS, R.; HAMELRYCK, T.; BOUCKAERT, J.; WYNS, L. Legume lectin structure. *Biochimica et Biophysica Acta*, 1383, 9-36, 1998.
- LORIS, R.; VAN WALLE, I.; DE GREVE, H.; BEECKMANS, S.; DEBOECK, F.; WYNS, L.; BOUCKAERT, J. Structural basis of oligomannose recognition by the *Pterocarpus angolensis* seed lectin. *Journal of Molecular Biology*, 335, 1227. 2004.
- LUO, S.; ZHANGSUN, D.; TANG, K. Functional GNA expressed in *Escherichia coli* with high efficiency and its effect on *Ceratovacuna lanigera* Zehntner. *Applied Genetics and Molecular Biotechnology*. 69, 184-191, 2005.
- LUTOMSKI, D.; IMAM-SGHIOUAR, N.; BLONDEAU, K.; CARON, M.; JOUBERTCARON, R. Purification of human galectin-1 produced in high-cell density cultures of recombinant *Escherichia coli*: a comparison with classic shake flask cultivation. *Journal of Chromatography B*, 808, 105-109, 2004.
- MAMMEN, M., CHOI S.-K. AND WHITESIDES, G.M. Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem., Int. Ed. Engl.*, 37, 2754-2794, 1998.
- MANDAL, D. K., KISHORE, N. & BREWER. C. F. Thermodynamics of lectin-carbohydrate interactions. Titration microcalorimetry! Measurements of the binding of N-linked carbohydrates and oval buniin to concanavalin A, *Biochemistry*, 33, 1149- 1156, 1994.
- MANN, K., FARIAS, C. M. S. A., GALLEGU DEL SOL, F., SANTOS, C. F., GRANGEIRO, T. B., NAGANO, C. S., CAVADA, B. S., CALVETE, J. J. The amino-acid sequence of the glucose/ mannose-specific lectin isolated from *Parkia platycephala* seeds reveal three tandemly arranged jacalin related domains. *Eur. J. Biochem.*, 218, 4414-4422, 2001.
- MANOJ, N., SUGUNA, K. Signature of quaternary structure in the sequences of legume lectins. *Protein Eng.* 14, 735-745, 2001.
- MARONDEDZE C., CHIDZWONDO F., SITHOLE-NIANG I., READ J.S. *Pterocarpus rotundifolius* lectin gene. Submitted to the EMBL/GenBank/DDBJ databases, 2004.
- MATTHEWS, B.W. Solvent content of protein crystals. *J. Mol. Biol.*, 33, 491-497, 1968.
- MCCOY, A. J., STORONI, L. C. & READ, R. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Cryst. D*, 60, 1220-1228, 2004.
- MOLEMA G. AND MEIJER, D. K. F., Targeting of drugs to various blood cell types using (neo)glycoproteins, antibodies and other protein carriers. *Adv. Drug Delivery Rev.*, 14(1), 25-50, 1994.
- MOOTHOO, D. N., CANAN, B., FIELD, R. A. & NAISMITH, J. H. Man(α 1-2)Man α -OMe-concanavalin A complex reveals a balance of forces involved in carbohydrate recognition. *Glycobiology*, 9, 539-545, 1999.
- MURSHUDOV, G. N.; VAGIN, A. A; DODSON, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 1 (53), 240-55, 1997.
- NAKAMURA, S., IKEGAMI, A., MATSUMURA, Y., NAKANISHI, T., NOMURA, K. Molecular Cloning and Expression of the Mannose/Glucose Specific Lectin from *Castanea crenata* Cotyledons1. *J. Biochem.*, 131, 241-246, 2002.

NASCIMENTO, K. S.; GRANGEIRO, T. B.; CAVADA, B. S.; SAMPAIO, A. H. Composição mínima, teores proteicos e detecção de atividade hemaglutinante em sementes de *Andira retusa* (Leguminosae, Papilionoidae, Dalbergieae); *Resumos do 53o Congresso Nacional de Botânica*; 53o Congresso Nacional de Botânica; Recife; BRASIL, 2002.

NELSON, D.L. & COX, M.M. Principles of Biochemistry, Fifth Edition (*Freeman Publishers*), New York, 2009.

NISHIGUCHI, M., YOSHIDA, K., SUMIZONO, T., TAZAKI, K. Studies by site-directed mutagenesis of the carbohydrate-binding properties of a bark lectin from *Robinia pseudoacacia*. *FEBS Lett.* 403(3), 294-298, 1997.

NOGUEIRA, N. A.; GRANGEIRO, M. B.; CUNHA, R. M.; RAMOS, M. V.; ALVES, M. A.; TEIXEIRA, E. H.; BARRAL-NETTO, M.; CALVETE, J. J.; CAVADA B. S.; GRANGEIRO, T. B. Expression and purification of the recombinant ConBr (*Canavalia brasiliensis* lectin) produced in *Escherichia coli*. *Protein Pet Lett.*, 9(1), 59-66, 2002.

NUNES, B. S., RENSONNET, N. S., DAL-SECCO, D., VIEIRA, S. M., CAVADA, B. S., TEIXEIRA, E. H., MOURA, T. R., TEIXEIRA, C. S., CLEMENTE-NAPIMOGA, J. T., CUNHA, F. Q., NAPIMOGA, M. H. Lectin extracted from *Canavalia grandiflora* seeds presents potential anti-inflammatory and analgesic effects. *Naunyn Schmiedebergs Arch Pharmacol.*, 379(6), 609-616, 2009.

OLIVEIRA, C., COSTA, S., TEIXEIRA, J. A., DOMINGUES, L. cDNA Cloning and Functional Expression of the α -d-Galactose-Binding Lectin Frutalin in *Escherichia coli*. *Molecular Biotechnology.* 43 (3), 212-220, 2009.

PEREIRA-JUNIOR, F.N. Purificação e Caracterização Parcial de uma Lectina Isolada de Sementes de *Platymiscium floribundum* Vogel. *Monografia de Graduação*, URCA, Crato, 45 p., 2008.

PEUMANS, W. J. AND VAN DAMME, E. J. M., Lectins as plant defense proteins. *Plant Physiol.*, 109, 347-352, 1995.

PIRES, A. F.; RODRIGUES, N. V. F. C.; SIQUEIRA, R. C. L.; SILVA, M. S. J.; MARINHO, M. M.; CAVADA, B. S.; ALENCAR, N. M. N.; ASSREUY, A. M. S. Atividade Antiinflamatória Da Lectina De Sementes De *Lonchocarpus araripensis*. 3ª Reunião Regional *FESBE*, Fortaleza, CE, 2008.

PRASTHOFER, T., PHILLIPS, S. R., SUDDATH, F. L., ENGLER, J. A. Design, expression, and crystallization of recombinant lectin from the garden pea (*Pisum sativum*). *J. Biol. Chem.*, 264 (12), 6793-6796, 1989.

PUSTZAI, A. *Plant Lectins*. *Cambridge University Press*. Cambridge, 1991.

RAEMAEKERS, J. M. R.; MURO, L.; GATEHOUSE, J. A.; FORDHAM-SKELTON, A. P. Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris*. Correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. *European Journal of Biochemistry*, 265, 394-403, 1999.

RANGEL, T. B. A., CARVALHO, A. U., NASCIMENTO, K. S., CAVADA, B. S., FERNANDES, P. M. B. Isolamento e caracterização parcial de lectina proveniente de sementes da espécie vegetal *Andira fraxinifolia* (LEGUMINOSAE-FABACEAE). *Anais/Resumos da 61ª Reunião Anual da SBPC-ISSN 2176-1221*, 2009.

- RINI, J. M., HARDMAN, K. D., EINSPAHR, H., SUDDATH, F. L. & CARVER, J. P. X-ray crystal structure of a pea lectin-trimannoside complex at 2.6 Å resolution. *J. Biol. Chem.*, 268, 10126–10132, 1993.
- RIVET, A. MAZEAU, K. IMBERTY, A. BETTLER, E. PEREZ, S. 3D lectin database Available from <http://www.cermav.cnrs.fr/lectines/>.
- ROZWARSKI, D. A., SWAMI, B. M., BREWER, C. F. & SACCHETTINI, J. C. Crystal structure of the lectin from *Dioclea grandiflora* complexed with core trimannoside of asparagine-linked carbohydrates. *J. Biol. Chem.* 273, 32818–32825, 1998.
- SANTOS, A. F., CAVADA, B. S., ROCHA, B. A., NASCIMENTO, K. S., SANT'ANA, A. E. Toxicity of some glucose/mannose-binding lectins to *Biomphalaria glabrata* and *Artemia salina*. *Bioresour Technol.* 101 (2), 794-798, 2010.
- SANZ-APARICIO, J., HERMOSO, J., GRANGEIRO, T. B., CALVETE, J. J., CAVADA, B. S. The crystal structure of *Canavalia brasiliensis* lectin suggests a correlation between its quaternary conformation and its distinct biological properties from concanavalin A. *FEBS Letters*, 405, 114–118, 1997.
- SHARMA V., SUROLIA A. Cloning by genomic PCR and production of peanut agglutinin in *Escherichia coli*. *Gene*, 148, 299-304, 1994.
- SHARMA, V., SUROLIA, A. Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J Mol Biol*, 267, 433–445, 1997.
- SHARON, N. & LIS, H. Legume Lectins – A large family of homologous proteins. *In FASEB Journal*, 4, 3198-3208, 1990.
- SHARON, N., AND LIS, H. Lectin's, 2nd Ed., *Kluwer Academic Publishers*, Dordrecht, The Netherlands, 2003.
- SHARON, N.; LIS, H. How proteins bind carbohydrates: lessons from legume lectins. *Journal of Agriculture and Food Chemistry*, 50, 6586-6591, 2002.
- SRINIVASAN, N., RUFINO, S. D., PEPYS, M. B., WOOD, S. P., BLUNDELL, T. L. A superfamily of proteins with the lectin fold. *Chemrracrs-Biochem Mol Biol.*, 6, 149-164, 1996.
- SRINIVAS, V. R., REDDY, G. B., AHMAD, N., SWAMINATHAN, C. P., MITRA, N. & SUROLIA, A. Legume lectin family, the “natural mutants of the quaternary state”, provide insights into the relationship between protein stability and oligomerization. *Biochim. Biophys. Acta*, 1527, 102–111, 2001.
- STANCOMBE, P. R.; ALEXANDER, F. C. G.; LING, R.; MATHENSON, M. A.; CHONE, C. C.; CHADDOCK, J. A. Isolation of the gene and large-scale expression and purification of recombinant *Erythrina cristagalli* lectin. *Protein expression and purification.* 30, 283-292. 2003.
- STREICHER, H. R. G.; SHARON, N. Recombinant Plant Lectins and Their Mutants. *Methods in enzymology.* 363, 47-77, 2003.
- STUBBS, M. E., CARVER, J. P., DUNN, R. J. Production of pea lectin in *Escherichia coli*. *J. Biol. Chem.*, 261 (14), 1986.

- TAYLOR, M., DRICKAMER, K. Introduction to Glycobiology. *Oxford University Press*, 2003.
- TEIXEIRA, E. H.; NAPIMOGA, M. H.; CARNEIRO, V. A.; DE OLIVEIRA, T. M.; CUNHA, R. M.; HAVT, A.; MARTINS, J. L.; PINTO, V. P.; GONÇALVES, R. B.; CAVADA, B. S. In vitro inhibition of Streptococci binding to enamel acquired pellicle by plant lectins. *J Appl Microbiol.*, 101, 111-116, 2006.
- TEIXEIRA, E. H.; NAPIMOGA, M. H.; CARNEIRO, V. A.; DE OLIVEIRA, T. M.; NASCIMENTO, K. S.; NAGANO, C. S.; SOUZA, J. B.; HAVT, A.; PINTO, V. P.; GONÇALVES, R. B.; FARIAS, W. R.; SAKER-SAMPAIO, S.; SAMPAIO, A. H.; CAVADA, B. S. In vitro inhibition of oral streptococci binding to the acquired pellicle by algal lectins. *J. Appl. Microbiol.*, 103, 1001-1006, 2007.
- TOONE, E. J. Structure and Energetics of Protein-Carbohydrate Complexes, edited by Bundle, D.R., Fiezi, T., *Current Opinion in Structural Biology*, 4, 719-728, 1994.
- UPADHYAY, S. K., SAURABH, S., RAI, P., SINGH R., CHANDRASHEKAR, K., VERMA, P. C., SINGH, P. K., TULI. R. SUMO fusion facilitates expression and purification of garlic leaf lectin but modifies some of its properties. *J Biotechnol.*, 146, 1-8, 2010.
- VAN DAMME, E. J. M., PEUMANS, W. J., BARRE, A. & ROUGÉ, P. Plant lectin: A composite of several distinct families of structural and evolutionary related proteins with diverse biological roles. *Crit. Rev. Plant Sciences*, 17(6), 575-692, 1998.
- VAREJÃO, N., ALMEIDA, M. D. A. S., DE CICCO, N. N., ATELLA, G. C., COELHO, L. C., CORREIA, M. A., FOGUEL, D. Heterologous expression and purification of a biologically active legume lectin from *Cratylia mollis* seeds (CRAMOLL 1). *Biochim Biophys Acta*. 1804(9), 1917-1924, 2010.
- VARROT, A., BLANCHARD, B. AND IMBERTY, A. Lectin Binding and its Structural Basis, in Carbohydrate Recognition: Biological Problems, Methods, and Applications (eds B. Wang and G.-J. Boons), *John Wiley & Sons, Inc.*, Hoboken, NJ, USA, 2011.
- WAH, D. A., ROMERO, A., GALLEGRO DEL SOL, F., CAVADA, B. S., RAMOS, M. V., GRANGEIRO, T. B. Crystal structure of native and Cd/Cd-substituted Dioclea guianensis seed lectin. A novel manganese-binding site and structural basis of dimer-tetramer association. *J. Mol. Biol.*, 310, 885-894, 2001.
- WALLACE, A. C., LASKOWSKI, R. A., THORNTON, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.*, 8, 127-134, 1995.
- WANG, L. C.; KANG, L.; HU, T. M.; WANG, J. L. Abrin-a A chain expressed as soluble form in *Escherichia coli* from a PCR-synthesized gene is catalytically and functionally active. *Biochimie*, 86, 327-333, 2004.
- WEIS, W. I., DRICKAMER, K. Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.*, 65, 441-473, 1996.
- YAMAUCHI, D. & MINAMIKAWA, T. Structure of the gene encoding concanavalin A from *Canavalia gladiata* and its expression in *Escherichia coli* cells. *FEBS Letters*, 260(1), 127-130, 1990.
- YIM, M.; ONO, T.; IRIMURA, T. Mutated plant lectin library useful to identify different cells. *Proc. Natl. Acad. Sci.*, 98, 2222-2225, 2001.

ZHU, K., HUESING, J. B., SHADE, R. E., BRESSAN, R. A., HASEGAWA, P. M., MURDOCK, L. L.
An insecticidal N-acetylglucosamine-specific lectin gene from *Griffonia simplicifolia* (Leguminosae).
Plant Physiol., 110, 195–202, 1996.

ANNEXES

ANNEXE 1:

Caracterização bioquímica e estrutural de uma lectina recombinante de sementes de *Platypodium elegans* Vogel

(Travail de thèse complète en portugais / Tese completa em língua portuguesa)



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR

**Caracterização bioquímica e estrutural de uma lectina
recombinante de sementes de *Platypodium elegans*
Vogel**

Raquel Guimarães Benevides

Orientadores:

Benildo Sousa Cavada (UFC, Brasil)

Anne Imberty (CNRS-UJF, França)

**Fortaleza
Dezembro – 2011**



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR

Caracterização bioquímica e estrutural de uma lectina recombinante de sementes de *Platypodium elegans* Vogel

Tese em regime co-tutela submetida à
Coordenação do Curso de Pós-Graduação em
Bioquímica como requisito para obtenção do
grau de Doutor em Bioquímica

Raquel Guimarães Benevides

Orientadores:

Benildo Sousa Cavada (UFC, Brasil)

Anne Imberty (CNRS-UJF, França)

**Fortaleza
Dezembro – 2011**

DEFESA DE TESE:

**Caracterização bioquímica e estrutural de uma lectina
recombinante de sementes de *Platypodium elegans* Vogel**

Tese submetida à Coordenação do Curso de Pós-Graduação em Bioquímica como requisito para
obtenção do grau de Doutor em Bioquímica

Raquel Guimarães Benevides
Bacharel e Licenciada em Biologia, Mestre em Bioquímica

BANCA EXAMINADORA:

Professor Doutor Benildo Sousa Cavada (Orientador)
Departamento de Bioquímica e Biologia Molecular
Universidade Federal do Ceará (UFC)

Dr. Anne Imberty (Orientadora)
Centre d'Étude et de Recherche de Macromolécules Végétales (CERMAV)
Université Joseph Fourier (CNRS-UJF)

Professor Doutor Alexandre Holanda Sampaio (Examinador)
Departamento de Engenharia de Pesca
Universidade Federal do Ceará (UFC)

Dr. Annabelle Varrot (Examinadora)
Centre d'Étude et de Recherche de Macromolécules Végétales (CERMAV)
Université Joseph Fourier (CNRS-UJF)

Professor Doutor Plínio Delatorre (Relator)
Departamento de Biologia Celular e Molecular
Universidade Federal da Paraíba (UFPB)

Dr. Isabelle André (Relatora)
Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés (LISBP)
(CNRS-INSA)

Fortaleza

Dezembro - 2011

A Deus,
A meu pai Ricardo (*in memoriam*)
A minha mãe Lindóia,
A minhas irmãs Rafaela e Rebeca,
A meu amado marido Elton,

Dedico

AGRADECIMENTOS...

De modo bem especial aos meus dois orientadores, Professor Doutor Benildo Sousa Cavada e Doutora Anne Imberty, pela maneira como me receberam, pela orientação e ensinamentos, não só para a conclusão deste trabalho, como para minha vida acadêmica. Sua muita dedicação, compreensão, amizade e oportunidade que me proporcionaram trilhar novos caminhos jamais poderão cessar os meus eternos agradecimentos. Obrigada por acreditarem em mim!

Agradeço aos doutores Plínio Delatorre e Isabelle André por terem aceitado julgar esse trabalho de tese tão prontamente. Meu obrigado também ao Professor Doutor Alexandre Holanda Sampaio por aceitar participar dessa banca de tese.

Agradeço à Dra. Anabelle Varrot, pela sua co-orientação e suporte em boa parte da tese, pela sua atenção e paciência, e por aceitar compor a banca dessa tese. Gostaria de agradecer também todo o apoio da Profa. Dra. Christelle Breton, por seus ensinamentos e orientações sempre que necessários. Agradeço também ao apoio, ensinamentos e discussões com o Prof Dr. Celso Shiniti Nagano.

Gostaria de agradecer à coordenação do Departamento de Bioquímica e Biologia Molecular (UFC), atualmente representada pelo Prof. Dr. Márcio Viana Ramos, pela recepção e oportunidade, assim como à direção do CERMAV (UJF), representado pelo Dr. Redoane Borsali, pela calorosa recepção. Agradeço igualmente a todos os professores, membros e funcionários de ambas as instituições pela grande oportunidade de desfrutar momentos de ensinamentos engrandecedores, científicos e de vida, abrindo-me novos horizontes.

Aos grandes amigos do BioMol-Lab, grandes incentivadores e apoiadores do meu trabalho. Sem vocês, sem dúvida alguma, não teria conseguido chegar até aqui: Talinhos (obrigada por tudo!), Profs. Bruno e Kyria, Júlia, Eduardo, Sâmia, Raniere, Ito, Alysson, Rafael, Helton, Júnior, Camila, Suzete, Bruno Lopes, Mayron, Theodora, Thaiz, Joana, Raphael, Cíntia, Claudener, Rômulo, Arthur, Guilherme, Fernando, Kássia, Mayara Torquato, Ivandra, Ronniery, Mayara Queiroz, Vanir, Vinícius, Cláudia, Lia, Davi, Caio, Clareane, Alfa, Cibele.

Agradeço igualmente a todos os membros da equipe *Glicobiologie Moléculaire*, pela grande ajuda nessa tese: em especial Catherine Gautier e Valérie Chazalet, pelos ensinamentos, suporte e grande ajuda nos momentos mais difíceis (muito obrigada pelo carinho!). Meu grande obrigada também ao Dr. Olivier Lexourel e aos meus colegas Bertrand Blanchard, Sophie Mathieu, Michael Reynolds, Alessandra Nurisso, Sami Cecioni, Ondraj Sulak, Aymeric Audfray, Nicolas Sapay, Géraldine Ganne, Anita Sarkar e Sumaira Kousar.

De maneira especial e particular, agradeço à minha família, minha mãe Lindóia e irmãs Rafaela e Rebeca e meu cunhado Muniz, que me deram todo o suporte em amor, compreensão e paciência durante essa fase, compartilhando meus obstáculos e apoiando: vocês são minha maior motivação.

A todas as pessoas amigas e companheiras que contribuíram direta ou indiretamente para a realização desse trabalho, com críticas construtivas, favores, incentivos e apoio.

Registro aqui também o meu muito obrigado, faltando-me palavras para expressar minha gratidão, amor e grande felicidade de poder ter ao meu lado Elton Menezes Leite, meu grande amor, companheiro, incentivador e, muitas vezes, guia, a quem devo muito do que hoje sou... muito obrigada por tudo.

SUMÁRIO

APRESENTAÇÃO	78
LISTA DAS PRINCIPAIS ABREVIATURAS E SIGLAS.....	79
LISTA DE FIGURAS.....	81
LISTA DE TABELAS	84
RESUMO	85
ABSTRACT.....	86
Capítulo 1: Introdução	87
1. Generalidades sobre Glicobiologia e Carboidratos	87
2. Generalidades sobre lectinas	90
3. Lectinas de Plantas	93
3.1. Funções Endógenas de Lectinas de Plantas	95
3.2. Classificação Estrutural de Lectinas	96
3.3. Classificação de Lectinas de Plantas em Famílias Evolutivamente Relacionadas	98
3.4. Lectinas de Leguminosas	103
4. Lectinas de Dalbergiae	113
5. Lectina de sementes de <i>Platypodium elegans</i>	115
6. Clonagem, expressão heteróloga e purificação de lectinas de leguminosas.....	117
6.1. Clonagem.....	121
6.2. Expressão e Purificação.....	121
7. Estudos Estruturais e Termodinâmicos de Interações Lectina-Carboidrato.....	123
8. Objetivos do Trabalho de Tese.....	127
8.1. Objetivo geral.....	127
8.2. Objetivos específicos.....	127
Capítulo 2: Materiais e Métodos	129
I. Clonagem e Expressão Heteróloga de uma lectina de sementes de <i>Platypodium elegans</i>.....	129
1. Obtenção das Sementes de <i>Platypodium elegans</i>	129
2. Isolamento do RNA Total de Sementes de <i>P. elegans</i>	130
3. Obtenção, Clonagem e Caracterização do gene da lectina de <i>P. elegans</i>	131
3.1. 3'RACE (System for Rapid Amplification of cDNA Ends)	132
3.2. 5'RACE (System for Rapid Amplification of cDNA Ends)	140
4. Determinação da Sequência Protéica e Alinhamento de Lectinas de Leguminosas	144
5. Construção do Vetor de Expressão e Transformação em células de células de <i>E. coli</i>	
BL21 competentes.....	145
5.1. Desenho de primers específicos para o gene de PELA adicionados de sítios de	
restrição	145
5.2. Amplificação do inserto para construção de vetor de expressão.....	146
5.3. Digestão enzimática do inserto purificado e do vetor pET-29a	148
5.4. Clonagem do inserto no vetor pET-29a.....	149
5.5. Transformação do plasmídeo de expressão em células de <i>E. coli</i> BL 21 competentes	
.....	150
6. Expressão recombinante da lectina de sementes de <i>P. elegans</i>	150
7. <i>Refolding</i> e purificação da lectina recombinante de <i>P. elegans</i>	151
II. Ferramentas para o estudo de interação lectina-caboidrato	153
1. Atividade hemaglutinante da lectina PELA	153

2.	Análise da especificidade de PELA por <i>Glycan-Array</i>	153
3.	Ensaio de microcalorimetria- <i>Isothermal Titration Microcalorimetry</i> (ITC).....	155
3.1.	Princípio da Técnica	156
3.2.	Noções de Termodinâmica	158
3.3.	Análise da afinidade da lectina de <i>P. elegans</i> por oligossacarídeos através de ITC	159
III. Estudos da estrutura cristalina de PELA através de Cristalografia por Difração em raios X		
161		
1.	Cristalização – Obtenção dos Cristais de Proteína	161
2.	Difração em Raios X, coleta e processamento de dados	164
3.	Refinamento e Determinação da Estrutura	168
4.	Resolução da estrutura cristalina de PELA em complexo com oligossacarídeos	170
4.1.	Cristalização	170
4.2.	Coleta de dados em difração de raios X.....	171
4.3.	Substituição Molecular.....	172
4.4.	Refinamento e Validação	172
Capítulo 3: Clonagem e expressão heteróloga de uma lectina de sementes de <i>Platypodium elegans</i> (PELa).....		
174		
1.	Extração e análise do RNA total de sementes de <i>P. elegans</i>	174
2.	Estrutura, caracterização e clonagem do mRNA de PELA	176
2.1.	Produto 3' RACE	176
2.2.	Produto 5' RACE	177
2.3.	Estrutura e caracterização do mRNA de PELA	178
3.	Determinação da sequência protéica e Alinhamento com lectinas de leguminosas	180
4.	Clonagem no vetor de expressão pET-29a e transformação em células de <i>E. coli</i> BL21	183
5.	Indução e Expressão de PELA	184
6.	<i>Refolding</i> e Purificação de PELA	185
Capítulo 4: Estudo das propriedades hemaglutinantes, especificidade e afinidade da lectina PELA.....		
188		
1.	Propriedades hemaglutinantes de PELA	188
2.	Estudos de especificidade e afinidade na interação lectina-carboidrato em PELA... ..	189
2.1.	Estudos de especificidade por Glycan Array	190
2.2.	Estudos de afinidade por microcalorimetria (ITC):	192
Capítulo 5: Estudos Estruturais da forma cristalina de PELA em complexo com oligossacarídeos		
197		
1.	Estrutura cristalina de PELA em complexo com trimanose	198
1.1.	Cristalização	198
1.2.	Coleta de dados	198
1.3.	Substituição molecular e refinamento da estrutura	200
1.4.	Análise da estrutura PELA em complexo com trimanose.....	201
2.	Estrutura cristalina de PELA em complexo com heptassacarídeo simétrico	209
2.1.	Cristalização	209
2.2.	Coleta de dados	210
2.3.	Substituição molecular e refinamento da estrutura	211
2.4.	Análise da estrutura de PELA em complexo com heptassacarídeo	213

CONCLUSÕES E PERSPECTIVAS.....	221
RESUMO DE OUTROS TRABALHOS REALIZADOS NO PERÍODO DO DOUTORADO:.....	222
REFERÊNCIAS BIBLIOGRÁFICAS.....	226

APRESENTAÇÃO

Os assuntos abordados nessa tese se situam dentro do domínio da Glicobiologia estrutural, um ramo da biologia que estuda as estruturas e as interações de glicomoléculas. Durante os quase quatro anos desta tese, passados um pouco mais da metade no Departamento de Bioquímica e Biologia Molecular (UFC) e o outro período no CERMAV (UJF, França), fui confrontada com sujeitos diferentes, mas complementares, tais quais biologia molecular, análise estrutural de complexos proteína-carboidratos e caracterização termodinâmica de interações lectina-glicídios.

Essa tese pode ser dividida em três partes principais. Uma primeira dedicada à clonagem e à produção da lectina recombinante; a segunda descreve os estudos de afinidade e especificidade dessa lectina frente a diferentes carboidratos; a terceira analisa a estrutura cristalina da lectina complexada com diferentes ligantes glicídicos.

O primeiro capítulo é uma introdução à glicobiologia e ao objeto de nosso estudo: as lectinas. No segundo capítulo, uma atenção particular é dada aos métodos experimentais utilizados, compreendendo técnicas de biologia molecular, “*glicochips*”, microcalorimetria e cristalografia. O terceiro capítulo é dedicado à produção de uma lectina de leguminosa sob forma recombinante a partir de material genético de sementes. O quarto capítulo é voltado às interações lectinas-carboidratos e descreve a especificidade desta lectina PELa frente a glicanos e a análise da afinidade por alguns oligossacarídeos. No quinto capítulo, o estudo estrutural dos complexos da lectina com diferentes carboidratos é abordado. Algumas conclusões e perspectivas fecham a apresentação desse trabalho. Ao fim, há um resumo sobre outros trabalhos desenvolvidos durante essa tese.

Essa tese foi redigida em língua portuguesa, mas compreende dois anexos: o resumo expandido do trabalho de tese em língua francesa e o manuscrito em língua inglesa de uma publicação que está sendo submetida ao *Journal of Biological Chemistry*.

LISTA DAS PRINCIPAIS ABREVIATURAS E SIGLAS

AAP	<i>abridged anchor primer</i>
ADNc	ácido desoxirribonucléico complementar
AP	<i>anchor primer</i>
ARNm	ácido ribonucléico mensageiro
CFL	lectina de <i>Cratylia floribunda</i>
CGL	lectina de <i>Canavalia gladiata</i>
ConA	lectina de <i>Canavalia ensiformes</i>
ConBol	lectina de <i>Canavalia boliviana</i>
ConBr	lectina de <i>Canavalia brasiliensis</i>
ConM	lectina de <i>Canavalia maritima</i>
CRD	domínio de reconhecimento a carboidrato
DGL	lectina de <i>Dioclea grandiflora</i>
DGL	lectina de <i>Dioclea grandiflora</i>
Dgui	lectina de <i>Dioclea guianenses</i>
DMF	dimetilformamida
DO ₆₀₀	densidade óptica a 600 nm
Dvir	lectina de <i>Dioclea virgata</i>
ESRF	<i>European Synchrotron Radiation Facility</i>
GNA	aglutinina de <i>Galanthus nivalis</i>
IPTG	isopropil- β -D tiogalactosídeo
ITC	microcalorimetria de titulação isotérmica
K _a	constante de associação
K _d	constante de dissociação
LCA	aglutinina de <i>Lens culinaris</i>
LOL	lectina de <i>Lathyrus ochrus</i>
PAL	lectina de <i>Pterocarpus angolensis</i>
pb	pares de bases
PDB	<i>protein data bank</i>
PELa	lectina de <i>Platypodium elegans</i>

PHA	fitohemaglutinina
RACE	rápida amplificação de extremidades de cDNA
SDS-PAGE	<i>sodium dodecyl sulfate polyacrylamide gel electrophoresis</i>
TFP	fenil-tetrafluoro
Tm	Temperatura de fusão
UH	unidade de hemaglutinação
UTR	região terminal não codificadora

LISTA DE FIGURAS

Figura 1. Conceitos Gerais de Estrutura de Carboidratos	88
Figura 2. Papéis de oligossacarídeos no reconhecimento e na adesão de superfície da célula	92
Figura 3. Classificação estrutural de lectinas de plantas	98
Figura 4. Classificação de lectinas de plantas em famílias evolucionamente relacionadas ..	100
Figura 5. Modificações pós-traducionais durante a biossíntese da Concanavalina A (ConA)	105
Figura 6. Monômero típico de uma lectina de leguminosa	110
Figura 7. Estrutura cristalina da lectina de <i>Canavalia gladiata</i> (CGL) complexada com α -metil-manosídeo	109
Figura 8. Reconhecimento de Man(α 1-3)Man e Man(α 1-6)Man por <i>Pterocarpus angolensis</i> lectin (PAL)	115
Figura 9. <i>Platypodium elegans</i>	116
Figura 10. Mapa do vetor pGEM-T easy (Promega)	136
Figura 11. Mapa do vetor pET-29a (Novagen)	147
Figura 12. <i>Glycan array</i> como uma imitação da superfície celular, expressando glicanos para estudo de interações multivalentes	154
Figura 13. Descrição de <i>glicochips</i>	155
Figura 14. Representação esquemática de um microcalorímetro	157
Figura 15. Princípio da microcalorimetria de titulação isotérmica	157
Figura 16. Principais passos na resolução da estrutura de uma biomolécula por cristalografia	162
Figura 17. Princípio de cristalização pelo método de difusão de vapor - gota suspensa	163
Figura 18. Formação e características de uma rede cristalina	165
Figura 19. Visão geral do modelo de Bragg	166
Figura 20. Esfera de Ewald: rede recíproca e a geometria de difração	167
Figura 21. Eletroforese em gel de agarose (1%). Extração do RNA total de sementes de <i>P. elegans</i>	175

Figura 22. Eletroforese em gel de Agarose (1,2%). Produtos 3'RACE	176
Figura 23. Eletroforese em gel de Agarose (1,2%). Produto 3' RACE purificado e reamplificado	177
Figura 24. Eletroforese em gel de Agarose (1%). Produto 5' RACE	178
Figura 25. Estrutura completa do mRNA codificante para a lectina de <i>P. elegans</i> (região codificadora: 786 nucleotídios), deduzida à partir dos produtos 3' e 5' RACE	179
Figura 26. Sequência nucleotídica codificadora do mRNA da lectina de <i>P. elegans</i>	179
Figura 27. Sequência polipeptídica da lectina de <i>P. elegans</i> deduzida a partir do mRNA correspondente	181
Figura 28. Alinhamento de sequências polipeptídicas de lectinas de leguminosas	182
Figura 29. Validação da transformação de células de <i>E. coli</i> com o plasmídeo pET-29a_PELa	183
Figura 30. SDS-PAGE 12% de extrato de células de <i>E. coli</i> BL21 induzidas com IPTG 0,5 mM	184
Figura 31. SDS-PAGE 12% do extrato celular de <i>E. coli</i> BL21 após indução e ruptura	185
Figura 32. Cromatografia de afinidade em matriz de Manose-Agarose: Purificação de PELa pós- <i>refolding</i>	187
Figura 33. SDS-PAGE 12%: <i>refolding</i> e purificação de PELa	187
Figura 34. Hemaglutinação de hemácias de coelho por PELa. (-)	189
Figura 35. <i>Glycan array</i>	191
Figura 36. Compostos H (heptassacarídeo simétrico) e E (heptassacarídeo assimétrico) em suas formas ligadas à Asn, utilizados nos experimentos de microcalorimetria	193
Figura 37. Curva de microcalorimetria obtida para PELa (0,06 mM) com o heptassacarídeo simétrico H (0,775 mM)	194
Figura 38. Curvas de microcalorimetria obtidas para PELa (0,6 mM) e os oligossacarídeos H (0,775 mM, à esquerda) e E (0,45 mM, à direita) em tampão Tris (pH 7,6) a 25 °C	195
Figura 39. Representação dos Compostos A (trimanose) e H (heptassacarídeo simétrico), utilizados nos experimentos de cristalografia	197
Figura 40. Cristais de PELa obtidos na presença de trimanose (α Man1-3(α Man1-6)Man ..	198
Figura 41. Visão geral da estrutura dimérica de PELa em complexo com trimanose,	202

formada pelos monômeros A em vermelho e monômero B em azul	
Figura 42. Estrutura monomérica de PELa	205
Figura 43. Estrutura cristalina de PELa em complexo com trimanose	204
Figura 44. Representação esquemática bidimensional (LigPlot) da rede de pontes de hidrogênio entre os resíduos chave envolvidos na interação com o ligante trimanose, compartilhado entre os monômeros A (resíduos em vermelho) e B (resíduos em azul) de dímeros distintos de PELa	205
Figura 45. Estrutura cristalina do complexo PELa/trimanose	206
Figura 46. Representação da arquitetura do sítio de ligação a carboidrato de PELa complexada com trimanose, evidenciando os 5 voltas (A-E) típicos de lectinas de leguminosas	207
Figura 47. Comparação do reconhecimento do Trissacarídeo (Man) por PELa (A) e outras lectinas de leguminosas	209
Figura 48. Cristais de PELa obtidos na presença de heptassacarídeo simétrico (GlcNAc β 1-2Man α 1-3(GlcNAc β 1-2Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAcB-Asn)	210
Figura 49. Visão geral da estrutura dimérica de PELa em complexo com o heptassacarídeo simétrico, formada pelos monômeros A em vermelho e B em azul	214
Figura 50. Representação da superfície do sítio de ligação de PELa em complexo com o heptassacarídeo	214
Figura 51. Representação esquemática bidimensional (LigPlot) da rede de ligações de hidrogênio entre os principais resíduos envolvidos na interação do sítio de ligação de PELa com o ligante heptassacarídeo	215
Figura 52. Estrutura cristalina do complexo PELa/heptassacarídeo. Representação do oligossacarídeo sobre a superfície da proteína	216
Figura 53. Representação da arquitetura do sítio de ligação a carboidrato de PELa complexada com heptassacarídeo, evidenciando as 5 voltas (A-E) típicas de lectinas de leguminosas	217
Figura 54. Comparação do complexo PELa/heptassacarídeo com outros de lectinas de leguminosas relacionadas	218
Figura 55. Comparação do reconhecimento dos ligantes trimanose e heptassacarídeo pela lectina PELa	219

LISTA DE TABELAS

Tabela 1. Exemplos de atividades biológicas descritas para lectinas extraídas a partir de espécies de leguminosas	44
Tabela 2. Especificidade por açúcares das lectinas da subtribo Dalbergieae	45
Tabela 3. Lectinas de Leguminosas clonadas e expressas em <i>E. coli</i>	51
Tabela 4. Desenho de <i>primers</i> degenerados para o gene de PELa	66
Tabela 5: Desenho de <i>primer</i> específico para o gene a partir da sequência nucleotídica do produto 3' RACE	73
Tabela 6. Iniciadores utilizados na clonagem do gene de PELa no vetor de expressão pET-29a	78
Tabela 7. Lista dos iniciadores (<i>primers</i>) utilizados nesse estudo	107
Tabela 8. Dados de ITC para a ligação dos oligossacarídeos a PELa	126
Tabela 9. Estatística da coleta e do processamento de dados de difração de raios X da lectina PELa em complexo com trimanose	131
Tabela 10. Estatística dos dados de substituição molecular e refinamento da estrutura de PELa em complexo com trimanose	133
Tabela 11. Estatística da coleta e do processamento de dados de difração de raios X da lectina PELa em complexo com heptassacarídeo simétrico	143
Tabela 12. Estatística dos dados de substituição molecular e refinamento da estrutura de PELa em complexo com heptassacarídeo simétrico	144

RESUMO

Uma atividade lectínica com especificidade para manose e glicose foi detectada em sementes de *Platypodium elegans*, uma leguminosa da subtribo Dalbergiae. O gene da lectina PELa foi clonado, resultante em uma proteína de 261 resíduos de aminoácidos pertencente à família de lectinas de leguminosas, com similaridade com a lectina de *Pterocarpus angolensis* (PAL). A lectina recombinante foi expressa em células de *E. coli* e refoldada à partir de corpos de inclusão. Uma análise da especificidade por *Glycan array* evidenciou uma preferência bastante não-usual para *N*-glicanos do tipo complexos com ramificações assimétricas. Um braço curto consistindo em um resíduo de manose é preferido no braço 1-6 do *N*-glicano, enquanto uma extensão por GlcNAc, Gal e NeuAc são favoráveis no braço 1-3. Valores de afinidade foram obtidos por microcalorimetria usando heptassacarídeos simétricos e assimétricos ligados a Asn preparados por método semi-enzimático. Uma alta afinidade de 5 μ M foi obtida para ambos os ligantes. Duas estruturas cristalinas de PELa, uma em complexo com um trimanose ramificado e outra com um heptassacarídeo complexo simétrico ligado a Asn foram resolvidas a resoluções de 2,1 e 1,65 Å, respectivamente. A lectina mostrou adotar uma organização dimérica canônica típica de lectinas de leguminosas. O trimanose faz uma ponte entre os sítios de ligação a carboidrato de dímeros vizinhos, resultando na formação de cadeias infinitas no cristal. O heptassacarídeo liga-se com o braço 1-6 no sítio primário de ligação e com contatos extensivos adicionais em ambos os braços. O GlcNAc do braço 1-3 está ligado em uma conformação restrita que pode justificar a alta afinidade que é observada em *chips* para oligossacarídeos com braço 1-3 curto que não apresentam esse monossacarídeos.

ABSTRACT

Lectin activity with specificity for mannose and glucose has been detected in the seed of *Platypodium elegans*, a legume from the *Dalbergiae* tribe. The gene of the lectin PELa has been cloned and the resulting 261 amino acid protein belongs to the legume lectin family with similarity with *Pterocarpus angolensis* agglutinin (PAL) from the same tribe. The recombinant lectin has been expressed in *E. coli* and refolded from inclusion bodies. Analysis of specificity by Glycan Array evidenced a very unusual preference for complex type *N*-glycans with asymmetrical branches. A short branch consisting of one mannose residue is preferred on the 6-arm of the *N*-glycan, while extension by GlcNAc, Gal and NeuAc are favorable on the 3-arm. Affinities have been obtained by microcalorimetry using symmetrical and asymmetrical Asn-linked heptasaccharide prepared by semi-enzymatic method. Strong affinity of 5 μ M was obtained for both ligands. Crystal structures of PELa complexed with branched trimannose and symmetrical complex type Asn-linked heptasaccharide have been solved at 2.1 and 1.65 Å resolution respectively. The lectin adopts the canonical dimeric organization of legume lectins. The trimannose bridges the binding sites of two neighbouring dimers, resulting in the formation of infinite chains in the crystal. The Asn-linked heptasaccharide binds with the 6-arm in the primary binding site and extensive additional contacts on both arms. The GlcNAc on the 3-arm is bound in a constrained conformation that may rationalize the higher affinity that is observed on chips for oligosaccharide with shorter 3-arm that do not present this monosaccharide.

Capítulo 1:

Introdução

1. Generalidades sobre Glicobiologia e Carboidratos

Glicobiologia é a ciência que estuda as estruturas, propriedades e funções dos carboidratos presentes em diversidade nos seres vivos. Estes estudos não se restringem apenas aos carboidratos, mas também a todas as outras classes de biomoléculas (proteínas, lipídios, nucleotídeos) que interagem com eles. O objetivo da glicobiologia estrutural é, naturalmente, determinar a estrutura tridimensional e estabelecer modelos para as interações envolvendo essas moléculas (TAYLOR & DRICKAMER, 2003).

Os carboidratos são a maior fonte de biomassa do planeta Terra e sua complexidade, especialmente em termos de variações estruturais, não tem igual em outras classes de biomoléculas. Do ponto de vista químico, os açúcares são poliídrixialdeídos ou poliídrixicetonas. Monossacarídeos têm uma fórmula geral do tipo $(\text{CH}_2\text{O})_n$ com um número de átomos de carbono da cadeia principal que pode variar de três a dez unidades. Uma característica geral de açúcares é que um dos átomos de carbono possui uma função do tipo ceto-aldeído, que permite a carboidratos com cinco (pentose) ou seis átomos de carbono (hexoses) uma condensação intramolecular causando a formação de um anel (furanose ou piranose), que é a sua forma predominante em solução. Esse processo denominado de ciclização dá origem a um novo centro quiral na molécula, o carbono anomérico C1, que pode ter duas configurações chamadas α ou β , que se interconvertem por um processo denominado mutarrotação. O carbono anomérico pode se condensar com a hidroxila de um segundo monossacarídeo para formar um dissacarídeo. Assim, a adição de monômeros resulta na formação de estruturas cuja complexidade aumenta drasticamente com o número de resíduos. As estruturas mais simples, lineares ou ramificadas, são formadas pela adição de 2-5 monômeros e são nomeados com os prefixos di, tri, tetra- e pentassacarídeo; oligossacarídeos, para as moléculas que contêm até vinte monossacarídeos; e

quando o tamanho e a complexidade da estrutura se tornam muito importantes (mais de 20 resíduos), em geral são chamadas de polissacarídeos (Figura 1) (NELSON & COX, 2009).

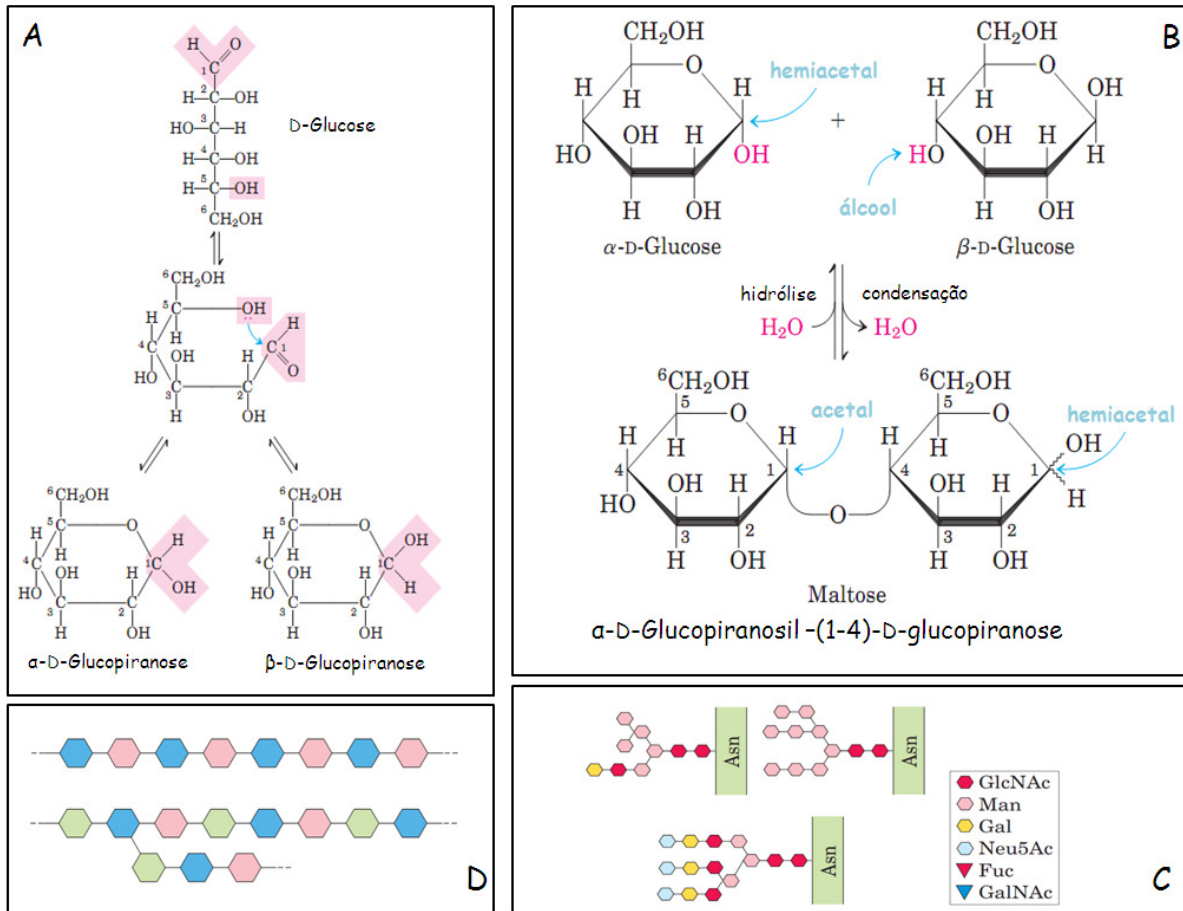


Figura 1. Conceitos Gerais de Estrutura de Carboidratos. A. Monossacarídeo glicose e mutarotação. B. Ligação Glicosídica: formação do dissacarídeo maltose. C. Oligossacarídeos. D. Polissacarídeos simples e ramificados (Adaptado de NELSON & COX, 2004).

O enorme potencial biológico codificado por carboidratos resulta da sua capacidade de criar estruturas complexas, incluindo pontos de ramificação a partir de monômeros simples (lineares). Esta característica dos carboidratos é única entre todos os polímeros biológicos. Por exemplo, dois aminoácidos diferentes podem se ligar apenas por uma de duas maneiras. Por outro lado, um monossacarídeo simples, como a glicose, pode se ligar com o seu carbono anomérico (α ou β) a cinco hidroxilas presentes em uma segunda glicose, dando origem a até onze diferentes isômeros estruturais. É compreensível, portanto, porque o número de variações estruturais fornecidas por este tipo de moléculas é enorme.

A estrutura de glicanos (oligossacarídeos, glicoconjugados, polissacarídeos) depende fortemente da sua composição e, além disso, a presença de ligações glicosídicas introduz uma flexibilidade estrutural e conformacional considerável. A consequência direta dessa variabilidade estrutural é uma grande diversidade no papel e localização de carboidratos nos organismos vivos: resistência estrutural, armazenamento de energia, reconhecimento molecular, crescimento e desenvolvimento, defesa, âncoragem membranar, endereçamento, dentre outros (TAYLOR & DRICKAMER, 2003).

Foi somente à partir de metade do século XX que os carboidratos foram reconhecidos como moléculas biológicas extremamente importantes; até então se admitia que todas as informações biológicas eram realizadas pelas duas principais classes de moléculas: proteínas e ácidos nucléicos (LIS & SHARON, 1998). Em 1900, surgiu um interesse principalmente em polissacarídeos de plantas, pela sua ampla utilização em muitas aplicações industriais. Conhecimentos na biologia dos carboidratos eram, ainda, limitados àqueles com papéis estruturais (celulose, quitina, etc ...) ou de acúmulo e transporte de energia (glicose, amido). A percepção e a importância de que glicanos, livre ou na forma de glicoconjugados, poderiam desempenhar papéis diversos foi a principal motivação que deu origem a glicobiologia moderna (SHARON & LIS, 1993). Grandes descobertas que marcaram a elucidação do papel dos oligossacarídeos: o estabelecimento da estrutura glicídica dos determinantes do sistema ABO; a demonstração da aglutinação de eritrócitos pelo vírus influenza pela interação entre a hemaglutinina e ácido siálico; a observação de que a defucosilação de linfócitos impedia sua migração pelos tecidos. Estas descobertas foram a fonte de muita pesquisa que tem ligado o papel dos carboidratos nos processos de comunicação e reconhecimento molecular (SHARON & LIS, 1993).

Funções biológicas associadas a porção de carboidratos de glicoconjugados são tão numerosos quanto a sua diversidade estrutural. Duas grandes famílias de funções podem ser destacadas: a função estrutural e de reconhecimento molecular. Em relação às propriedades estruturais e físicas de modulação, estas moléculas geralmente têm uma alta carga eletrostática, além uma grande superfície de hidratação. Estas propriedades físicas lhe dão um papel na estabilização estrutural, solubilização e proteção contra a proteólise em glicoproteínas. Glicoconjugados são também parte da camada de alta densidade molecular que cobre a superfície de células epiteliais em eucariotos, o glicocálix. Esta área está envolvida na coesão

celular e na proteção da célula contra choques físicos e contra o ataque de microorganismos. Ligado ao papel de reconhecimento molecular, os oligossacarídeos estão envolvidos em muitos fenômenos de reconhecimento molecular do tipo receptor-ligante envolvendo lectinas e muitos estudos confirmam papéis na comunicação entre células (Figura 2) (VARKI, 1993; DENNIS *et al.*, 1999). Além disso, foi demonstrado que anormalidades no padrão de glicosilação frequentemente acompanham condições anormais fisiológicas ou patológicas, como o câncer. Portanto, este objeto de estudo agora está recebendo atualmente muita atenção e estudos têm sido realizados a fim de elucidar as complexas estruturas e papéis dessas glicanos (HAKOMORI, 1985; DENNIS *et al.*, 1999).

2. Generalidades sobre lectinas

O termo lectina derivado do verbo latino “legere”, que significa "selecionar" ou "escolher" é um nome apropriado para esta classe muito importante de proteínas. Lectinas têm sido definidas como proteínas de origem não-imune que se ligam especificamente e reversivelmente aos açúcares e não mostram atividade enzimática para estes substratos (KOCOUREK & HOREJSI, 1981). Lectinas são moléculas ubíquas, encontradas em todas as classes de organismos: microorganismos (vírus, bactérias), plantas, fungos, insetos e animais. Também são chamadas de aglutininas por serem capazes de aglutinar células (eritrócitos, por exemplo) e glicoconjugados. Esta característica é devido ao fato destas proteínas serem geralmente multivalentes.

O começo da documentação da pesquisa sobre lectinas data de 1860. Silas Weir Mitchekk (1829-1913), um experiente psiquiatra e neurologista, observou que uma gota do veneno da serpente *Crotalus durissus* e uma gota de sangue da asa ferida de um pombo coagularam rapidamente em menos de três minutos (MITCHELL, 1860). Após mais de 100 anos após essa observação pioneira, Ogilvie & Gartner (1984) foram capazes de provar a existência de lectinas Ca^{++} dependentes em venenos de Crotalidae responsáveis por nessa atividade. Tecnicamente, aglutinação dependente de açúcar tem sido o ensaio comum mais utilizado para se detectar e monitorar atividade lectínica, e lectinas tem sido classificadas de acordo com sua especificidade a monossacarídeos mensurada em ensaios sistemáticos de inibição dessa atividade. A introdução de cromatografia de afinidade para purificação de lectinas foi feita em 1965 e o desenvolvimento de técnicas de sequenciamento protéico e genômico tem permitido a determinação de suas

estruturas primárias, assim como a aplicação de técnicas biofísicas para determinação do seu enovelamento, abrindo caminho para refinar o critério para agrupar lectinas em famílias (GABIUS, 2001; GABIUS *et al.*, 2004; SHARON & LIS, 2001).

Um dos maiores motivos que hoje incentiva a pesquisa sobre lectinas provavelmente está relacionada a sua capacidade única de "ler" as informações biológicas que são codificadas na estrutura tridimensional de açúcares. Lectinas são, de fato, os receptores para as interações proteína-açúcar, tendo um papel fundamental em vários processos de reconhecimento molecular e sinalização celular (AMBROSI *et al.*, 2005) (Figura 2). Lectinas localizadas na superfície de bactérias, vírus e parasitas intestinais reconhecem glicoconjugados presentes na superfície das células epiteliais, facilitando os processos de colonização e infecção. Em relação aos papéis das lectinas de plantas, uma boa hipótese tem sido o papel de defesa da planta contra patógenos de plantas (ou contra os animais que podem se alimentar das mesmas), mas outras possibilidades são consideradas (RUDIGER & GABIUS, 2001). As funções das lectinas animais, por outro lado, já têm sido estudadas em mais detalhes.

É interessante notar que a maioria das lectinas são específicas para um pequeno número de açúcares e, na maioria dos casos, esses açúcares estão presentes dentro ou na superfície da célula, especialmente na forma de glicoconjugados. Podemos identificar duas classes de lectinas com relação a sua especificidade: aquelas que reconhecem um determinado monossacarídeo e aquelas que reconhecem apenas os oligossacarídeos (SHARON & LIS, 2003). Entre os monossacarídeos geralmente reconhecido por lectinas encontramos manose, fucose, galactose/GalNAc e *N*-acetilglicosamina.

Na primeira classe, com poucas exceções, estas lectinas têm uma elevada afinidade para seus ligantes monossacarídeos com uma constante de afinidade (K_a) da ordem de 10^3 a 10^4 M^{-1} . Em geral, a especificidade é alta para um monossacarídeo e poucas substituições são permitidas. A afinidade é geralmente modulada pela presença de um substituinte no carbono anomérico do monossacarídeo e, portanto, essas lectinas também reconhecem di- e trissacarídeos (GOLDSTEIN & PORETZ, 1986; LEE & LEE, 1995).

A segunda classe de lectinas é formado por aquelas que necessitam de oligossacarídeos específicos, como a *N*-cyanovirina, que reconhece glicanos do tipo oligomanose, como a glicoproteína gp120 do HIV (BOTOS *et al.*, 2002) ou a toxina da cólera que é específica para GM1 presentes na superfície das células epiteliais (MERRITT & HOL, 1995). A afinidade

mostrada para oligossacarídeos geralmente é muito maior do que para os monossacarídeos pela presença de sítios de ligação mais profundo formando mais contatos com o ligante. Por vezes, relatam-se lectinas com um notável K_a que pode alcançar a ordem de $10^7 - 10^8 \text{ M}^{-1}$ (GOLDSTEIN & PORETZ, 1986; LEE & LEE, 1995).

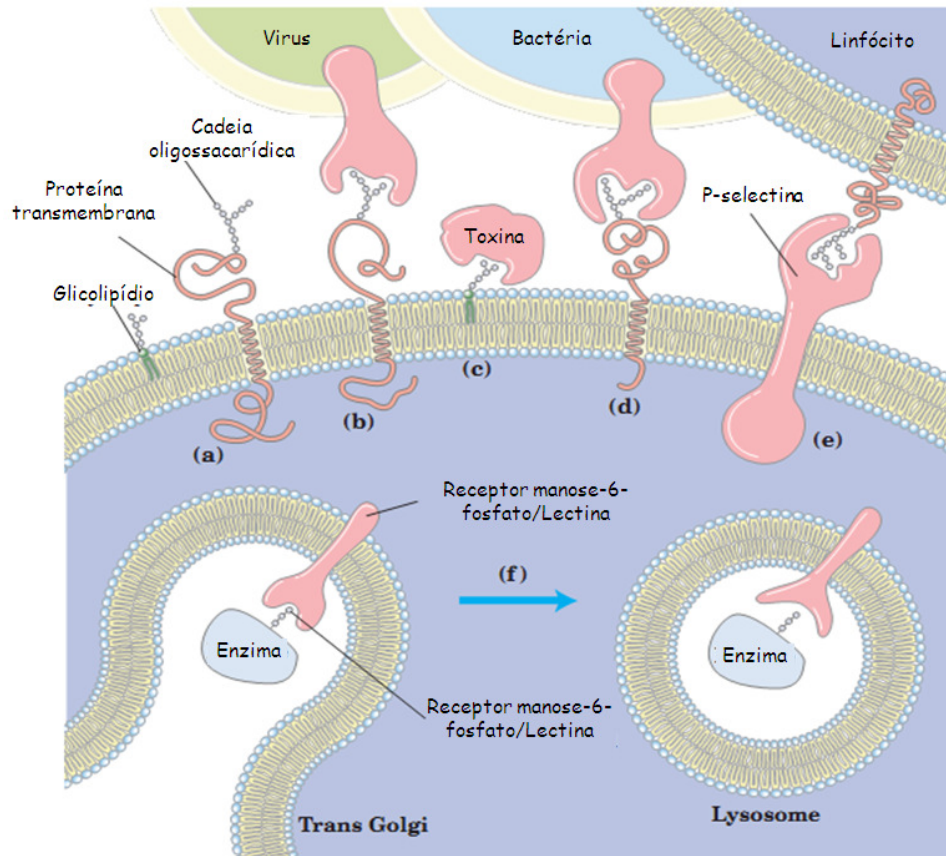


Figura 2. Papéis de oligossacarídeos no reconhecimento e na adesão de superfície da célula. (a) Oligossacarídeos com estruturas únicas (representados como hexágonos), componentes de uma variedade de glicoproteínas ou glicolípídeos na superfície externa da membrana plasmática, interagem com elevada especificidade e afinidade com lectinas do meio extracelular. (b) Vírus que infectam células animais, tais como o vírus da gripe, se ligam a glicoproteínas da superfície celular como primeiro passo para a infecção. (c) Toxinas bacterianas, como as da cólera e pertússis, se ligam a um glicolípídeo de superfície antes de entrar em uma célula. (d) Algumas bactérias, como *Helicobacter pylori*, aderem à célula e depois as colonizam ou infectam. (e) Selectinas (lectinas) na membrana plasmática de certas células mediam interações célula-célula, como as de linfócitos T com as células endoteliais da parede capilar em um sítio de infecção. (f) O receptor de manose 6-fosfato / lectina do complexo de Golgi se liga ao oligossacarídeo de enzimas lisossomais, orientando-as para a transferência ao lisossomo. (Adaptado de NELSON & COX, 2009).

3. Lectinas de Plantas

Lectinas foram primeiramente identificadas em plantas no início do século XX, mas a comunidade científica teve um interesse maior por esta classe de proteínas a partir dos anos sessenta, coincidindo com o nascimento da glicobiologia. O interesse foi motivado principalmente pelo uso de lectinas na detecção, isolamento e caracterização de oligossacarídeos, na tipagem sanguínea e identificação de glicoconjugados, especialmente glicoproteínas. A grande descoberta de que certas condições fisiológicas e patológicas estavam associadas a alguma mudança no estado de glicosilação das células só foi possível através da utilização de lectinas. O número de publicações sobre lectinas tem tido um grande crescimento, principalmente devido à abundância de lectinas em todos os organismos vivos, com uma certa facilidade de purificação. Nos últimos anos a comunidade científica começou a considerá-las como moléculas bioativas, aumentando o interesse nas funções biológicas destas moléculas. Percebe-se de imediato a extrema variabilidade das lectinas, primeiro em termos de estrutura primária, e, como resultado das primeiras estruturas resolvidas por difração de raios X, em termos de estrutura tridimensional (SHARON, 2007, LORIS, 2002).

A habilidade para o reconhecimento e ligação a carboidratos específicos distinguem as lectinas de todas as outras proteínas de plantas. Elas são classicamente consideradas um grupo heterogêneo de proteínas, pois apresentam propriedades bioquímicas e atividades biológicas acentuadamente diferentes, distintas estruturas moleculares e especificidades (PEUMANS & VAN DAMME, 1995)

Apesar das lectinas serem amplamente distribuídas no reino vegetal, abrangendo principalmente as famílias Leguminosae, Gramineae, Euphorbiaceae, entre outras, a família Leguminosae é a que apresenta o maior número de lectinas isoladas e destas destacam-se principalmente as de sementes (SHARON & LIS, 1993; PRAKASHKUMAR *et al.*, 1998; LORIS *et al.*, 1998), embora possam ser encontradas também em outras partes das plantas, como folhas (RATANAPO *et al.*, 2001), frutos (SAMPIETRO *et al.*, 2001), raízes de algumas Convolvulaceae (PEUMANS *et al.*, 1997; VAN DAMME *et al.*, 1997) e tubérculos (SUSEELAN *et al.*, 2002).

Muitas funções têm sido propostas para as lectinas vegetais, tais como proteção contra patógenos e insetos, transporte e armazenamento de carboidratos, reconhecimento celular (dentro

da célula, entre células ou entre organismos), proteínas de reserva ou reguladores de crescimento (PUSZTAI, 1991).

Atualmente estas proteínas têm mostrado aplicabilidades em diversas áreas científicas, abrangendo áreas que vão da medicina à agricultura, sendo muitas destas proteínas consideradas como ferramentas básicas para áreas de ponta como a biotecnologia, tendo em vista suas várias propriedades, tais como: atividade anti-inflamatória e pró-inflamatória (FREIRE *et al.* 2003, BENJAMIN *et al.*, 1997, ASSREUY *et al.*, 1997, ALENCAR *et al.*, 2005), atividade inseticida (MACEDO *et al.*, 2007; OHIZUMI *et al.*, 2009, BENETEAU *et al.*, 2010), efeito tóxico para moluscos hospedeiros de *Schistosoma mansoni* (SANTOS *et al.*, 2010), agentes em *drug delivery* (MOLEMA & MEIJER, 1994), efeito vasodilatador (ASSREUY *et al.*, 2009), indução do fenômeno de apoptose (BARBOSA *et al.*, 2001), dentre outros.

A ligação específica de lectinas a carboidratos tem um importante papel também na interação de plantas com outros microorganismos (PEUMANS & VAN DAMME, 1995). Existem evidências crescentes de que lectinas de plantas poderiam agir como moléculas de defesa contra insetos e outros herbívoros (VASCONCELOS & OLIVEIRA, 2004). A atividade inseticida de muitas lectinas de plantas tem sido documentada em ensaios de nutrição *in vitro* e em estudos com plantas transgênicas (CARLINI & GROSSI-DE-SÁ, 2002; MACEDO *et al.*, 2004, MACEDO *et al.*, 2007). Isso se constitui um potencial econômico contra pestes, já que, sendo as lectinas um produto metabólico primário, seus genes seriam bons candidatos para conferir resistência a insetos em linhagens transgênicas. A purificação e caracterização de lectinas de novas fontes, portanto, podem revelar genes com grande potencial para engenharia genética (KAUR *et al.*, 2006; SINGH *et al.*, 2004).

Embora lectinas apresentem a propriedade comum de se ligar, reversivelmente, a carboidratos (LIS & SHARON, 1998), elas apresentam características próprias, principalmente no que diz respeito a aplicações biológicas. Assim, mesmo lectinas que apresentem alto grau de identidade de sequências entre si, mostram-se diferentes quanto às várias propriedades biológicas. Isto faz com que, via de regra, cada lectina tenha as suas potencialidades de aplicação, o que justifica que cada uma delas, por mais semelhante que possa parecer com outra lectina, mereça ser estudada isoladamente. Fica evidente, então, que a descoberta, o isolamento e a caracterização química, físico-química e biológica de novas lectinas se revestem de grande importância, na medida em que novas lectinas com diferentes aplicabilidades podem ser

encontradas. Além disto, visto hoje em dia ser forte a corrente que defende o papel de moléculas de comunicação para as lectinas, a descoberta de novas lectinas com especificidades diferentes se torna atrativa, uma vez que podem servir de novos modelos para o entendimento destes processos dinâmicos de comunicação célula-célula, célula-microrganismo ou célula-vírus (PEUMANS & VAN DAMME, 1995).

3.1. Funções Endógenas de Lectinas de Plantas

Desde a sua descoberta as lectinas têm gerado polêmica com relação as suas possíveis funções, entretanto, mesmo para lectinas com homologia torna-se difícil atribuir uma função em comum, tendo em vista que as mesmas podem apresentar variações em relação à especificidade, localização intracelular e período de síntese (RUDIGER & GABIUS, 2001).

A capacidade de reconhecer carboidratos de forma específica permite que as lectinas atuem na decodificação de glicocódigos, representados pelos polissacarídeos e glicoconjugados. Tais moléculas podem existir tanto nas formas livre, quanto ancoradas à membrana celular ou a matriz extracelular (PEUMANS e VAN DAMME, 1998).

Em vegetais, uma das premissas mais aceitas trata sobre a possível função de defesa das lectinas em detrimento de simples reservas de nitrogênio. Essa hipótese é sustentada pelo fato de muitas dessas proteínas apresentarem a capacidade de reconhecer especificamente carboidratos inexistentes em seus vegetais de origem, mas abundantes em outros organismos (PEUMANS & VAN DAMME, 1995). Suas destacadas resistências a proteólise e estabilidade mediante variações consideráveis de temperatura e pH (VASCONCELOS & OLIVEIRA, 2004) são outras características que corroboram para essa idéia.

Essas “funções externas” das lectinas conferem aos vegetais uma maior adaptabilidade ao seu meio. Dentre elas, podemos citar:

Proteção contra herbívoros: Essa propriedade é conferida pela capacidade de algumas lectinas vegetais de interagir com glicanos presentes na superfície celular do trato digestivo de insetos e mamíferos (VASCONCELOS & OLIVEIRA, 2004). Diversos trabalhos reportam vários efeitos tóxicos e/ou antinutricionais causados por essas proteínas, como a WGA, ConBr e GSA, todas capazes de inibir o crescimento de *Callosobruchus maculatus*, o gorgulho do feijão-de-corda (GRANGEIRO *et al.*, 1999a; GRANGEIRO *et al.*, 1999b; MURDOCK *et al.*, 1990;

ZHU *et al.*,1996); além da GNA, que apresenta efeito tóxico contra o afídeo *Aulacorthum solani* (DOWN *et al.*, 1996).

Atividade antifúngica: A quitina, polímero constituído por resíduos de *N*-acetilglucosamina (GlcNac) unidos por ligações $\beta(1-4)$, é o principal constituinte da parede celular fúngica. Dessa forma, lectinas ligantes de GlcNac podem exercer efeito inibitório sobre o crescimento celular fúngico, como por exemplo a WGA (RUDIGER, 1997).

Simbiose com bactérias nitrificantes: Várias plantas, em particular as leguminosas, são conhecidas pela sua habilidade em estabelecer simbiose com bactérias capazes de fixar o nitrogênio atmosférico, como as do gênero *Rhizobium* e outros relacionados, tornando-as independentes da adição de fertilizantes nitrogenados (RUDIGER, 1984). Essa relação pode ocorrer mediante a interação entre lectinas presentes nas raízes e glicoconjugados (lipoquitooligossacarídeos) da parede celular bacteriana (VAN EIJSDEN *et al.*, 1995).

Apesar de apresentarem importantes papéis no meio externo, as lectinas também possuem funções indispensáveis no controle da homeostase vegetal. Dentre elas podemos citar o papel das lectinas como reservas de nitrogênio durante o desenvolvimento inicial da planta, já que o estoque dessas proteínas diminui progressivamente durante o crescimento da plântula. Essa função é acoplada ao possível papel na defesa, uma vez que quantidades consideráveis de lectinas presentes na semente podem ajudar a evitar que essa seja predada ou danificada, porém, se não requisitadas, são então consumidas (PEUMANS & VAN DAMME, 1995).

Outro papel das lectinas encontra-se na interação com enzimas, direcionando sua atuação a sítios específicos e ao mesmo tempo modulando sua atividade (TAIPALESUU *et al.*, 1997; WIERZBA-ARABSKA & MORAWIECKA, 1987).

3.2. Classificação Estrutural de Lectinas

As lectinas formam um grupo de proteínas muito heterogêneas. Muitas lectinas podem ser agrupadas em famílias caracterizadas por sequências similares ou organização estrutural. Van Damme *et al.* (1998) subdividiu as lectinas vegetais de acordo com os tipos estruturais em quatro classes estruturais principais: merolectinas, hololectinas, quimerolectinas e superlectinas (Figura 3).

As merolectinas são proteínas que consistem exclusivamente de um único domínio de ligação a carboidrato. Devido a seu caráter monovalente, as merolectinas são incapazes de precipitar glicoconjugados ou aglutinar células. A heveína, uma proteína do látex de seringueira (*Hevea brasiliensis*), que se liga à quitina, é uma típica merolectina (VAN PARIJS *et al.*, 1991).

As hololectinas também são compostas exclusivamente de domínios carboidrato ligantes, mas com dois ou mais sítios ligantes que são idênticos ou parecidos. Como as hololectinas são divalentes ou multivalentes, elas podem aglutinar células e/ou precipitar glicoconjugados. As hololectinas comportam-se como verdadeiras aglutininas e compreendem a maioria das lectinas de plantas, representando a classe de lectinas mais bem estudada. São exemplos típicos de hololectinas aquelas encontradas em sementes de plantas pertencentes à subtribo Diocleinae (CAVADA *et al.*, 2001).

As quimerolectinas são uma fusão de proteínas compostas de um ou mais domínios de carboidratos ligantes e de um domínio não relacionado, com uma atividade catalítica bem definida, e que age independentemente dos domínios de ligação a carboidrato. Dependendo do número de sítios de ligação a carboidratos, as quimerolectinas comportam-se como merolectinas ou hololectinas. Um representante deste grupo são as proteínas inativadoras de ribossomos (RIP tipo 2), como por exemplo a ricina (toxina da mamona), que possui dois domínios de ligação para carboidratos e um domínio para a inativação do ribossomo (PEUMANS & VAN DAMME, 1998). Outro exemplo é a PPL-2, uma lectina quitina-ligante isolada de sementes de *Parkia platycephala* que possui, além do sítio ligante a carboidrato, um sítio catalítico com atividade endoquitinásica (CAVADA *et al.*, 2006).

As superlectinas consistem de no mínimo dois domínios de ligação para carboidratos. Diferente das hololectinas os domínios de ligação para carboidratos das superlectinas reconhecem açúcares estruturalmente e funcionalmente diferentes. Como exemplo, a lectina do bulbo de tulipa (TxLCI) que é formada por dois domínios de ligação a carboidrato, reconhecendo manose e *N*-acetil galactosamina, respectivamente; e a banana-lectin, que possui dois sítios distintos de reconhecimento de laminaribiose e Xyl- β 1,3-Man- α -O-Methyl em domínios distintos (MEAGHER *et al.*, 2005).



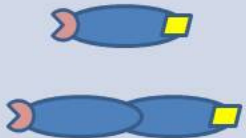
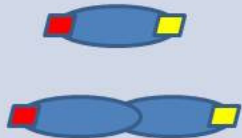




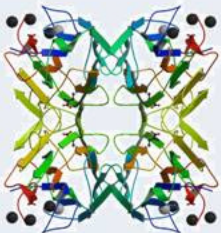

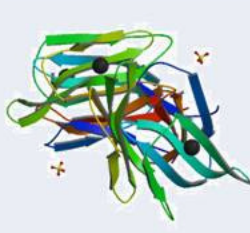




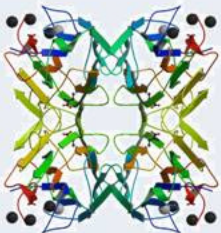

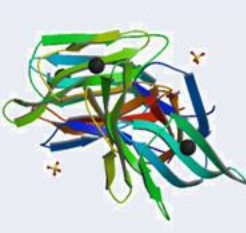
Merolectinas	Hololectinas	Quimerolectinas	Superlectinas
			
   	  	  	
			
Heveína (1HEV)	ConBr (3JU9)	PPL2 (2GSJ)	Banana lectin (2BMY)

Figura 3. Classificação estrutural de lectinas de plantas em Merolectinas, Hololectinas, Quimerolectinas e Superlectinas (Adaptado de Van Damme *et al.*, 1998). DRC: domínio de reconhecimento a carboidratos.

3.3. Classificação de Lectinas de Plantas em Famílias Evolutivamente Relacionadas

O avanço da biologia estrutural em conjunto com modernas técnicas de DNA recombinante deu um salto significativo na elucidação das mais diversas moléculas presentes em estruturas dos sistemas biológicos. Atualmente, mais de 300 estruturas de lectinas de vegetais encontram-se depositadas no *Protein Data Bank* (BERMAN *et al.*, 2000). Destas, além de uma grande maioria possuir um detalhamento estrutural a nível atômico (acima de 2,0 Å de resolução), muitas já possuem seus sítios de reconhecimento a carboidratos bem estabelecidos, apresentando semelhanças quanto a aspectos estruturais. Outras características como sequência primária, reconhecimento a açúcares e até mesmo quanto a função desempenhada, já são bem conhecidas e estudadas (VAN DAMME *et al.*, 1998).

Isso ocorreu de uma forma que tornou possível classificar lectinas não somente quanto aos aspectos estruturais vistos anteriormente, mas como também em famílias que estejam evolutivamente relacionadas, ou seja, que não possuam apenas estruturas quaternárias semelhantes, mas que compartilham outras características que possam estar diretamente

relacionadas com a função desempenhada por cada uma delas (VAN DAMME *et al.*, 1998). Van Damme e colaboradores (1998) classificaram as lectinas de plantas em sete grupos, ou famílias, de acordo com as semelhanças estruturais e filogenéticas. Dentre as famílias evolutivamente relacionadas de lectinas de plantas podemos destacar (Figura 4):

a. Lectinas de monocotiledôneas do tipo manose

O termo “lectinas de monocotiledôneas que se ligam a manose” refere-se a uma família de lectinas estreitamente relacionadas que são estritamente reconhecedoras de manose, e que são encontradas em monocotiledôneas. Foram classificadas, até então, em seis diferentes famílias de plantas: Alliaceae, Amaryllidaceae, Araceae, Bromeliaceae, Liliaceae e Orchidaceae.

Estudos comparativos indicam que essas proteínas compartilham um alto grau de similaridade de sequência primária, porém com uma completa heterogeneidade de estrutura molecular (VAN DAMME *et al.*, 1998). Um exemplar com estrutura resolvida está representado na Figura 4a.

b. Lectinas específicas a quitina e homólogas a heveína

As lectinas que se ligam a quitina são capazes de se ligar a resíduos de *N*-acetilglicosamina (principal unidade monomérica constituinte da quitina) e têm como característica principal o fato de possuir um domínio heveínico em suas moléculas (VAN DAMME *et al.*, 1998). O termo “domínio heveínico” refere-se a uma pequena proteína (uma merolectina) de 43 resíduos de aminoácidos, extraída do látex da *Hevea brasiliensis*, a popular seringueira, que reconhece quitina (WALJUNO *et al.*, 1975) e é o protótipo de proteínas dessa classe. Lectinas pertencentes a esse grupo têm sido encontradas em várias famílias de plantas não relacionadas taxonomicamente, tais como Gramineae, Leguminosae, Solanaceae, Urticaceae, Papaveraceae e Amaranthaceae. Lectinas que se ligam à quitina e não possuem o domínio heveínico são classificadas em outras famílias de lectinas. A aglutinina de germen de trigo (WGA) é a principal representante dessa classe, é composta de duas subunidades de 18 kDa consistindo de quatro domínios de Heveína estruturalmente semelhantes. (RAIKHEL *et al.*, 1993). Um exemplar com estrutura resolvida está representado na Figura 4b.

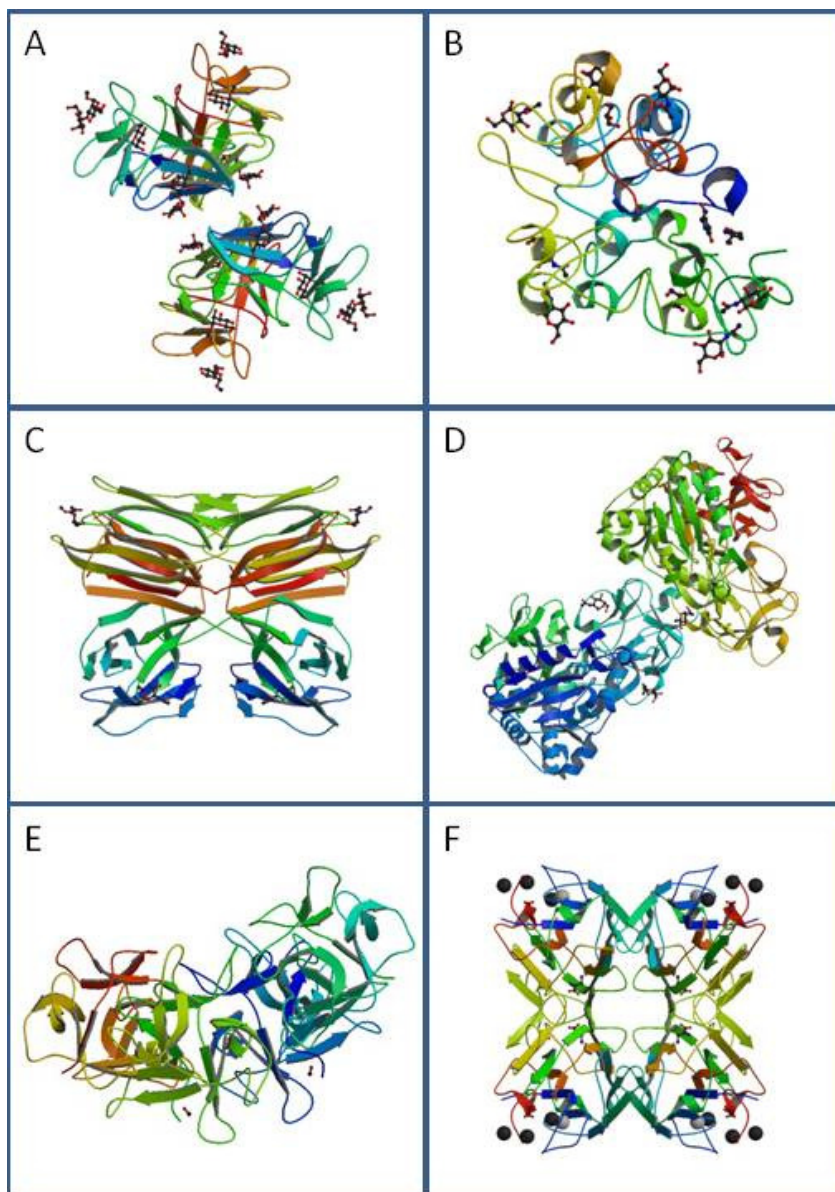


Figura 4. Classificação de lectinas de plantas em famílias evolucionariamente relacionadas: representantes com estrutura tridimensional. A. Lectinas de monocotiledôneas do tipo manose: *Galanthus nivalis* agglutinin (GNA, código PDB: 1NIV). B. Lectinas específicas a quitina e homólogas a heveína: Wheat germ agglutinin (WGA, código PDB: 2UVO). C. Lectinas homólogas a Jacalina: Jacalina (1HU0). D. Lectinas homólogas ao tipo RIP-2 (Ricina): Abrina (código PDB: 2ZR1). E. Lectinas da família das Amarantaceae: Amarantina (código PDB: 1LJY). F. Lectinas de leguminosas: *Canavalia brasiliensis* lectin (ConBr, código PDB: 3JU9). A família de lectinas de floema de Curcubitaceae não possui representante com estrutura tridimensional.

c. Lectinas homólogas a Jacalina

A expressão “lectinas relacionadas à Jacalina” é usada para nomear um grupo de lectinas que são evolucionária e estruturalmente relacionadas a aglutininas presente nas

sementes de *Artocarpus integrifolia* (jaca). Basicamente, a família da jacalina compreende dois subgrupos específicos: i) que reconhecem galactose e ii) que reconhecem manose. Esse tipo de lectina é encontrada principalmente em espécies de Moreaceae e Convolvulaceae (VAN DAMME *et al.*, 1998). Esse primeiro grupo possui uma alta afinidade pelo antígeno-T [Gal β (1,3)GalNAc] (SASTRY *et al.*, 1986; SARKAR *et al.*, 1981). Em relação as lectinas específicas por manose relacionadas a jacalina, estas foram identificadas em muitas espécies em grupos taxonomicamente diferentes. Sabe-se, até então, que as lectinas relacionadas à jacalina que se ligam a manose possuem duas ou quatro subunidades não glicosiladas de aproximadamente 150 resíduos de aminoácidos. Um exemplar com estrutura resolvida está representado na Figura 4c.

d. Lectinas homólogas ao tipo RIP-2 (Ricina)

As RIPs tipo 2 constituem uma típica família de quimerolectinas constituídas por duas cadeias polipeptídicas sintetizadas a partir de uma única molécula precursora, a qual é pós-transducionalmente processada através da excisão de uma ligação entre as cadeias. O termo “proteína que inativa ribossomos” (RIP) refere-se à capacidade destas de inativarem cataliticamente os ribossomos, que ocorre através da remoção enzimática de um resíduo de adenina específico de uma volta na grande subunidade do RNA ribossomal 28 S. A primeira dessas cadeias (cadeia A) é formada por um polipeptídeo constituído de um domínio adenosina glicosidase e a outra cadeia (cadeia B) constituída por um domínio de ligação a carboidratos, sendo que as cadeias permanecem unidas por pontes dissulfeto. Ao penetrar nas células, esta ligação dissulfeto é quebrada e a cadeia A adquire uma forte atividade *N*-glicosidase tornando-se capaz de inativar cataliticamente ribossomos eucarióticos e procarióticos (VAN DAMME *et al.*, 1998). A Ricina (*Ricinus communis*) e Abrina (*Abrus precatorius*) são exemplos clássicos de RIPs tipo 2 com potentes atividades biológicas, já tendo sido identificadas em espécies de várias famílias de plantas (PEUMANS *et al.*, 2001; VAN DAMME *et al.*, 1998). A atividade e especificidade de ligação a carboidrato das RIPs são estabelecidas somente por sítios de ligação a carboidratos situados na cadeia B. Estudos de especificidade indicaram que a maioria das RIPs são efetivamente inibidas por Galactose ou GalNAc ou por ambos. Todavia os inibidores mais potentes que quaisquer outros monossacarídeos são os dissacarídeos ou oligossacarídeos ricos

em galactose. A lectina de *Abrus pulchelus* possui especificidade por galactose, mas apresenta uma maior afinidade para sequências de lactose [β -D-Gal-(1,4)-D-Glc], *N*-acetilactosamina [β -D-Gal-(1,4)-D-GlcNAc] e lacto-*N*-biose [β -D-Gal-(1,3)-D-GlcNAc] (VAN DAMME *et al.*, 1998). Um exemplar com estrutura resolvida está representado na Figura 4d.

e. Lectinas da família das Amarantaceae

A amarantina, uma lectina extraída de *Amaranthus caudatos*, não se assemelha a nenhuma outra lectina de planta, tanto com relação a sequência de aminoácidos quanto com relação a estrutura tridimensional. Baseado nessas informações, a amarantina é considerada um protótipo da família de lectinas das Amaranthaceae. Várias outras lectinas extraídas de plantas do gênero *Amaranthus*, como por exemplo *Amarantus leucocarpus*, contêm lectinas que são muito similares à amarantina, sendo todas elas lectinas ligantes específicas por *N*-acetilgalactosamina (VAN DAMME *et al.*, 1998). Um exemplar com estrutura resolvida está representado na Figura 4e.

f. Lectinas de floema de Cucurbitaceae

As lectinas do floema de Cucurbitaceae são proteínas diméricas composta de duas subunidades idênticas de 24 kDa que se ligam a resíduos de *N*-acetilglicosamina. Todas as lectinas conhecidas dos floemas de Cucurbitaceae mostram um grau de similaridade sequencial, mas não apresentam qualquer similaridade sequencial com outras lectinas de plantas ou proteínas (VAN DAMME *et al.*, 1998). Até o presente, não existe representante com estrutura tridimensional disponível.

g. Lectinas de leguminosas

As lectinas de leguminosas correspondem à família de lectinas mais estudadas até o presente, com dezenas de representantes já purificados e caracterizados. As proteínas desse grupo caracterizam-se por possuírem uma alta similaridade estrutural, porém, uma especificidade por carboidratos e propriedades biológicas distintas. Se enquadram nesse grupo a Concanavalina-

A (ConA), extraída de sementes de *Canavalia ensiformes*, e as lectinas extraídas de outras plantas da subtribo Diocleinae (CAVADA *et al.*, 2001). Um tópico adiante irá descrever com maiores detalhes esse grupo. Um exemplar com estrutura resolvida está representado na Figura 4d.

Apesar de parecer clara, esta classificação das lectinas vegetais tende a deixar de ser baseada em aspectos evolutivos e passa a levar em conta somente os aspectos estruturais. Isso se dá pelo fato de que, à medida que novas lectinas vão sendo descobertas e caracterizadas, deixa de existir um limite taxonômico bem definido entre as famílias de lectinas vegetais. Um exemplo claro deste fato é a lectina específica para glicose e manose extraída de sementes da leguminosa *Parkia platycephala*, cuja estrutura é composta por três domínios repetidos relacionados à jacalina (MANN *et al.*, 2001; GALEGO DEL SOL *et al.*, 2005).

3.4. Lectinas de Leguminosas

Novas lectinas têm sido descobertas a cada dia em células de mamíferos, em microorganismos ou em plantas, sendo encontradas envolvidas em intrincados processos bioquímicos e biológicos. Lectinas de plantas, em particular, têm sido intensivamente estudadas em inúmeros aspectos, particularmente por seus potenciais antifúngicos e inseticidas (CARLINI & GROSSI-DE-SÁ, 2002; PEUMANS & VAN DAMME, 1996).

As lectinas de leguminosas representam o grupo mais bem estudado dentre as lectinas de plantas, sendo a maioria dos estudos de lectinas extraídas de membros da subfamília Papilionoideae, principalmente da tribo Phaseolae (MANN, *et al.*, 2001). Constituem uma ampla família de proteínas estreitamente relacionadas, a maior parte encontrada exclusivamente em espécies representativas da família Leguminosae, com exceção de sequências similares que foram encontradas em *Arabidopsis thaliana* (HERVÉ *et al.*, 1999). Elas estão presentes em sementes, folhas, caules e raízes (SHARON & LIS, 1990; LORIS *et al.*, 1998).

Até o momento, dezenas de lectinas desta família já foram isoladas e caracterizadas, principalmente lectinas de sementes, onde o percentual pode variar de 1 a 10% do total de proteínas solúveis, embora algumas espécies possam apresentar concentrações acima de 50%, ou

abaixo de 0,1% do total de proteínas solúveis presentes nas sementes. (SHARON & LIS, 1990; VAN DAMME, *et al.*, 1998).

3.4.1. Características gerais

Sabe-se que algumas lectinas de Leguminosas, especialmente as da subtribo Diocleinae, são sintetizadas no retículo endoplasmático na forma de pré-pro-lectinas ou pro-lectina glicosilada (Figura 5). A região correspondente ao cDNA contém 29 resíduos de uma sequência sinal, que é removido posteriormente durante o transporte da pré-pro-lectina para o lúmen do retículo endoplasmático, originando uma pró-lectina. Em seguida encontra-se uma região codificante referente aos aminoácidos 119-237 da Concanavalina A, e uma outra região codificante correspondente aos aminoácidos 1-118 da lectina, seguida por último por uma extensão C-terminal de 9 resíduos. Além disso, há entre os dois peptídios uma unidade oligomanose *N*-ligada a uma pentadecapeptídeo interno, o qual é excisado durante a maturação, produzindo os dois peptídios. Acredita-se que a presença da glicosilação iniba a atividade de ligação a carboidrato da lectina. A lectina madura e ativa, portanto, é não-glicosilada. A maior parte dos dois peptídios formados são então covalentemente ligados por uma reação de transpeptidação, concomitante com a clivagem da extensão C-terminal, formando a lectina madura, onde o alinhamento dos resíduos 1-118 e 119-237 é o inverso de como se apresenta no precursor (SHARON & LIS, 1990; SHARON, 2007).

Do ponto de vista estrutural, as lectinas de leguminosas são geralmente compostas de 2 ou 4 subunidades, iguais ou diferentes com massa molecular em torno de 25 a 30 kDa, compostas geralmente de uma cadeia polipeptídica simples com cerca de 250 aminoácidos, cuja união é estabelecida por forças não covalentes como ligações de hidrogênio, interações hidrofóbicas e eletrostáticas, que formam dímeros canônicos e estabilizam o tetrâmero pela união destes dímeros. Cada uma destas subunidades apresenta um único sítio de ligação a carboidratos com a mesma especificidade, além de um ou mais sítio de ligação a íons, podendo ainda apresentar um ou dois oligossacarídeos *N*-ligados. Entretanto, algumas lectinas possuem subunidades formadas por 2 cadeias polipeptídicas, entre elas as lectinas da tribo Viciae, gêneros *Pisum*, *Vicia*, *Lathyrus*, e *Lens*. Estas lectinas são dímeros formados por duas

subunidades iguais e cada subunidade é constituída de uma cadeia α (5 a 7 kDa) e uma β (15 a 19 kDa) mantidas por ligações não-covalentes (SHARON & LIS, 1989).

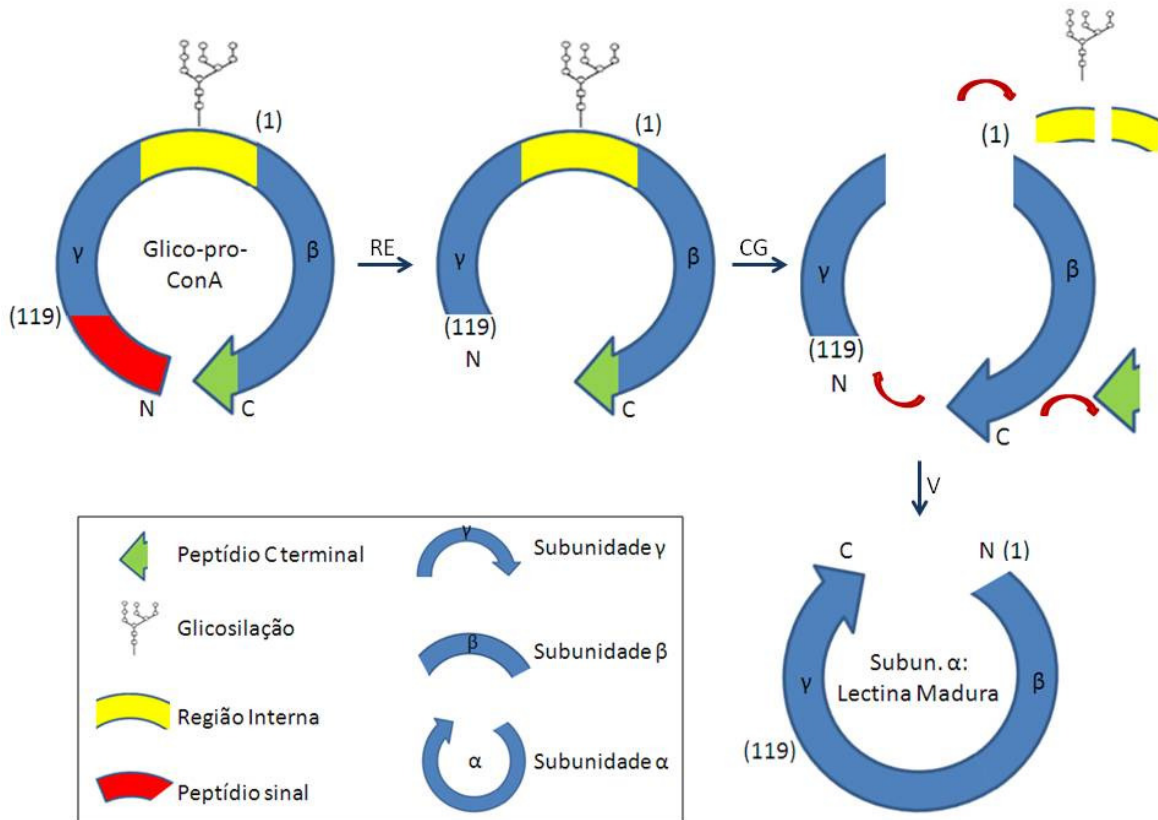


Figura 5. Modificações pós-traducionais durante a biossíntese da Concanavalina A (ConA). Sumário dos eventos de processamento convertendo a pro-concanavalina A glicosilada (pré-pro-conA) em lectina madura. Extremidades amino e carboxi terminal da lectina madura estão indicados por N e C, respectivamente, e o número em parênteses refere-se a resíduos da concanavalina A madura. Durante o processamento na planta, a pro-proteína glicosilada inativa é deglicosilada resultando no surgimento da atividade de ligação a carboidrato. A ação de uma endopeptidase que cliva o nonapeptídeo carboxiterminal e o espaçador da deglicosilação está mostrado em verde e amarelo, respectivamente. Resíduos 118 e 119 são ligados enzimaticamente. *Splicing* então resulta em uma transposição do arranjo linear dos domínios protéicos γ e β (Adaptado de CARRINGTON, 1985). RE: retículo endoplasmático, CG: complexo de golgi, V: vacúolo.

Algumas lectinas de leguminosas podem ser encontradas em diferentes isoformas, como, por exemplo, a lectina de *Phaseolus vulgaris* (PHA) que é formada por duas subunidades gênicas diferentes, denominadas E (31,7 kDa) e L (29,9 kDa), que combinadas dão origem a uma família de cinco isolectinas tetraméricas: E4, E3L, E2L2, EL3, L4 (FELDSTED *et al.*, 1977). Além da heterogeneidade genética, ou seja, mais de um gene codificando uma lectina, a

ocorrência de isoformas pode ser atribuída a processamentos pós-traducionais incompletos (YOUNG *et al.*, 1995).

Outras lectinas de diferentes partes vegetais existem como uma complexa mistura de isoformas, como por exemplo, as lectinas de *Robinia pseudoacacia* (VAN DAMME *et al.*, 1995), do gênero *Erythrina* (BONNEIL *et al.*, 2004), de *Acacia constricta* (GUZMAN-PARTIDA *et al.*, 2004) e de *Amaranthus leucocarpus* (HERNANDEZ *et al.*, 2001). A presença de isoformas têm alto impacto sobre a função biológica de lectinas, e a existência de várias isoformas poderiam oferecer uma estratégia alternativa ou uma adaptação evolutiva para compensar a baixa especificidade.

3.4.2. Características estruturais

As estruturas cristalinas de um grande número de lectinas de leguminosas já foram resolvidas (210 estruturas disponíveis de 35 lectinas diferentes) (VARROT *et al.*, 2011). As estruturas dos monômeros de diferentes lectinas de leguminosas são extremamente similares (Figura 6). Aproximadamente 20% dos resíduos de aminoácidos são invariantes em todas as lectinas de leguminosas e outros 20% são similares. Entre os aminoácidos conservados incluem-se aqueles envolvidos na interação com o carboidrato e quase todos os resíduos que coordenam íons metais (BANERJEE *et al.*, 1996; HAMELRYCK *et al.*, 1998). A resolução de estruturas tridimensionais de lectinas de leguminosas tem mostrado que cada subunidade é constituída de cerca de 60% de fitas-beta antiparalelas arranjadas em duas ou mais folhas beta interconectadas por voltas, mais conhecido como enovelamento do tipo “ β -sanduíche”. Para todas as lectinas de leguminosas conhecidas, a estrutura terciária é constituída de duas folhas- β anti-paralelas, uma com seis fitas β (atrás) e outra com sete fitas β (frente), conectadas por uma outra folha- β de 5 fitas (superior), gerando o conhecido motivo "jellyroll", também referido como "lectin fold". (SRINIVASAN *et al.*, 1996; HAMELRYCK *et al.*, 1998).

O sítio de ligação a carboidratos está localizado no lado côncavo do β -sanduíche, próximo ao sítio de ligação a metais. O sítio de reconhecimento a carboidratos consiste de diversas voltas com diferentes graus de variabilidade (SHARMA & SUROLIA, 1997). As conformações destas voltas são determinadas pela presença na estrutura de íons de metais de transição, como o cálcio e o manganês (LORIS *et al.*, 1998; BOUCKAERT *et al.*, 2000). O sítio

de ligação a íons é amplamente conservado nestas proteínas. Sabe-se que a atividade de reconhecimento a açúcar depende da presença do íon cálcio e de um outro íon, geralmente o manganês. Esses metais se localizam a uma distância de aproximadamente 4,5 Å e ambos são coordenados por cadeias laterais de pelo menos quatro aminoácidos e duas moléculas de água. O reconhecimento das lectinas de leguminosas por íons é reversível e a remoção desses metais resulta em importantes mudanças conformacionais. Em ConA, por exemplo, com a falta dos íons cálcio e manganês, o peptídeo Ala207-Asp208, que para participar efetivamente do reconhecimento ao seu carboidrato específico necessita estar na conformação trans, passa a assumir uma conformação do tipo cis (BOUCKAERT *et al.*, 1995). A ausência desses metais, portanto, resulta em uma instabilidade local e na perda da capacidade de ligar-se a carboidratos (LORIS *et al.*, 2004).

O sítio primário de ligação a carboidratos das lectinas de leguminosas consiste em uma região na superfície da molécula composto por quatro voltas que proporcionam a formação de uma cavidade na superfície da mesma, onde o reconhecimento específico de carboidratos se dá por intermédio de ligações de hidrogênio entre a lectina e o açúcar. Nestas voltas encontram-se três resíduos altamente conservados, Asp, Asn e Gly/Arg, os quais são responsáveis pela formação de quatro ligações de hidrogênio com as hidroxilas dos carbonos três e quatro presentes em monossacarídeos como manose ou glicose. Também participam do processo de reconhecimento interações de Van der Waals estabelecidas entre o anel aromático do açúcar e as cadeias laterais de aminoácidos hidrofóbicos: a estabilização dos monossacarídeos ocorre também através de interações hidrofóbicas com os resíduos de aminoácido como Phe, Tyr, Trp ou Leu (Moreno, 2008). Contatos adicionais entre o carboidrato e a superfície da proteína (diretos ou mediados por moléculas de água) contribuem para a alta seletividade (WEIS & DRICKAMER, 1996; ELGAVISH & SANAN, 1997). Apesar da riqueza de dados obtidos por cristalografia por raios X, um número relativamente pequeno de lectinas apresentam-se complexadas com oligossacarídeos longos (BERMAN *et al.*, 2002; <http://www.cermav.cnrs.fr/lectines/>).

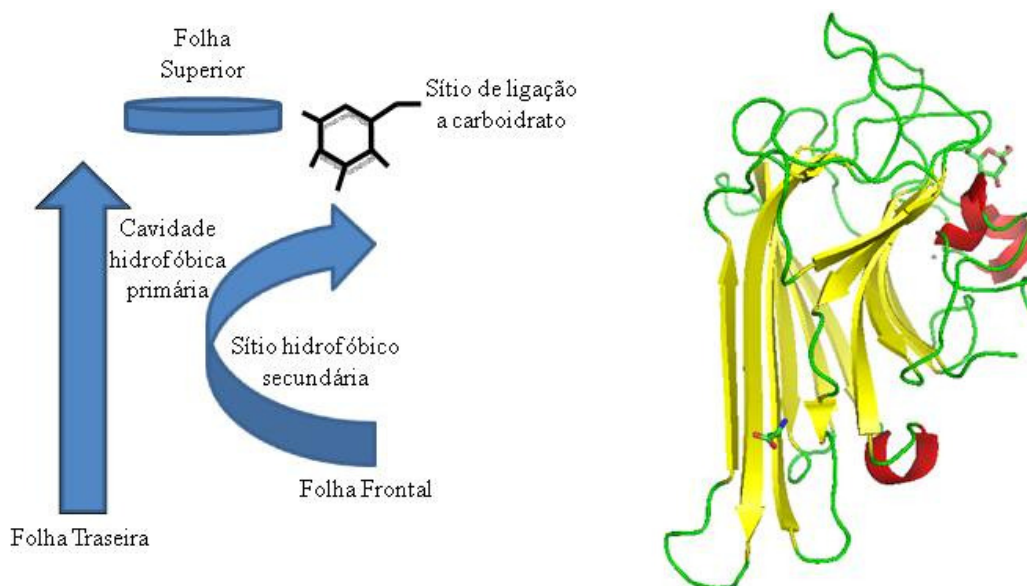


Figura 6. Monômero típico de uma lectina de leguminosa. À esquerda, diagrama esquemático da estrutura terciária de um monômero típico de lectinas de leguminosas. Do lado direito, estrutura terciária do monômero da lectina de *Canavalia gladiata* (CGL) complexada com α -metil-manosídeo e ácido aminobutírico (Abu) (ambos em verde) (código PDB:2D7F).

Além de um sítio de ligação a carboidratos, altamente conservado, juntamente com um sítio de ligação a metais, que é responsável pela estabilização da volta que compõe o sítio de ligação a carboidratos, estudos recentes mostraram a capacidade dessas proteínas interagirem com compostos hidrofóbicos. Algumas lectinas de leguminosas podem interagir com outras espécies de moléculas, que não carboidratos, tais como adenina (HAMELRYCK *et al.*, 1999) e ácido aminobutírico (Abu) (Figura 7) (DELATORRE *et al.*, 2007). A ligação das lectinas a essas moléculas se dá através de uma região estrutural hidrofóbica, altamente conservada em muitas lectinas de leguminosas e que, devido a esse alto grau de conservação, pode estar envolvida em algum papel biológico importante (DELATORRE *et al.*, 2007).

A despeito de suas similaridades ao nível monomérico primário, secundário, terciário e quaternário, lectinas de leguminosas exibem considerável variação em sua estrutura quaternária: pequenas diferenças na sequência de aminoácidos na interface monômero-monômero e a presença ou ausência da glicosilação afetam os modos de associação dos monômeros (SRINIVAS *et al.*, 2001). Em geral, lectinas de leguminosas são essencialmente dímeros de dímeros, e diferentes modos de associação são a consequência de uma combinação diversa de interfaces diméricas vistas nessas estruturas quaternárias. (MORENO *et al.*, 2006; DELATORRE *et al.*, 2007).

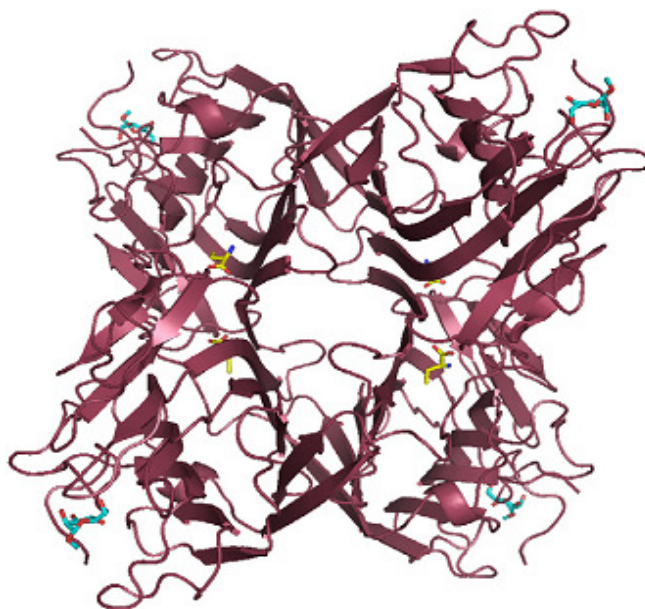


Figura 7. Estrutura cristalina da lectina de *Canavalia gladiata* (CGL) complexada com α -metil-manosídeo (em azul). Ácido aminobutírico (Abu, em amarelo) pode ser visto em cada dímero canônico entre os monômeros na cavidade hidrofóbica (DELATORRE *et al.*, 2007).

Adicionalmente, o modo de tetramerização de lectinas de leguminosas está intimamente relacionado a sua afinidade por oligossacarídeos complexos. Diferenças nas propriedades de ligação-cruzada de algumas lectinas em processos de reconhecimento celular ou transdução de sinais são refletidos em propriedades biológicas (DAM *et al.*, 2005). Então, um detalhado entendimento dos mecanismos de interações biológicas entre lectinas e carboidratos complexos é a chave para o estabelecimento de bases químicas de importantes funções de reconhecimento lógico desempenhado por lectinas de plantas (MORENO *et al.*, 2008).

Algumas lectinas de leguminosas, como as lectinas extraídas de espécies da subtribo Diocleinae, exibem uma oligomerização dependente de pH, onde a proporção entre proteínas no estado dimérico e no estado tetramérico pode ser alterada de acordo com pH do meio (CALVETE *et al.*, 1999; NAGANO *et al.*, 2008). A alteração desse equilíbrio dímero/tetrâmero dependente de pH pode influenciar diretamente nas atividades biológicas dessas lectinas, pelo fato dessas proteínas serem capazes de se ligar aos receptores glicoconjugados da superfície das membranas com uma maior afinidade quando na forma tetramérica (DELATORRE *et al.*, 2006).

Muitas lectinas de leguminosas são expressas nas formas glicosiladas e não glicosiladas. É o caso da lectina de *Vatairea macrocarpa* que apresenta cadeias ativas com um único ponto ou dois pontos de glicosilação, além de fragmentos, também ativos com frações

glicosilados e não glicosiladas (CALVETE *et al.*, 1998). Algumas variações também são encontradas nas lectinas de sementes de *Phaseolus vulgaris*, *Griffonia simplicifolia* e *Vicia villosa* (VAN DAMME *et al.*, 1998). Alguns pesquisadores citam que a glicosilação pode ser uma estratégia celular para que novas estruturas tridimensionais surjam devido a impedimento espacial causado pela presença do glicano (SHAANAN *et al.*, 1991). Outros relatam que a glicosilação em lectinas pode estar envolvida em outras funções celulares, como o envelhecimento diferenciado (TURTON *et al.*, 2004).

3.4.3. Especificidade de Lectinas de Leguminosas

As lectinas de leguminosas possuem capacidade de interagir com açúcares simples e complexos (DAM & BREWER, 2002). Essa interação depende, exclusivamente, de um sítio de ligação primário a carboidratos presente na superfície molecular das mesmas. Esse sítio interage diretamente com um único açúcar simples, daí sua denominação. Diversos grupos de lectinas foram descritos quanto à especificidade por monossacarídeos. Dentro das lectinas de leguminosas, os principais grupos de especificidades encontrados, de acordo com RUDIGER (1997), são: manose, glicose e seus α -glicosídeos; galactose e *N*-acetilgalactosamina; *N*-acetilglucosamina; α -L-fucose; e oligossacarídeos complexos.

Entretanto, a maioria das lectinas, que compartilham entre si a característica de reconhecerem primariamente o mesmo tipo de monossacarídeo, diferem entre si em relação a afinidade por oligossacarídeos. Em outras palavras, isso quer dizer que a interação com monossacarídeos é bem menos complexa do que por oligossacarídeos. Portanto, lectinas de plantas leguminosas, embora pertençam a uma mesma família de proteínas, apresentam notável divergência de estrutura quaternária e de especificidade por carboidratos. A elucidação do mecanismo pelo qual essas lectinas possuem uma ampla gama de especificidades de ligação enquanto mantêm uma estrutura monomérica e tridimensional muito similar é a chave para o entendimento da essência das interações carboidrato-proteína (KANEDA *et al.*, 2002; BRINDA *et al.*, 2005). Desta forma, a busca e a caracterização de lectinas inéditas possibilitam a descoberta de novas propriedades funcionais e aplicações.

3.4.4. Propriedades químicas e biológicas de lectinas de leguminosas

De acordo com ELGAVISH & SHAANA (1998), carboidratos interagem com lectinas através de ligações de hidrogênio, coordenação por moléculas de água, Forças de Van Der Waals e interações hidrofóbicas. O grande número de hidroxilas presentes nos açúcares em geral os permite interagir através de formações de ligações de hidrogênio (hidroxilas funcionando como doadores ou receptores de hidrogênio).

A capacidade de interação com carboidrato em uma lectina é atribuída a uma porção limitada da molécula, denominada domínio globular de reconhecimento de carboidrato (CRD), geralmente com menos de 200 aminoácidos. Os CRDs de muitas lectinas conhecidas se relacionam quanto à sequência primária (WEIS & DRICKAMER, 1996).

Demonstrou-se já que lectinas de especificidades semelhantes por carboidratos podem, em alguns casos, apresentar afinidade pelo ligante específico de maneira diferenciada, como se alguma pequena diferença estrutural entre as lectinas pudesse ocasionar diferenças de afinidade em cada complexo estabelecido entre lectina-ligante, e isto se referir na potência das atividades biológicas desempenhadas por estas proteínas (CAVADA *et al.*, 2001).

A conservação estrutural das lectinas de leguminosas mostra que realmente essas proteínas estão evolutivamente relacionadas, podendo ser relatadas como proteínas homólogas que surgiram a partir de um único ancestral comum. Contudo, a principal distinção entre as lectinas pertencentes à classe das leguminosas está relacionado à afinidade fina por açúcares. Inicialmente acreditava-se que essas proteínas tinham afinidades apenas por monossacarídeo e que este, por sua vez, era o principal mediador de toda a interação lectina-açúcar. Todavia, novos experimentos apontaram que conforme houvessem mudanças nas cadeias subsequentes dos glicanos, diferentes valores para K_a eram observados nos complexos lectina-açúcares. Isso indica que apesar de diferentes lectinas se ligarem especificamente a um único tipo de açúcar, não quer dizer que uma cadeia mais complexa composta de diversos açúcares não possa também interagir em sítios estendidos, adjacentes ao sítio primário. Isso foi interessante no sentido de se entender que algumas lectinas possuem especificidades diferentes em relação a diferentes glicanos e isso pode ser obviamente extrapolado para se entender o porquê que uma única lectina pode apresentar diferenças em suas atividades biológicas quando utilizadas em diferentes meios celulares (DAM & BREWER, 2002).

As possibilidades do uso de lectinas vegetais como ferramentas biotecnológicas são consideradas devido ao grande número de trabalhos científicos exibindo atividades biológicas relevantes para essas moléculas (KITADA *et al.*, 2010; KIMBLE *et al.*, 2010). A propriedade que muitas lectinas possuem de aglutinar células talvez seja seu aspecto imunológico mais importante. Devido à multivalência de reconhecimento de carboidratos, muitas lectinas fazem ligações cruzadas entre células e seus receptores (DUBOIS *et al.*, 1998). Exemplos importantes advindos dessa propriedade, dentre outras, vêm sendo descritas para lectinas extraídas de espécies de leguminosas (Tabela 1).

Tabela 1. Exemplos de atividades biológicas descritas para lectinas extraídas a partir de espécies de leguminosas

Lectina	Atividade Biológica	Referência
Dgui (<i>Dioclea guianensis</i>)	Atividade antifúngica	ARAUJO-FILHO, <i>et al.</i> , 2010
ConBr (<i>Canavalia brasiliensis</i>), CFL (<i>Cratylia floribunda</i>), Dgui, DGL (<i>D. grandiflora</i>), Dvir (<i>Dioclea virgata</i>)	Efeito tóxico sobre moluscos	SANTOS <i>et al.</i> , 2010
CGL (<i>Canavalia gladiata</i>)	Atividade antiinflamatória e analgésica	NUNES <i>et al.</i> , 2009
ConBol (<i>Canavalia boliviana</i>)	Atividade antinociceptiva	FIGUEIREDO <i>et al.</i> , 2009
ConBr, CGL, ConM (<i>Canavalia marítima</i>)	Efeito vasodilatador	ASSREUY <i>et al.</i> , 2009
ConM	Relaxação da aorta e liberação de óxido nítrico	GADELHA <i>et al.</i> , 2005
ConBr, DGL, DVL (<i>D. Violacea</i>)	Ativação de linfócitos	BARBOSA <i>et al.</i> , 2001
ConBr	Atividade antidepressiva	BARAUNA <i>et al.</i> , 2006
ConA, ConBr, CFL, DVL, DGL	Interferências no processo de formação de biofilmes bacterianos	TEIXEIRA <i>et al.</i> , 2006

4. Lectinas de Dalbergiae

A subtribo Dalbergiae, pertencente a subfamília Papilionoideae, família Leguminosae, compreende 48 gêneros e cerca de 1200 espécies. Até o presente momento foram relatadas a presença de lectinas com especificidades distintas em 11 espécies dessa subtribo (Tabela 2).

Tabela 2. Especificidade por açúcares das lectinas da subtribo Dalbergiae

Lectina	Especificidade	Referência
<i>Andira fraxinifolia</i>	Man	RANGEL <i>et al.</i> , 2009
<i>Andira retusa</i>	Man	NASCIMENTO <i>et al.</i> , 2002
<i>Lonchocarpus araripensis</i>	GlcNAc	PIRES <i>et al.</i> , 2008
<i>Lonchocarpus carpassa</i>	Gal/GalNAc/Glc	JOUBERT <i>et al.</i> , 1986
<i>Lonchocarpus sericeus</i>	GlcNAc	ALENCAR <i>et al.</i> , 1999
<i>Machaerium acutifolium</i>	GlcNAc /Glc	BEZERRA <i>et al.</i> , 2003
<i>Platymiscium floribundum</i>	Man/ GlcNAc	PEREIRA-JÚNIOR, 2008
<i>Platypodium elegans</i>	Man/Glc	BENEVIDES <i>et al.</i> , 2008
<i>Pterocarpus angolensis</i>	Man/Glc	LORIS <i>et al.</i> , 2003
<i>Pterocarpus rotundifolius</i>	Man/Glc	MARONDEDZE <i>et al.</i> , 2004
<i>Vatairea macrocarpa</i>	Gal/GalNAc	CAVADA <i>et al.</i> , 1998

A lectina extraída de *Vatairea macrocarpa* é uma proteína ligante a galactose/*N*-Acetilgalactosamina, formada por 4 subunidades de 26 kDa, que formam um tetrâmero em solução independente de pH (RAMOS *et al.*, 2000). Seus efeitos biológicos descritos incluem indução da infiltração de leucócitos em edema de pata (ALENCAR *et al.*, 2004), liberação de mediadores quimiotáticos por macrófagos (ALENCAR *et al.*, 2007), aumento da resistência vascular renal, taxa de filtração glomerular e fluxo urinário (MARTINS *et al.*, 2005), migração de neutrófilos *in vivo* por mecanismo indireto (ALENCAR *et al.*, 2003) e apresenta especificidade para resíduos de antígeno do tipo Tn (DAM *et al.*, 2007).

Algumas atividades biológicas também foram evidenciadas em uma lectina extraída de sementes de *Lonchocarpus sericeus*, como alta atividade antiinflamatória em modelo de edema

de pata (ALENCAR *et al.*, 1999), efeito antiinflamatório e antibacteriano em um modelo de peritonite infecciosa (ALENCAR *et al.*, 2005), diminuição da migração leucocitária e hipernocicepção mecânica pela inibição da produção de citocinas e quimiocinas (NAPIMOGA *et al.*, 2007).

Segundo Pires *et al.* 2008, a lectina de *Lonchocarpus araripensis* (LaL) possui ação antiinflamatória sobre o edema de pata induzido por carragenina, mas é desprovida de ação pró-inflamatória quando injetada localmente.

A lectina glicose/manose específica de *Platypodium elegans* (PEL) apresentou atividade termiticida contra operários e soldados de *Nasutitermis corniger*, além de potencial fungicida contra os fungos *Fusarium solani*, *F. oxysporium* e *F. lateritium* (BENEVIDES *et al.*, 2008). Essas duas atividades reforçam a participação de lectinas na defesa vegetal, atuando como fitopatógenos.

Dentre as lectinas da subtribo Dalbergiae, a lectina Man/Glc específica de *Pterocarpus angolensis*, denominada PAL, é a única que apresenta sua estrutura tridimensional determinada (LORIS *et al.*, 2003). A sua estrutura contém uma volta clássica de especificidade a manose, mas sua volta de ligação a metal se assemelha ao de lectinas com especificidade não-relacionada, como *Ulex europaeus* e *Maackia amurensis*. As interações com manose no sítio de ligação primário são conservadas, mas detalhes da extensão desse sítio de ligação a carboidratos difere de outros vistos em complexos a carboidratos equivalentes em ConA (Figura 8). Essas observações explicam as diferenças em seu respectivo perfil de especificidade fina a oligomanoses. (LORIS *et al.*, 2004). As estruturas cristalinas de PAL em complexo com uma série de oligomanoses e N-glicanos foram resolvidas (LORIS *et al.*, 2003, BUTS *et al.*, 2006, GARCIA-PINO *et al.*, 2007). Os oligomanoses se ligam de maneira preferencial pela sequência tetrassacarídica de manose no braço 1-6. Experimentos de ITC mostraram afinidades de 150 μ M pelo oligomanose M5, 16 μ M pelo complexo biantenado GN2M3 e 12 μ M pela glicoproteína aglutinadora de soja.

Outros estudos realizados investigaram a interação entre o sitio de ligação de metal, atividade de reconhecimento e ligação a carboidratos, estabilidade e estrutura da lectina de *P. angolensis*. Estes resultados juntos com dados termodinâmicos contribuem para o melhor entendimento das interações lectina-carboidrato (GARCIA-PINO *et al.*, 2006; GARCIA-PINO *et al.*, 2007; BUTS *et al.*, 2006).

Estudos estruturais de lectinas da subtribo Dalbergieae ainda são carentes, o que torna limitante a tentativa de se descrever certos mecanismos de ação baseado na afinidade diferencial frente a ligantes específicos. Além disso, a presença de isoformas e glicofomas impedem o estudo individual de cada espécie, tanto do ponto de vista estrutural como da caracterização biológica.

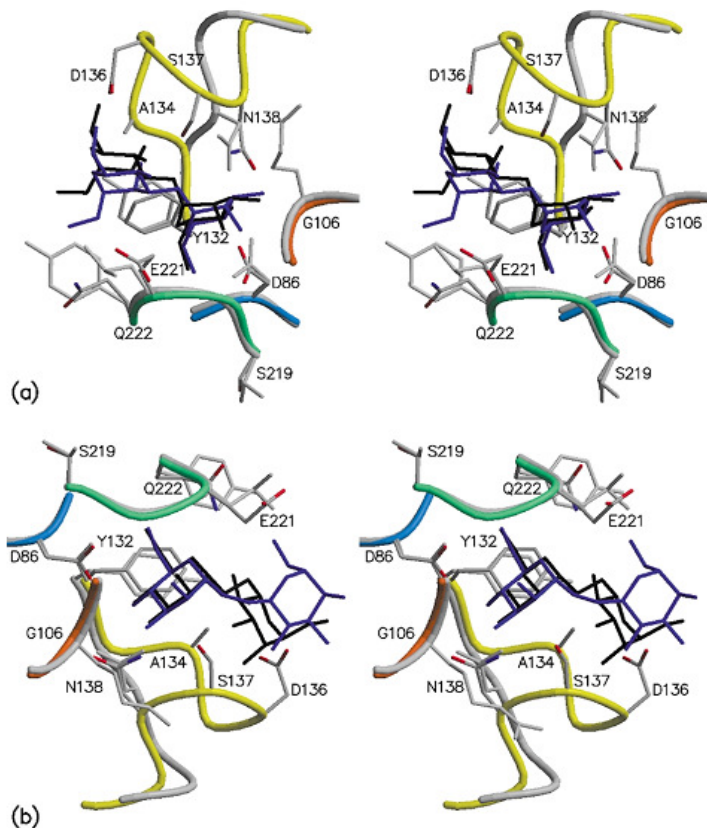


Figura 8. Reconhecimento de Man(α 1-3)Man e Man(α 1-6)Man por *Pterocarpus angolensis* lectin (PAL). A. Visão estereo do sítio de ligação a carboidrato em complexo com Man(α 1-3)Man. O dissacarídeo é mostrado em azul e as voltas de PAL estão nas cores laranja e amarelo. A volta equivalente e as cadeias laterais do complexo ConA-Man(α 1-3)Man estão mostrados em cinza claro. O dissacarídeo ligado a ConA é mostrado em preto. B. Visão estereo similar do complexo PAL-Man(α 1-6)Man superposto sobre o complexo ConA-Man(α 1-6)Man (Adaptado de Loris *et al.*, 2004).

5. Lectina de sementes de *Platypodium elegans*

O gênero *Platypodium*, da tribo Dalbergieae, compreende duas espécies de árvores leguminosas. *Platypodium elegans* Vog. (Figura 9) é uma espécie nativa de ocorrência natural no Cerrado do Brasil e em áreas de transição entre cerrado e floresta estacional,



Figura 9. *Platypodium elegans*. A. Representação Esquemática de folhas e frutos B. Representação esquemática da exsiccata. C. Detalhes do fruto imaturo e maduro. D. Copa. Floração amarela, folhas compostas paripinadas. E. Tronco. F. Fruto imaturo.

encontrada em terrenos bem drenados (LORENZI, 1998). Sua madeira, moderadamente pesada, dura, porém de tecido frouxo, durável e resistente quando em ambientes internos, é empregada na carpintaria, marcenaria, obras internas, cabos de ferramentas, etc.; além disso, esta espécie apresenta potencial ornamental, sendo utilizada na arborização de ruas e avenidas; e, como planta pioneira e rústica, não pode faltar nos plantios mistos destinados a recomposição de áreas degradadas de preservação permanente (SCHORR *et al.*, 2004). Não há nenhuma aplicação medicamentosa tradicional reportada. Essa espécie já foi reportada como nodulada por HALLIDAY & NAKAO (1982).

Filogenia:

Família Leguminosae (Fabaceae)

Subfamília Papilionoideae

Tribo Dalbergieae DC.

Gênero *Platypodium* Vogel

Platypodium elegans Vogel

De acordo com BENEVIDES (2008), sementes de *Platypodium elegans* possuem pelo menos uma lectina específica por manose/glicose e seus derivados (dados sobre inibição da atividade hemaglutinante já definiram certa especificidade por glicose, manose, GlcNac e Metil- α -D-Glucopiranosídeo). Alguns dados deste trabalho descrevem o promissor potencial fungicida e termiticida da fração lectínica pré-purificada em DEAE-Sephacel. A presença de isoformas e a dificuldade em separá-las têm limitado o estudo dessa lectina.

6. Clonagem, expressão heteróloga e purificação de lectinas de leguminosas

A dificuldade de separação de diferentes isoformas de lectinas por técnicas convencionais tem levado a definição de propriedades funcionais e atividades biológicas para uma mistura de proteínas (ECK *et al.*, 1999; RAEMAEEKERS *et al.*, 1999), impedindo um estudo detalhado das propriedades de muitas lectinas. A expressão de lectinas de plantas em bactéria demonstra ser um meio de seleção satisfatório para o estudo isolado de cada isoforma (LUO *et al.*, 2005). A expressão de lectinas vegetais em sistemas heterólogos oferece um meio de se produzir proteínas de sequências definidas de aminoácidos, e permite que se trace uma relação específica entre sequência e função através de mutações sítio dirigidas (RAEMAEEKERS *et al.*, 1999). Além

disso, essa estratégia visa solucionar problemas em função da dificuldade de se obter essas proteínas diretamente da planta, uma vez que este processo inclui etapas laboriosas e onerosas de extração e purificação e pode não garantir alto rendimento da proteína. Nesse contexto, a clonagem dos genes que codificam para essas proteínas pode ser considerada uma etapa crucial para a obtenção e produção destas proteínas em larga escala (LUO *et al.*, 2005).

Escherichia coli é o organismo mais utilizado para a produção de proteínas recombinantes em função de seu rápido crescimento, que alcança altas densidades celulares em substratos de baixo custo, por sua genética e proteômica bem caracterizadas e pela disponibilidade de numerosos vetores de clonagem (LUTOMSKI *et al.*, 2004). No caso de lectinas recombinantes, *E. coli* também é o sistema heterólogo mais utilizado (BRANCO *et al.*, 2004; WANG *et al.*, 2004). Além da expressão em *E. coli*, muitas lectinas de planta têm sido produzidas com sucesso em leveduras, como as lectinas de leguminosas PHA (*Phaseolus vulgaris* agglutinin) e GNA (*Galanthus nivalis* agglutinin) (RAEMAEEKERS *et al.*, 1999) e a lectina da leguminosa *Canavalia brasiliensis* (ConBr) em *Pichia pastoris* (BEZERRA *et al.*, 2006).

Diversos experimentos envolvendo a clonagem e expressão heteróloga de lectinas de leguminosas, algumas das quais sob a forma de precursores, têm possibilitado a realização de análises sobre a estrutura e função dessas moléculas. A expressão do precursor da lectina de *Canavalia ensiformis* (ConA) em *E. coli* forneceu diversos relatos a respeito da funcionalidade e das modificações sofridas pelas lectinas da subtribo Diocleinae durante seu processo de maturação (YAMAUCHI & MINAMIKAWA, 1990; MIN & JONES, 1992; MIN *et al.*, 1992; MIN & JONES, 1994; DINCTURK *et al.*, 2001). Outro exemplo é a modificação da especificidade por açúcar da lectina de *Erythrina corallodendron*: um mutante da lectina de *Erythrina corallodendron* foi engenhado com o objetivo de aumentar a afinidade por *N*-Acetilgalactosamina. Um resíduo de tirosina próximo ao sítio de ligação da lectina foi mutado para glicina para facilitar interações mais fortes entre o grupo acetamido do açúcar e a lectina, prevenindo a cadeia lateral da tirosina na lectina selvagem. Como resultado, a lectina mutante apresentou uma maior afinidade para *N*-Acetilgalactosamina comparada com a galactose, em contraste com a afinidade quase igual para os dois açúcares com a lectina tipo selvagem (THAMOTHARAN, 2011).

Nas últimas duas décadas, o rápido desenvolvimento das técnicas em engenharia genética (tecnologia do DNA recombinante) tem alavancado as pesquisas envolvendo lectinas vegetais por possibilitar o acesso a sequências completas de cDNA de várias dessas proteínas. Diversas delas têm sido expressas em sistemas heterólogos além de bactérias (*Escherichia coli*), como leveduras (*Pichia pastoris* e *Saccharomyces cerevisiae*), plantas (*Nicotiana tabacum*), ou mesmo células de mamíferos (células de rins de macaco). A clonagem e expressão recombinante de lectinas vegetais servem as seguintes propostas: determinação de suas estruturas primárias; análise de sua genética, filogenética e biosíntese; avaliação do papel de diferentes resíduos de aminoácidos no reconhecimento de carboidratos e na associação de monômeros; produção de lectinas com novas especificidades e; para estudo das funções das lectinas nos organismos vegetais.

A clonagem desses e de vários outros genes tem sido facilitada pela disponibilidade de kits comerciais para essa proposta. Nos últimos anos, diversas lectinas de leguminosas têm sido clonadas e expressas em sistemas heterólogos (Tabela 3) (STREICHER & SHARON, 2003).

Tabela 3. Lectinas de Leguminosas clonadas e expressas em *E. coli*

Espécie	Símbolo	Especificidade	Rendimento	Referência
<i>Arachis hypogea</i>	PNA	GalNAc/Gal	12 µg/L	SHARMA & SUROLIA, 1994
<i>Artocarpus incisa</i>	Frutalina	Gal	16 mg/L	OLIVEIRA <i>et al.</i> 2009
<i>Bauhinia purpurea</i>	BPA	Gal	ND*	KUSUI <i>et al.</i> , 1991
<i>Canavalia brasiliensis</i>	Prepro ConBr	Man/Glc	30 mg/L	NOGUEIRA <i>et al.</i> , 2002
<i>C. ensiformis</i>	Prepro ConA	Man/Glc	1 mg/L	YAMAUCHI & MINAMIKAWA, 1990
<i>Castanea crenata</i>	CCA	Man/Glc	4,2 mg/L	NAKAMURA <i>et al.</i> , 2002
<i>Cratylia mollis</i>	CRAMOLL 1	Man/Glc	6 mg/L	VAREJÃO <i>et al.</i> , 2010
<i>Cycas revoluta</i>	CRLI	Man	18 mg/L	HARAGUCHI <i>et al.</i> , 2006
<i>Dolichus biflorus</i>	DBL	GalNAc	20 mg/L	CHAO <i>et al.</i> , 1994

<i>Erythrina crystagalli</i>	ECL	GalNAc/Gal	294 mg/L	STANCOMBE <i>et al.</i> , 2003
<i>E. corallodendron</i>	EcorL	GalNAc/Gal	4,7 mg/L	ARANGO <i>et al.</i> , 1992
<i>Glycine max</i>	SBA	GalNAc/Gal	100 µg/L	ADAR <i>et al.</i> , 1997
<i>Griffonia simplicifolia</i>	GS-II	Man/Glc	ND	ZHU <i>et al.</i> , 1996
<i>Maackia amurensis</i>	MAH	Neu5Ac2-3Gal1-3 (Neu5Ac2-6) GalNAc	ND	YIM <i>et al.</i> , 2001
<i>Phaseolus vulgaris</i>	PHA-L	Gal 1-4 GlcNAc 1-6(Gal1-4 GlcNAc1-2)	0,4-1 mg/L	HOFFMAN & DONALDSON, 1987
<i>P. lunatus</i>	LBL	GalNAc/Gal	5 mg/L	JORDAN & GOLDSTEIN, 1994
<i>Pisum sativum</i>	PSL	Man/Glc	2-5 mg/L 20 mg/L	STRUBBS <i>et al.</i> , 1986; PRASTHOFER <i>et al.</i> , 1989
<i>Robinia pseudoacacia</i>	RBL	GalNAc/Gal	10 mg/L	NISHIGUSHI <i>et al.</i> , 1997

Em geral, as lectinas recombinantes apresentam especificidades e afinidades muito similares às nativas, entretanto, podem diferir de forma mais significativa a nível estrutural. Isso ocorre principalmente porque originalmente algumas lectinas vegetais são submetidas a processamentos pós-traducionais que muitas vezes não são usuais. Por exemplo, as lectinas de *Pisum sativum* e *Vicia faba* são clivadas em duas cadeias, enquanto que a ConA (assim como outras lectinas de Diocleinae) sofre um processo de permutação que a torna circularmente homóloga a outras lectinas de leguminosas, como descrito anteriormente. Além disso, lectinas nativas glicosiladas, uma vez expressas em bactérias, apresentam-se desprovidas de seus oligossacarídeos (já que as mesmas apresentam-se incapazes de realizar a maioria dos processamentos pós-traducionais), e mesmo quando expressas em sistemas eucarióticos podem apresentar glicosilações diferenciadas das originais (ETZLER, 2000; PUSTZAI, 1991; RUDIGER, 1997).

6.1. Clonagem

O procedimento de clonagem de lectinas vegetais, assim como de outras proteínas, inicia-se com o isolamento do RNA total ou RNA poli-A do tecido em análise. No caso de lectinas de sementes, uma fonte interessante de material são as sementes em desenvolvimento, onde o nível de lectina aumenta rapidamente, assim como o do seu mRNA. A partir desses mRNAs pode ser construída então uma biblioteca de cDNA (DNA complementar) com o auxílio da enzima transcriptase reversa, e por meio da reação em cadeia da polimerase, utilizando iniciadores ou *primers* específicos, a biblioteca é ligada a vetores bacterianos ou plasmídeos. A detecção de clones de interesse é geralmente feita através da utilização de sondas de DNA específicas para hibridização. O passo crucial no procedimento de clonagem é a busca por clones positivos. Para o desenho de sondas de DNA é necessário ter conhecimento completo ou parcial sobre a sequência de aminoácidos da lectina de interesse; entretanto, como o código genético é redundante, a utilização de *primers* degenerados tem auxiliado nessa proposta. Caso não exista informação a respeito da sequência de aminoácidos, a utilização de sondas para lectinas relacionadas torna-se uma alternativa viável (STREICHER & SHARON, 2003; SAMBROOK, 1989).

6.2. Expressão e Purificação

O primeiro sistema utilizado para a expressão recombinante de proteínas foi o modelo procariótico *Escherichia coli*. Esse organismo foi inicialmente escolhido em decorrência do conhecimento aprofundado sobre o seu metabolismo, além de já haver sido utilizado anteriormente de forma bem sucedida para procedimentos de clonagem. Além disso, esse modelo oferece diversas rotas alternativas para a expressão de proteínas (MERGULHÃO *et al.*, 2005; CHOI & LEE, 2005), algumas das quais já foram utilizadas na produção recombinante de lectinas (Tabela 2).

Para a expressão de proteínas recombinantes, é necessário inicialmente subclonar o cDNA de interesse em um plasmídeo específico para a expressão, o qual estará sob o controle de um forte promotor induzível. Um dos sistemas de expressão mais populares é o pET, que engloba um família de plasmídeos detentores do promotor T7. Esse promotor normalmente não é reconhecido pela RNA polimerase da *E.coli*, sendo necessária para a expressão do gene sob o

seu controle a presença de uma RNA polimerase adequada. Uma vez subclonado, o plasmídeo é transferido para uma cepa apropriada de *E.coli* (geralmente BL21), desenvolvida para a expressão recombinante, a qual contém uma cópia cromossômica do gene para a T7 RNA polimerase sob o controle do promotor *lac*. Dessa forma, a produção da T7 RNA polimerase pode ser promovida pelo acréscimo do indutor isopropil- β -tiogalactosidase (IPTG), promovendo altos níveis de expressão da lectina recombinante (SØRENSEN & MORTENSEN, 2005a).

Um problema comum durante a expressão recombinante de proteínas é a formação dos chamados “corpos de inclusão”. Esses corpos consistem em agregados insolúveis que se formam no citoplasma bacteriano quando é detectada a presença de altos níveis de proteínas potencialmente tóxicas, exógenas ou inativas. Isso acaba por diminuir substancialmente a quantidade solúvel de proteína recombinante. Entretanto, esses agregados podem ser solubilizados (após a extração protéica) por meio de agentes desnaturantes (tampões contendo uréia 8 M ou cloridrato de guanidina 6 M), porém, é necessário submeter posteriormente a lectina a um processo de *refolding* (renaturação). Tal processo é realizado mediante a remoção progressiva do agente desnaturante por ultracentrifugação, diluição lenta ou sucessivas diálises. Frequentemente mais de 90% das lectinas formam corpos de inclusão, o que pode ser utilizado como uma forma prática de purificação, uma vez que a maior parte desses agregados é constituída por um único tipo de proteína (SØRENSEN & MORTENSEN 2005b). Isso ajuda a isolar e purificar agregados protéicos antes de solubilizá-los e renaturá-los.

O passo final no processo de purificação consiste geralmente em cromatografias de afinidade seguidas por diálise ou cromatografias de exclusão molecular. Uma das estratégias alternativas para facilitar esse processo consiste na expressão de proteínas de fusão. Para isso, o gene para lectina é ligado ao gene de outro peptídeo ou proteína (já presente no plasmídeo de expressão) que apresente propriedades de interesse. Dentre esses parceiros, podemos citar a tioredoxina, capaz de conferir uma maior solubilidade a proteína que se deseja expressar; ou a cauda de histidina (seis resíduos de histidina fusionados a um dos terminais da lectina). Essa cauda de histidina é capaz de formar interações reversíveis com metais de transição imobilizados (Co^{2+} ou Ni^{2+}), possibilitando a purificação da proteína de interesse em um único passo cromatográfico (SØRENSEN & MORTENSEN, 2005a).

7. Estudos Estruturais e Termodinâmicos de Interações Lectina-Carboidrato

Tornar-se pequeno o suficiente, até a escala atômica, e ver com seus próprios olhos uma interação molecular em tempo real é o sonho de todo bioquímico, embora infelizmente não seja ainda possível. Uma combinação de técnicas experimentais, no entanto, pode chegar perto. Destas, a cristalografia de proteínas por difração em raios X é atualmente a alternativa mais eficiente na investigação estrutural de moléculas biológicas, pois resulta em uma imagem tridimensional da macromolécula. Se o conhecimento estrutural permite entender as topologias de superfície, as forças de interação e os impedimentos estéricos, por outro lado traz um aspecto "estático" de uma interação, que nem sempre retrata fielmente o mecanismo da reação. Ela oferece apenas uma visão relativamente fixa do fenômeno, já que se trata de uma "fotografia instantânea" da interação, e que a proteína na forma cristalina pode estar longe das suas condições biológicas naturais. Assim a combinação com outras técnicas, tais como microcalorimetria, que permite o acesso a grandezas termodinâmicas em curso na interação, fornecem informações complementares necessárias para entender o mecanismo de reação. Outras técnicas, como testes de concorrência entre ligantes, nos falam também sobre a especificidade da proteína. É, portanto, a complementaridade de métodos que permitem elucidar o papel e a função de uma proteína e visualizar, por exemplo, o desenho de moléculas terapêuticas.

Um completo entendimento a nível molecular do mecanismo de associação proteína-carboidrato requer a elucidação tanto de aspectos estruturais como energéticos do processo. Uma grande quantidade de dados estruturais está disponível, mas a ligação entre propriedades estruturais e energéticas permanece ainda obscura para permitir uma descrição energética da interação a partir da informação estrutural (TOONE *et al.*, 1994). A acessibilidade de microcalorímetros de alta resolução têm provido um método direto de avaliação de dados de acurácia termodinâmica que podem complementar a informação estrutural (AMBROSI *et al.*, 2005).

Lectinas em geral são fáceis de purificar por cromatografia de afinidade e um largo número dessas proteínas têm sido caracterizadas por cristalografia. Base de dados de lectinas em 3-D, disponíveis na internet em <http://www.cermav.cnrs.fr/lectines/>, contém mais de 500 estruturas tridimensionais, muitas dessas em complexo com carboidratos ligantes. Lectinas apresentam uma larga variedade de estruturas, variando em *fold* do monômero, estado de

oligomerização ou arquitetura do sítio de ligação. Lectinas de plantas e animais têm sido bem caracterizadas nas últimas décadas (LIS & SHARON, 1998; LORIS, 2002), tanto como lectinas de bactérias e de fungos (IMBERTY *et al.*, 2005).

Interações lectina-carboidrato são geralmente caracterizadas por uma baixa afinidade a seus ligantes monovalentes, um inconveniente balanceado pela multivalência que provém de sua alta afinidade por substratos com diversidade potencial de ligantes, como em glicanos complexos ou superfícies celulares. Em geral, uma afinidade milimolar é observada para lectinas que se ligam a monossacarídeos; quando oligossacarídeos mais longos agem como ligantes, correspondendo a um envolvimento de um sítio de ligação estendido sobre a superfície da lectina, a afinidade aumenta para um nível micromolar.

Nesse contexto, microcalorimetria por titulação têm sido usada para caracterização das associações termodinâmicas de proteínas-carboidratos (AMBROSI *et al.*, 2005; DAM & BREWER, 2002). Em geral, essas associações são tipificadas por um termo entálpico favorável, devido ao alto número de ligações de hidrogênio e contatos de Van der Waals entre faces hidrofóbicas do açúcar e aminoácidos aromáticos, que é contrabalanceada por uma contribuição entrópica desfavorável, que tem sido atribuída tanto ao rearranjo de moléculas de solvente, como pela perda de flexibilidade conformacional / mobilidade do ligante (DAM & BREWER, 2002).

As medições de titulação calorimétrica isotérmica (ITC) tornaram-se um método importante e amplamente utilizada para investigar a termodinâmica de interações carboidrato-lectina. Assim, nenhum outro método oferece a facilidade de determinar diretamente parâmetros termodinâmicos de ligação sem modificar os carboidratos ou proteínas de interesse. Como exemplo, a interação entre a lectina de leguminosa de *Erythrina corallodendron* e lactose tem sido caracterizado tanto por microcalorimetria (SUROLIA *et al.*, 1996) como por cristalografia de raios X (SHAANAN *et al.*, 1991). A rede de ligação de hidrogênio típica para o sítio de ligação a carboidrato de uma lectina de leguminosa é observada em conjunto com uma interação hidrofóbica com uma fenilalanina. Neste caso, a baixa afinidade ($K_a = 9,7 \cdot 10^3 \text{ M}^{-1}$, $K_d = 1 \cdot 10^{-4} \text{ M}$) corresponde ao termo entálpico favorável ($\Delta H = -47,1 \text{ kJ / mol}$) compensando um termo entrópico desfavorável ($T\Delta S = -24,4 \text{ kJ / mol}$).

Em um outro trabalho, sete lectinas de leguminosas foram utilizadas para um estudo termodinâmico detalhado. As lectinas foram isoladas das espécies *Canavalia brasiliensis*, *C. bonariensis*, *Cratylia floribunda*, *Dioclea rostrata*, *D. virgata*, *D. violacea* e *D. guianensis*.

Apesar de suas proximidades filogenéticas e sequências conservadas, as lectinas acima exibem diferentes atividades biológicas. Estudos de hemaglutinação e termodinâmicos revelou semelhanças e diferenças importantes nas propriedades de carboidratos de ligação desses lectinas. Especificidades diferenciais a carboidratos foram encontradas sendo correlacionadas com a atividade biológica dessas lectinas. (DAN & BREWER, 2002)

Gomes *et al.* (1994) investigaram a correlação significativa entre as propriedades de liberação de histamina dessas lectinas e suas constantes de afinidade para um carboidrato complexo biantenário, mostrando que os valores de K_a dessas lectinas para o mesmo e a quantidade de histamina liberada pelas lectinas a 10 $\mu\text{g/mL}$ estão correlacionados. A forte indução de histamina pelas lectinas ConA, *C. brasiliensis*, *D. guianensis* e *D. virgata* exibem afinidades relativamente alta (K_a) para esse composto. Por outro lado, as lectinas restantes relativamente inativas possuem afinidades mais baixas para esse carboidrato complexo. Parece, portanto, que a indução de histamina liberada por mastócitos peritoniais de rato por ConA, *C. brasiliensis*, *D. guianensis* e *D. virgata* envolve ligação das lectinas a um carboidrato complexo biantenário como epítopo presente em superfície da célula.

Uma afinidade muito mais elevada foi observada para as lectinas solúveis de bactérias que interagem com os açúcares presentes em tecidos do hospedeiro. Um estudo recente em lectinas bacterianas demonstrou seu papel no reconhecimento do hospedeiro, formação do biofilme, adesão ao tecido e virulência. *Pseudomonas aeruginosa* e *Burkholderia cenocepacia* são patógenos oportunistas responsáveis por infecções nos pulmões que são grandes ameaças em pacientes com fibrose cística ou indivíduos imuno-comprometidos. Ambas as bactérias possuem diversas lectinas cálcio-dependentes que exibem alta afinidade a diversos oligossacarídeos que estão presentes em tecidos humanos. Usou-se uma combinação de abordagens como microcalorimetria de titulação, cristalografia por raios X e modelagem molecular para decifrar as bases termodinâmicas e estruturais para a alta afinidade de ligação de lectinas bacterianas a carboidratos do hospedeiro (CIOCI *et al.*, 2003). As observações estruturais e termodinâmicas servem como base para o desenho e síntese de glicomiméticos com afinidade sub-micromolar para o sítio lectínico (MAROTTE *et al.*, 2007a,b).

Ultimamente reforça-se que o completo entendimento da interação lectina-carboidrato é essencial para as lectinas que sejam exploradas em aplicações biomédicas (JANAKIRAMAN *et al.*, 1994, MAMMEN *et al.*, 1998, MOLEMA & MEIJER, 1994). O profundo conhecimento de

ambos os aspectos estrutural e energético do processo é necessário. Certamente, a falha de um número de fármacos baseadas em carboidrato pode ser atribuída ao pobre entendimento de seus supostos mecanismos de ação. (JANAKIRAMAN *et al.*, 1994). Quando há um bom entendimento, os resultados se mostram impressionantes: a toxina de Shiga, por exemplo, consiste de uma subunidade A enzimaticamente ativa que consegue entrar em células susceptíveis após reconhecimento de oligossacarídeos por um homopentâmero a uma baixa afinidade, contornada pela ligação simultânea a vários receptores oligossacarídeos da superfície celular. Através de estudos cristalográficos e de afinidade, foi comprovada a eficiência de um carboidrato oligovalente análogo ao receptor, que apresentou uma atividade inibitória subnanomolar cerca de 10 milhões de vezes superior a qualquer ligante univalente reportado, capaz de envolver o homopentâmero de até duas toxinas (KITOV *et al.*, 2000). Em outro estudo, Fan *et al.* (2000) desenvolveram um ligante pentavalente para a enterotoxina de *E. coli*, baseado em galactose. O ligante foi desenhado combinando o arranjo especial do receptor (tendo uma simetria em 5) com um suporte sintético rígido no qual galactoses ligantes são adicionadas via espaçadores flexíveis de comprimento variável. Eles demonstraram que o ligante com espaçadores combinados na geometria do sítio de ligação do melhor alvo mostrou alta afinidade, com um ganho de 1000 x quando comparado a ligantes monovalentes.

8. Objetivos do Trabalho de Tese

8.1. Objetivo geral

No presente estudo, objetivou-se demonstrar semelhanças e diferenças da lectina recombinante de *Platypodium elegans* em relação a outros membros desta família de lectinas. Descrevemos a clonagem, a produção recombinante, a especificidade, a afinidade frente a oligossacarídeos complexos e a estrutura de PELa (lectina recombinante A de *Platypodium elegans*), o produto de um gene presente em sementes desta espécie.

8.2. Objetivos específicos

- Iniciar estudos envolvendo a clonagem e caracterização molecular do gene da lectina de sementes de *Platypodium elegans*;
- Desenvolver técnicas de expressão e purificação da lectina em *E. coli*, com a obtenção final da proteína recombinante (PELa) pura;
- Avaliar sua afinidade e especificidade para diferentes oligossacarídeos por microcalorimetria de titulação isotérmica (ITC) e em experimentos de *Glycan Array*.
- Cristalizar e determinar a sua estrutura tridimensional, obtendo a primeira estrutura cristalina de uma lectina recombinante da subtribo Dalbergiae complexada com oligossacarídeos. Isto irá fornecer novas estruturas de uma lectina da família das leguminosas, cuja subtribo Dalbergiae só apresenta atualmente um representante estrutural descrito.
- Por fim, determinar as semelhanças e diferenças na especificidade, estrutura e afinidade entre esta lectina e as demais da mesma família.

O estudo dessa lectina presente em sementes de *P. elegans* e sua caracterização bioquímica irá fornecer informações estruturais importantes sobre as interações proteína-açúcar necessária para compreender os mecanismos de ação, o reconhecimento, especificidade e afinidade a oligossacarídeos complexos em termos moleculares. A longo prazo, esta lectina

produzida na forma recombinante poderá ser testada na sua capacidade de reconhecer glicanos envolvidos em eventos biológicos interessantes, desde a investigação de sua função no organismo de origem, como em aplicações biotecnológicas em áreas como farmacologia, imunologia, dentre outras.

Capítulo 2:

Materiais e Métodos

Esse capítulo foi dividido em três partes experimentais principais: a primeira descreve métodos relacionados à clonagem e expressão heteróloga de uma lectina de sementes de *Platypodium elegans*, assim como sua purificação a partir de corpos de inclusão; a segunda traz experimentos que descrevem a funcionalidade, a especificidade e a afinidade dessa lectina recombinante frente a oligossacarídeos diversos utilizando-se técnicas como atividade hemaglutinante, *glycan array* e microcalorimetria; e a terceira reúne a metodologia utilizada para a determinação da estrutura cristalográfica da lectina em complexo com oligossacarídeos de interesse.

I. Clonagem e Expressão Heteróloga de uma lectina de sementes de ***Platypodium elegans***

Sem o conhecimento de início da sequência primária da lectina de *Platypodium elegans* (PELa), a estratégia adotada para a obtenção e clonagem do gene correspondente à partir de sementes foi feita da obtenção do mRNA extraído de sementes, com a produção subsequente do cDNA pelas técnicas 3'e 5' RACE, como descrito a seguir. Em seguida, o gene foi clonado no vetor de expressão pET-29a, visando a produção heteróloga da proteína em células de *E. coli*.

1. Obtenção das Sementes de *Platypodium elegans*

Sementes de *Platypodium elegans* foram adquiridas comercialmente da Reserva de Reflorestamento Flora Tietê, São Paulo. O material foi enviado ainda nas vagens, embora já

maduras e parcialmente partidas, chegando-se apenas as porções que continham a semente. Para experimentos de extração de RNA total, as sementes foram estocadas em freezer à -70 °C.

2. Isolamento do RNA Total de Sementes de *P. elegans*

Para o isolamento do RNA total a partir de sementes de *P. elegans* foi utilizado o kit *Plant RNA Reagent* (Invitrogen). O procedimento consistiu em adicionar 5 mL de *Plant RNA Reagent* a cada 1 g de sementes congeladas e maceradas na presença de nitrogênio líquido (-173 °C). Em seguida, a mistura foi homogeneizada com uso de vortex, incubada por 5 minutos à temperatura ambiente e centrifugada por 2 minutos a 12.000 x g em microcentrífuga. O sobrenadante foi transferido para um tubo de centrífuga livre de RNase e a ele foi adicionado 1 mL de NaCl 5 M. Em seguida, a amostra foi misturada gentilmente por inversão, e a ela foram adicionados 3 mL de clorofórmio. Logo após, a amostra foi homogeneizada e centrifugada a 4 °C por 10 minutos a 12.000 x g para que houvesse a separação das fases. A fase aquosa (fase superior) foi transferida para um tubo de centrífuga livre de RNase e a ela foi adicionado igual volume de álcool isopropílico. A amostra foi deixada em repouso por 10 minutos a temperatura ambiente e, em seguida, centrifugada a 4 °C por 10 minutos a 12.000 x g. O sobrenadante foi descartado e ao *pellet* foi adicionado 10 mL de etanol 75%. Logo após, a amostra foi centrifugada à temperatura ambiente por 1 minuto a 12.000 x g e o sobrenadante foi cuidadosamente descartado. O *pellet* foi dissolvido em 100 µL de água estéril contendo DEPC (dietilpirocarbonato). Em seguida, a quantificação, determinação da pureza e integridade do RNA total extraído foram verificadas através de eletroforese em gel de agarose a 1% (pré-corado com GelRed 10000 x diluído), e espectrofotometria através da leitura dos comprimentos de onda a 260 nm (1 unidade de absorbância a 260 nm corresponde a cerca de 40 µg/mL de RNA) e 280 nm. As leituras a 260 e 280 nm, respectivamente, serviram como parâmetro na estimativa do grau de contaminação do ácido nucléico com proteínas. Os valores considerados satisfatórios foram aqueles entre 1,7 a 2,0. O RNA total extraído foi armazenado a -70 °C.

3. Obtenção, Clonagem e Caracterização do gene da lectina de *P. elegans*

Sem o conhecimento da sequência primária da lectina de *P. elegans*, a estratégia adotada para obtenção e clonagem do gene da lectina de *P. elegans* partindo-se de sementes foi feita a partir da obtenção do mRNA extraído de sementes, com a conseguinte produção de cDNA por técnicas como 3'RACE e 5'RACE, como descrito a seguir.

A amplificação rápida de extremidades de cDNA (RACE) é um procedimento para o aumento de sequências de ácido nucleico a partir de um molde de RNA mensageiro entre um sítio interno definido e a extremidade 3' ou 5' final do mRNA (FROHMAN *et al.*, 1988). Esta metodologia de amplificação com especificidade para uma única fita tem sido descrita como PCR "unilateral" (OHARA *et al.*, 1989). Uma PCR (reação em cadeia da polimerase) requer dois *primers* sequência-específicos que flanqueiem a sequência a ser amplificada (SAIKI *et al.*, 1985; 1988). No entanto, para amplificar e caracterizar regiões de sequências desconhecidas, essa exigência impõe uma limitação.

3' RACE aproveita a cauda poli(A) natural encontrado em mRNA como um sítio iniciador genérico para PCR. Neste procedimento, mRNAs são convertidos em cDNAs utilizando a transcriptase reversa (RT) e um *primer* adaptador oligo-dT. Um cDNA específico é então amplificado por PCR usando um *primer* específico ao gene que anela a uma região de sequência éxon conhecida, e um *primer* adaptador, que tem como alvo a região da cauda poli (A). Isto permite a captura de sequências desconhecidas de 3'-mRNA que se situam entre o éxon e a cauda poli (A).

5' RACE usa um *primer* antisenso específico ao gene para a síntese de cDNA específico pela transcriptase reversa. Antes da PCR, um passo de adição de uma cauda homopolimérica pela enzima TdT acrescenta uma sequência adaptadora para a extremidade 5' desconhecida do cDNA. O cDNA específico é então amplificado por PCR usando um *primer* específico ao gene que anele em uma região de sequência éxon conhecida e um *primer* adaptador que tem como alvo o término da extremidade 5', com a cauda homopolimérica.

A técnica RACE tem sido utilizada para a amplificação e clonagem de mRNAs raros e pode ser aplicada a bibliotecas de cDNA existentes (BERCHTOLD, 1989). Além disso, os produtos RACE podem ser diretamente sequenciados sem passos intermediários de clonagem (BROW, 1990; ADAMS & BLAKERLEY, 1991), ou serem usados para preparar sondas

(FROHMAN & MARTIN, 1989). Produtos gerados por procedimentos 3' e 5' RACE podem ser combinados para gerar cDNAs de cadeia completa (FROHMAN & MARTIN, 1989; FROHMAN, 1990). Por último, os procedimentos RACE podem ser utilizados em conjunto com o métodos de captura de éxons (BUCKLER *et al.*, 1991) para permitir a amplificação e posterior caracterização de sequências de codificação desconhecida.

3.1. 3'RACE (System for Rapid Amplification of cDNA Ends)

Em modo geral, foi realizada uma transcrição reversa utilizando-se oligo (dT) contendo em sua extremidade 5' um sítio adicional (5'-GGCCACGCGTCGACTAGTACTTTTTTTTTTT TTTTTTT-3'), cuja finalidade foi de servir de sítio para o anelamento posterior de iniciadores adaptadores *Ancor primer* (AP). Dessa maneira, após a transcrição reversa da população de mRNA, o transcrito específico do gene foi obtido através de PCR utilizando um iniciador degenerado para a região interna do gene (descrito a seguir) e o *Ancor primer* (AP), de acordo com as recomendações do fabricante do 3' RACE (Invitrogen). Esse procedimento foi realizado em 3 etapas: síntese do cDNA, desenho de *primer* para região interna do gene e amplificação do gene. O produto amplificado foi clonado, sequenciado e caracterizado.

3.1.1. Síntese da primeira fita de cDNA

A reação de síntese da primeira fita de cDNA à partir da população de mRNA foi catalisada pela transcriptase reversa SuperScriptTM II RT (Invitrogen), uma DNA polimerase dependente de RNA que utiliza como fita molde RNA, DNA ou RNA híbrido com DNA para sintetizar uma fita de DNA complementar (cDNA). Essa enzima é uma mutante da *Moloney Murine Leukemia Virus* Transcriptase Reversa (M-MLV RT), que foi engenhada para reduzir a atividade Rnase H, resultando em melhores rendimentos e em uma síntese mais completa (GERARD *et al.*, 1989; D'ALESSIO *et al.*, 1990; GERARD *et al.*, 1992), podendo ser usada para a síntese de cDNA à partir de uma preparação de RNA total. A enzima possui atividade exonucleásica 3'-5', e para a construção da primeira fita de cDNA foi utilizada as seguintes condições, seguindo recomendações do fabricante:

REAGENTE	QUANTIDADE
RNA Total	5 µg
Oligo dT 10 µM (<i>Ancor Primer</i>)	1 µL
10x PCR Buffer	2 µL
MgCl ₂ 25 mM	2 µL
dNTP mix 10 mM	1 µL
DTT 0,1 M	2 µL
SS II RT (200 u/µL)	1 µL
Água nuclease-free	q.s.p. 20µL

O RNA total foi desnaturado a 65 °C por 10 minutos e após isso se adicionou o *primer* Oligo dT (*Ancor Primer*), incubando-os a 70 °C por 10 minutos. Logo em seguida, todos os reagentes foram adicionados e a reação completa foi incubada a 42 °C por 50 minutos. Após isso, a enzima foi então desnaturada a 70 °C por 15 minutos.

Após a síntese do cDNA, o molde de RNA foi removido da molécula híbrida cDNA:RNA por digestão após adição de 2 u de RNase H a 37 °C por 20 minutos.

3.1.2. Desenho de iniciadores degenerados (*primers*) para o gene da lectina de *P. elegans*

Dois *primers* degenerados para o gene da lectina presente em sementes de *P. elegans* foram desenhados visando o mínimo de degeneração (PLAT1 e PLAT2), a partir da sequência primária da lectina de *Pterocarpus angolensis* (código de acesso: CAD198110). Foram utilizadas duas sequências conservadas distintas correspondentes aos sítios de ligação a metais de lectinas de leguminosas (Tabela 4). Para verificar a estrutura de cada *primer*, porcentagem de conteúdo C/G, bem como temperatura de *melting* (T_m) foi utilizado o simulador gratuito de construção de *primers* disponível na página <http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx>.

Tabela 4. Desenho de *primers* degenerados para o gene de PELa.

Sequência de Referência	Sequência de Aminoácidos	Posição na Sequência	Primer	Sequência do Primer*	Temperatura de <i>melting</i> (Tm)
<i>Pterocarpus angolensis</i> lectin (CAD19811)	(V/L/I)AVEFDT	145	PLAT1	5'-TNGCNGTNGART TYGAYAC-3'	54,5
<i>Pterocarpus angolensis</i> lectin (CAD19811)	HIGIDVN	166	PLAT2	5'-CAYATHGGNATHGA YGTNAA-3'	52,8

* R, Y e N indicam (A ou G), (C ou T), and (A, C, G ou T), respectivamente.

3.1.3. Amplificação do produto 3' RACE da lectina de *P. elegans*

Para a amplificação do cDNA foi utilizada a estratégia de reação de polimerização em cadeia (PCR). Como fita molde foi utilizada a primeira fita de cDNA e, como iniciadores, um *primer* degenerado para o gene (PLAT1 ou PLAT2, como senso) e oligo-dT (AP, anti-senso). Para a reação de PCR foi utilizado o kit GoTaq® Green Master Mix: tampão de reação, dNTP mix 10 mM (cada), iniciadores para o gene (senso e anti-senso), DNA molde (cDNA), enzima GoTaq polimerase 5U/μL, água estéril, seguindo instruções do fabricante:

REAGENTE	QUANTIDADE (μL)
cDNA (<0,5μg)	5
Iniciador senso PLAT1 ou PLAT2 (1μg/μl)	0,5
Iniciador anti-senso AP (1μg/μl)	0,5
5x GoTaq Buffer	10
PCR Nucleotide Mix (10μm cada)	1
GoTaq (5u/μl)	0,250
Água estéril	q.s.p. 50

A reação de PCR conteve a seguinte condição de reação:

PASSO	TEMPERATURA (°C)	TEMPO	N° CICLOS
Desnaturação Inicial	95	3'	1
Desnaturação	95	45''	35
Anelamento	50*	1'	35
Extensão	72	3'	35
Extensão Final	72	5'	1
Fim	16	-	1

*Obs.: A temperatura de anelamento varia conforme a temperatura de *melting* (T_m) dos oligonucleotídeos. A temperatura de anelamento deve ser em torno de 5°C menor que a do T_m mais baixo.

O produto final foi analisado em gel de agarose a 1,2% e foi reamplificado, nas mesmas condições.

3.1.4. Purificação do produto 3' RACE da lectina de *P. elegans*

O produto reamplificado foi purificado utilizando-se QIAquick Gel Extraction Kit (Qiagen), através de extração em gel de agarose 1,2%, seguindo-se as instruções do fabricante. As porções de gel contendo DNA correspondente ao tamanho molecular esperado do cDNA do gene foram cortados e pesados em um único tubo de microcentrífuga. Foram adicionados 3 volumes de gel de *QG buffer*, sendo volume de gel o correspondente ao peso obtido em miligramas das porções de gel cortadas. A mistura foi incubada a 50 °C por 10 minutos, a fim de fundir o gel de agarose, e depois se adicionou 1 volume de gel de isopropanol. A coluna do “kit” foi colocada em um tubo coletor e a agarose dissolvida foi aplicada na coluna. O sistema foi brevemente centrifugado por 1 minuto (12000 x g) e o tubo coletor esvaziado. Foram adicionados à coluna 500 µL do *QG buffer* e o sistema foi centrifugado novamente. Adicionou-se à coluna 750 µL do *PE buffer* e o sistema foi deixado incubando por 5 min, sendo, após isso, centrifugado novamente. O tubo coletor foi trocado por um tubo de microcentrífugação novo e 50 µL de TE (10mM Tris pH 8,0; 1mM EDTA pH 8,0) foram aplicados na coluna a fim de eluir o fragmento purificado. O sistema foi incubado por 5 minutos à temperatura ambiente e centrifugado a 12000 x g por 1 minuto. Após a purificação, o gene foi analisado qualitativamente

através da relação das absorvâncias a 260 e 280 nm e quantitativamente através da absorvância a 260nm (1 unidade de absorvância a 260 nm correspondente a cerca de 50 µg/mL de DNA).

3.1.5. Clonagem do produto 3' RACE da lectina de *P. elegans* no vetor de clonagem pGEM-t easy

O sistema de vetores pGEM®-T e pGEM®-T easy (Promega) é um sistema conveniente para a clonagem de produtos de PCR. Os vetores são preparadas cortando-os com EcoRV e adicionando-se uma timidina 3' terminal em ambas as extremidades. Esta única 3'-T terminal no local de inserção melhora significativamente a eficiência de ligação de um produto PCR para os plasmídeos por duas razões: impedindo a recirculização do vetor; fornecendo uma extremidade compatível com produtos de PCR gerados por alguns polimerases termoestáveis (MEZEI & STORTS, 1994; ROBLES & DOERS, 1994), já que estas frequentemente adicionam uma desoxiadenosina (A) terminal, na extremidade 3' de fragmentos amplificados (CLARK, 1988; NEWTON & GRAHAM, 1994) (Figura 10). Por esta razão, o produto 3' RACE amplificado e purificado foi sequencialmente clonado no plasmídeo pGEM®-T easy.

O processo de clonagem do gene foi realizado com auxílio da enzima T4 DNA ligase, tampão de ligação e vetor pGEM®-T easy, numa proporção 3:1 inserto/vetor. O procedimento foi executado de acordo com recomendações do fabricante. Após clonagem, o plasmídeo resultante da ligação foi inserido em células de *E. coli* XL1 blue competentes, segundo protocolo de transformação.

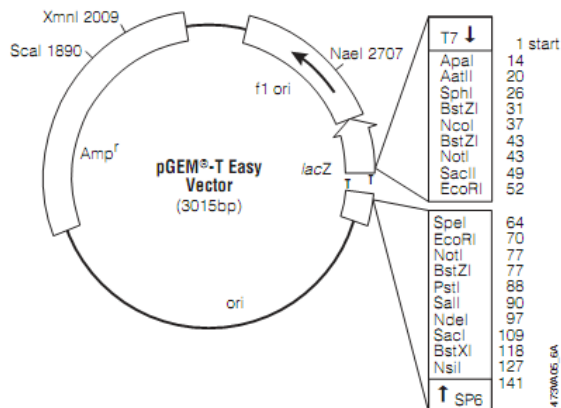


Figura 10. Mapa do vetor pGEM-T easy (Promega).

3.1.6. Preparação de células competentes

Uma bactéria competente é uma bactéria que está apta a receber DNA exógeno. A competência pode ocorrer naturalmente ou pode ser produzida através de diferentes técnicas, que envolvem tratamento com soluções, como cloreto de cálcio, e mudanças bruscas de temperatura. Os íons de metal e as mudanças bruscas de temperatura alteram a permeabilidade da parede celular, fazendo com que estas células permitam a entrada de DNA exógeno através da membrana plasmática.

As células de *E. coli* não são naturalmente competentes. A penetração do DNA pode ser possível, por exemplo, graças ao *stripping* de células com Ca^{2+} e Mg^{2+} . Estes cátions divalentes reduzem a repulsão eletrostática entre as moléculas de DNA e a superfície celular da bactéria, deixando-a competente à introdução de DNA. Durante o choque de calor, abrem-se os poros na membrana do hospedeiro, permitindo a passagem de DNA. Quando as células são rapidamente resfriadas, fecham-se os poros, prendendo o DNA introduzido nas células.

A preparação de células competentes para este trabalho foi realizada de acordo com o método modificado de Hanahan, conforme descrito por Sambrook (2002), ou método de cloreto de cálcio, conforme descrito abaixo:

1. Um pré-inóculo de células de *E. coli* (XL1 Blue) foi preparado em 5 mL de meio líquido LB (Luria-Bertani) Broth e incubado a 37 °C durante 16 horas, sob agitação constante.

2. A cultura foi diluída (1/100) em 500 mL de LB líquido e incubada novamente a 37°C até atingir $A_{600\text{nm}} = 0,4-0,6$. Após isso, foi incubada em banho de gelo por 5 minutos

3. As células foram centrifugadas a 1.600 x g a 4 °C por 10 minutos e o *pellet* suspenso cuidadosamente em 50 mL de MgCl_2 100 mM gelado.

4. A ressuspensão foi incubada em banho de gelo por 30-45 min e então centrifugada a 1.600 x g a 4°C por 10 minutos, descartando o sobrenadante.

5. O *pellet* foi cuidadosamente suspenso em 10 mL de tampão CaCl_2 100 mM gelado e incubado em banho de gelo por 90-120 min.

6. Dois mL de glicerol foram adicionados (15-20% final) e distribuiu-se rapidamente a suspensão de células em tubos de microcentrífuga (em banho de gelo), em alíquotas de 100 µL. Os tubos foram rapidamente congelados em nitrogênio líquido e estocados em freezer -80 °C.

3.1.7. Transformação em células de *E. coli* XL1 blue competentes com o vetor de clonagem

A transformação é a introdução de moléculas de DNA em uma célula hospedeira (bactérias ou células eucarióticas, incluindo células vegetais e animais). Este procedimento permite a entrada do produto 3'RACE clonado em vetor plasmidial para o interior de células de *E. coli*.

Para isso, 10 µL do produto clonado no vetor pGEM-T easy foi adicionado a 100µL de células competentes de *E. coli* XL1 blue. Em seguida, as células foram incubadas em gelo por 30 minutos, sucedidas de um choque térmico em banho-maria a 42 °C por 45 segundos, e mais dois minutos no banho de gelo. Finalmente as células foram incubadas a 37 °C por 1 hora em meio LB Broth (SAMBROOK, 2002).

3.1.8. Validação da transformação

A clonagem bem sucedida de uma inserção dentro do pGEM ®-T interrompe a sequência de codificação de β-galactosidase; clones recombinantes, portanto, podem ser identificados pela triagem de cores em placas indicadoras contendo ampicilina, IPTG e X-GAL. O vetor confere naturalmente resistência a ampicilina. A ausência de coloração azul é um indicativo de que algum inserto do DNA se inseriu no sítio de clonagem (plasmídeo), fato que impede a expressão do gene da galactosidase induzida pelo IPTG, com consequente inibição do catabolismo do X-Gal (a quebra do X-Gal gera um composto de coloração azul).

Para validação da transformação, portanto, células de *E. coli* após transformação foram plaqueadas em meio sólido LB ágar contendo 50 µg/mL de ampicilina, 80 µg/mL X-GAL e 0,5 mM IPTG. As placas foram incubadas a 37°C por 16 horas. Após o crescimento, as colônias brancas foram selecionadas individualmente para a purificação dos plasmídios.

3.1.9. Purificação e digestão enzimática dos plasmídios recombinantes

Purificações de DNA plasmidial em pequena escala, mais conhecido como *minipreps*, são comumente utilizados em procedimentos de biologia molecular. Ao longo dos anos, muitos protocolos *miniprep* têm sido usados (AUSUBEL, 1989), mas poucos têm se mostrado consistente e confiável. O procedimento *miniprep* pode ser tanto trabalhoso e demorado, especialmente quando um grande número são realizados em paralelo. O Wizard® Plus Minipreps DNA Purification System (Promega), o qual elimina muitos dos problemas associados com os procedimentos frequentes em *miniprep*, fornece um método simples e confiável para rapidamente isolar DNA plasmidial.

Para esta etapa foi utilizado o kit Wizard® Plus SV Minipreps DNA purification System (Quick protocol, Promega). O procedimento foi executado de acordo com o manual do fabricante.

Cada colônia positiva testada foi inoculada em 10 mL de meio líquido LB Broth (Luria-Bertani) a 37°C, 16 h, 180 rpm. As células foram então precipitadas por centrifugação a 1000 x g por 5 min em tubos de microcentrifugação e o sobrenadante foi descartado. O *pellet* final foi ressuspenso em 250 µL de *Cell Resuspension Solution*, homogeneizado, e então foi adicionado 250 µL de *Cell Lysis Solution*. Ao homogeneizado, acrescentou-se 10 µL de *Alkaline Solution*, deixando em incubação a temperatura ambiente. Ao fim de 5 min, 350 µL de *Neutralization Solution* foram adicionados, homogeneizados e submetidos a uma nova centrifugação a 13000 x g por 10 min. O lisado clarificado, correspondente ao sobrenadante, foi então aplicado em uma coluna de sílica disponível no “kit”, centrifugando-a em um tubo coletor a 13000 x g por 1 min. Após isso, 750 µL e depois 250 µL de *Wash Solution* foram aplicados na coluna, intercalados por centrifugação a 13000 x g por 1 min. Em um tubo de microcentrifugação limpo, o DNA plasmidial purificado foi então eluído da coluna com a adição de 50 µL de água nuclease-free, e analisado qualitativa e quantitativamente através da relação das absorbâncias a 260 e 280 nm.

Depois de purificados, uma amostra dos plasmídios sofreu digestão enzimática pela enzima NotI (10 u/µg DNA plasmidial) a 37 °C por 90 min. A digestão foi verificada em gel de agarose a 0,8%. Foram selecionados os plasmídios que, após digestão, liberaram um inserto de tamanho referente ao produto isolado, e foram então enviados para sequenciamento.

3.1.10. Análise da sequência do produto 3' RACE

As sequências obtidas foram analisadas quanto à qualidade de bases utilizando-se o programa Phred-Phrap-Consed (EWING *et al.*, 1998) e escore superior a 20. O software Consed foi a interface de visualização das sequências analisadas pelo Phred-Phrap. A sequência final obtida foi comparada com as sequências disponíveis nos bancos públicos não redundantes (nr) do GenBank (www.ncbi.nlm.nih.gov) por meio do algoritmo Blastx (Basic Local Alignment Search Tool).

3.2. 5' RACE (System for Rapid Amplification of cDNA Ends)

A estratégia geral da técnica 5'RACE consiste na síntese de cDNA à partir de uma população de RNA total, usando um *primer* antisense gene específico (GSP). Isto permite a conversão de mRNA específico em cDNA, ou famílias relacionadas de mRNAs, e maximiza o potencial para extensão completa da sequência 5' terminal. Após a síntese de cDNA, o primeira fita é purificada de dNTPs não incorporados, enzimas e *primers*. Ao fim, a TdT (*Terminal deoxynucleotidyl transferase*) é usada para adicionar uma cauda homopolimérica poli d(C) na extremidade 3' do cDNA purificado, e uma combinação de um *primer* adaptador complementar a esse homopolímero e do *primer* gene-específico correspondente permitem a amplificação de sequências desconhecidas entre na extremidade 5' do mRNA.

Esse procedimento foi realizado em 4 etapas principais: desenho de *primer* para região 3' do gene, síntese e purificação do cDNA, adição da cauda poli d(C) e amplificação do gene. O produto amplificado foi clonado, sequenciado e caracterizado, como descrito a seguir.

3.2.1. Desenho de iniciadores específicos (*primers*) para a extremidade 3' do gene da lectina de *P. elegans*

Um iniciador dito específico para o experimento 5' RACE foi desenhado à partir da sequência da extremidade 3' do cDNA (obtido pela estratégia 3' RACE). Esse iniciador (PLAT3) foi escolhido com o objetivo de obter o gene completo até sua extremidade 5' (Tabela 5). Para verificar a estrutura do iniciador, porcentagem de conteúdo C/G, bem como temperatura

de *melting* (T_m) foi utilizado o simulador gratuito de construção de *primers* disponível na página <http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/Default.aspx> .

Tabela 5. Desenho de *primer* específico para o gene a partir da sequência nucleotídica do produto 3' RACE.

Sequência de Referência	Sequência de Nucleotídios	Primer	Sequência do Primer	Temperatura de <i>melting</i> (T_m)
Extremidade 3'RACE	5'-TCCAGAGGGA TCTCTATATCA-3'	PLAT3	5'-TGATATAGAGATC CCTCTGGA-3'	55,9

3.2.2. Síntese da primeira fita de cDNA

A reação de síntese da primeira fita de cDNA à partir da população de mRNA foi catalisada pela SuperScript™ II RT (Invitrogen), e para a construção da primeira fita de cDNA foi utilizada as seguintes condições, seguindo recomendações do fabricante:

REAGENTES	QUANTIDADE
RNA Total	5 µg
PLAT3 10 µM	1µL
10x PCR Buffer	2,5 µL
MgCl ₂ 25 mM	2,5 µL
dNTP mix 10 mM	1µL
DTT 0,1 M	2,5µL
SS II RT (200 u/µL)	1µL
Água nuclease-free	q.s.p. 25µL

O RNA foi desnaturado a 70°C por 10 minutos, depois deixado em banho de gelo por um minuto, sendo submetido, após isso, a uma centrifugação de 13000 x g por um minuto. Logo em seguida, todos os reagentes foram adicionados com exceção da enzima, sendo equilibrados a 42 °C por um minuto, adicionando-se, por fim, 1 µL da enzima SS II RT e a reação completa foi incubada a 42 °C por 50 minutos. Logo após, a enzima foi desnaturada a 70°C por 15 minutos.

Após a síntese do cDNA, o molde de RNA foi removido da molécula híbrida cDNA:RNA por digestão pela RNase H (2 u) a 37 °C por 30 minutos.

3.2.3. Purificação do produto da primeira fita de cDNA em coluna S.N.A.P.TM

O excesso de nucleotídeos e de iniciadores devem ser removidos do produto da primeira fita de cDNA. Por outro lado, após a adição da cauda pela TdT, o iniciador específico poderia competir com o Abridged Anchor *primer* durante o início da PCR. Por conta da grande quantidade do iniciador específico em detrimento ao produto de cDNA, o procedimento com a coluna S.N.A.P.TM, adaptado por Vogelstein & Gillespie (1979), provê um meio rápido e eficiente de purificar o produto da primeira fita. De modo geral, na presença de um agente caotrópico, um cDNA com mais de 200 bases é ligado a uma membrana de sílica. Componentes tamponantes, dNTPs, enzimas e oligonucleotídeos remanescentes mantêm-se na solução e são removidos por centrifugação com o efluente. Impurezas residuais e iodato de sódio são removidos passando diversos volumes de tampão de lavagem, seguido por etanol 70% lavando a coluna. O cDNA purificado é então recuperado em água destilada e pode ser usado diretamente na reação de adição da cauda homopolimérica por TdT.

A partir do produto da primeira fita de cDNA (25 µL), adicionou-se 120 µL de *Binding Solution* (NaI 6 M) e aplicou-se o homogeinado na coluna S.N.A.P.TM, procedendo-se uma centrifugação a 13000 x *g* por 20 segundos. Lavou-se 4 x a coluna com 400 µL de *Wash buffer* (4 °C) e 2 x com etanol 70% (4°C). Ao fim, eluiu-se o cDNA purificado utilizando-se 50 µL de água destilada (65°C).

3.2.4. Adição da cauda poli (C) à extremidade 3' do cDNA purificado

A reação de adição da cauda poli (C) por TdT (Terminal deoxynucleotidyl Transferase) do sistema 5' RACE cria o sítio de ligação para o *Abridged Anchor Primer* na extremidade 3' do cDNA purificado. Para isto, foi utilizado o *tailling buffer* (Tris-HCl 10 mM pH 8,4, KCl 25 mM e MgCl₂ 5 mM), suplementado com 200 µM de dCTP para uma adição homopolimérica de cauda na primeira fita de cDNA, de acordo com as instruções do fabricante.

REAGENTE	QUANTIDADE
5x Tailing buffer	5 μ L
2mM dCTP	2.5 μ L
cDNA purificado SNAP	10 μ L
Água q.s.p.	24 μ L

O cDNA purificado adicionado de todos os reagentes acima (24 μ L) foi aquecido a 94°C durante 3 minutos e foi incubado a 0°C por 1 minuto, seguido de uma centrifugação a 13000 x g por 1 min e outra incubação a 0°C por 1 minuto. Adicionou-se 1 μ L de TdT e incubou-se a mistura a 27°C por 10 minutos e, após isso, a 65°C por mais 1 minuto, a fim de inativar a enzima. Centrifugou-se a 13000 x g por 1 minuto e incubou-se a 0°C.

3.2.5. Amplificação do produto 5' RACE da lectina de *P. elegans*

Após purificação e adição da cauda poli (C) ao cDNA, foi utilizada a estratégia de reação de polimerização em cadeia (PCR) para a amplificação do gene. Como fita molde foi utilizada a fita de cDNA à partir da técnica 5' RACE purificado e com cauda poli (C); como iniciadores, um iniciador específico para o gene (PLAT3, como antisense) e um iniciador oligo-dG (*Abridged Ancor primer*, AAP, sense). Para a reação de PCR foi utilizado o kit GoTaq® Green Master Mix: tampão de reação, dNTP mix 10 mM (cada), iniciadores para o gene (sense e anti-sense), DNA molde (cDNA), enzima GoTaq polimerase 5 U/ μ L, água estéril, seguindo instruções do fabricante:

Componente	Quantidade (μ L)
cDNA (< 0,5 μ g)	5
AAP (sense)	0,5
PLAT3 (anti-sense)	0,5
5 x GoTaq Buffer	10
PCR Nucleotide Mix (10 μ m cada)	1
GoTaq (5 u/ μ L)	0,250
Água estéril	q.s.p. 50

A reação de PCR conteve a seguinte condição de reação:

PASSO	TEMPERATURA (°C)	TEMPO	Nº CICLOS
Desnaturação Inicial	95	3'	1
Desnaturação	95	45''	35
Anelamento	50*	1'	35
Extensão	72	3'	35
Extensão Final	72	5'	1
Fim	16	-	1

*Obs.: A temperatura de anelamento varia conforme a temperatura de *melting* (Tm) dos oligonucleotídeos. A temperatura de anelamento deve ser em torno de 5° C menor que a do Tm mais baixo.

O produto final foi analisado em gel de agarose a 1,2%.

3.2.6. Purificação, Clonagem, Sequenciamento e Análise do produto amplificado 5' RACE da lectina de *P. elegans*

A purificação, clonagem no vetor pGEM-T easy, transformação em células competentes de *E. coli* XL1 blue, sequenciamento e análise do produto 5' RACE foram todos realizados de forma similar ao descrito para o produto 3' RACE.

O produto 5' RACE clonado em pGEM-T easy continha toda a sequência codificante completa para a lectina de *P. elegans* (PELa), sendo então usado para subclonagem em vetor de expressão.

4. Determinação da Sequência Protéica e Alinhamento de Lectinas de Leguminosas

A sequência de aminoácidos de PELa (número de acesso: AEK69351) foi deduzida com base na sequência de seu mRNA, obtida nos passos anteriores pelas técnicas 3' e 5' RACE, e comparada contra um banco de sequências de lectinas de leguminosas pela utilização do algoritmo BLAST, sendo analisadas similaridades entre a sequência dessa lectina com as de outras tribos. Algumas lectinas intimamente relacionadas (maior similaridade) que continham

representantes estruturais foram alinhadas para a análise da conservação dos resíduos envolvidos na estruturação dos sítios de ligação a metal (cálcio e manganês) e a carboidrato.

Foram selecionadas as seguintes lectinas: lectina de sementes de *Pterocarpus angolensis* (PAL – número de acesso: CAD19804), lectina de sementes de *Canavalia ensiformes* (ConA – número de acesso: CAA25787), lectina de sementes de *Arachia hypogaeae* (PNA – número de acesso: 2PEL_A), lectina da casca de *Robinia pseudoacacia* (RPbAI – número de acesso: BAA36415) e lectina de *Erythrina collarodendrom* (EcorL – número de acesso: CAA36986), sendo a primeira pertencente a subtribo Dalbergiae, as duas seguintes a tribo Phaseolae-Dioleinae (as duas primeiras específicas para Man/Glc; a terceira para Gal), a próxima Robineae (específica para Gal) e a última Phaseolae-Dioleinae (específica para Gal/GalNAc). O alinhamento foi feito na ferramenta on line Esript (<http://esript.ibcp.fr/ESript/cgi-bin/ESript.cgi>).

5. Construção do Vetor de Expressão e Transformação em células de células de *E. coli* BL21 competentes

Para a expressão da proteína recombinante foi selecionado o vetor pET-29a (Invitrogen) (Figura 11), o qual pode originar uma proteína de fusão formada pela proteína alvo acrescida da proteína N-terminal S Tag / sítio de trombina ou uma cauda c-terminal de histidina (histag – 6 X histidina), embora a estratégia adotada tenha dispensado qualquer das duas possibilidades.

Com o intuito de construir um vetor que expressasse apenas a proteína alvo, foi utilizada a estratégia de inserir o gene inteiro de PELA no vetor pET-29a, numa região interna àquelas codificadoras de proteína de fusão, entre os sítios das enzimas de restrição *NdeI* e *HindIII*. Para isso, foi necessário amplificar o gene inteiro com dois iniciadores que foram definidos, possuindo em suas extremidades 5' os sítios de clonagem *NdeI* e *HindIII*, permitindo a inserção orientada do inserto no vetor, como descrito a seguir.

5.1. Desenho de primers específicos para o gene de PELA adicionados de sítios de restrição

Além dos dois sítios de restrição específicos, *NdeI* e *HindIII*, para que só a proteína alvo fosse expressa, foi necessário, juntamente com o inserto, adicionar códons de iniciação (*start*) e

de parada (*stop*) para a correta leitura do código e expressão da proteína. Dois iniciadores foram então desenhados para flanquearem e amplificarem o gene com essas informações (Tabela 6): PLAT4 (5'GGAATTCCCATATGCTACTGAATAGGGCAAAC3') e PLAT5 (5'GGCAAAGCTTTCACATATCACGTGCAAGATAC3'). PLAT 4 possui um sítio de restrição para a enzima *NdeI* (sublinhado), apresentando internamente um códon *start* (ATG), seguido do código para a sequência aminoacídica MLLNRA. PLAT5 possui, de modo reverso e complementar, um sítio de restrição para a enzima *HindIII* (sublinhado), um códon *stop* e o código para a sequência aminoacídica LARDM. Além disso, as enzimas de restrição exigem adicionais estruturais de suporte para a ligação e consequente corte de DNA, que deve ser fornecida por nucleotídeos vizinhos. Dessa maneira, 3-7 nucleotídeos foram adicionados na extremidade 5' de cada iniciador, antes do sítio de restrição, com sequências da preferência requerida para cada enzima.

Tabela 6. Iniciadores utilizados na clonagem do gene de PELa no vetor de expressão pET-29a

Iniciador	5'- Sequência – 3'*	Sequência de aminoácidos correspondentes
PLAT4 (senso)	GGAATTCC <u>CATATG</u> CTACTGAATAGGGCAAAC	(<i>NdeI</i>)MLLNRA
PLAT5 (antisenso)	GGCA <u>AAGCTT</u> TCACATATCACGTGCAAGATAC	LARDM Stop (<i>HindIII</i>)

*As sequência sublinhadas indicam a posição dos sítios *NdeI* et *HindIII*, respectivamente.

5.2. Amplificação do inserto para construção de vetor de expressão

A partir do produto 5' RACE, a região codificadora do gene de PELa foi clonado no vetor pGEM-T *easy*, utilizando-se a estratégia de reação de polimerização em cadeia (PCR) para a amplificação do inserto para vetor de expressão com os iniciadores PLAT4 (senso) e PLAT5 (antisenso).

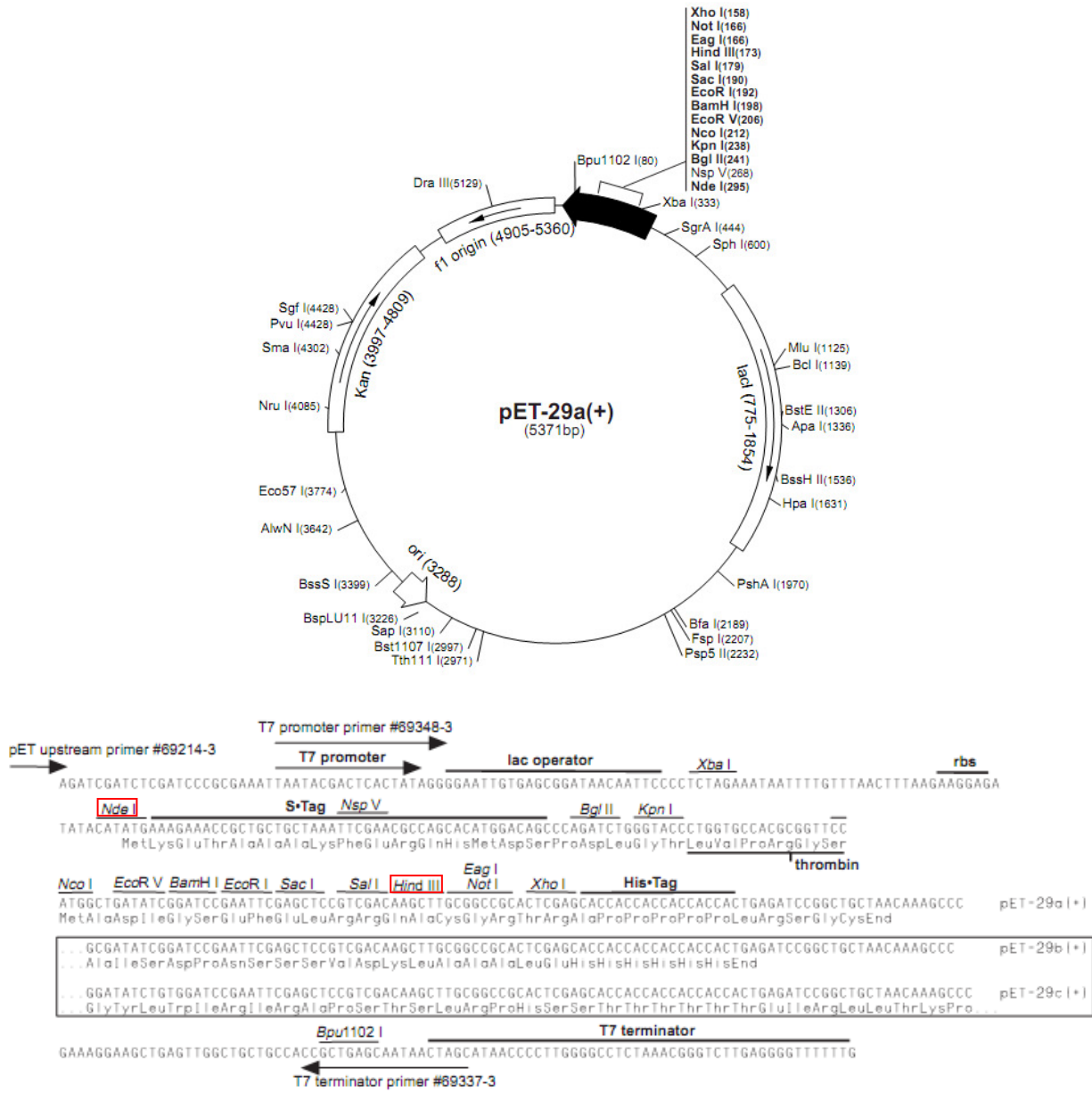


Figura 11. Mapa do vetor pET-29a (Novagen). Sítios de restrição para as enzimas *Nde*I e *Hind*III em vermelho.

Para a reação de PCR foi utilizado o kit GoTaq® Green Master Mix: tampão de reação, dNTP mix 10 mM (cada), iniciadores para o gene (senso e anti-senso), DNA molde (cDNA), enzima GoTaq polimerase 5 U/μL, água estéril, seguindo instruções do fabricante:

Componente	Quantidade (µL)
Plasmídeos	5
PLAT4 1 µg/µL (senso)	0,5
PLAT5 1 µg/µL (anti-senso)	0,5
5 x GoTaq Buffer	10
PCR Nucleotide Mix (10 µM cada)	1
GoTaq (5 u/µL)	0,250
Água estéril	q.s.p. 50

A reação de PCR conteve a seguinte condição de reação:

PASSO	TEMPERATURA (°C)	TEMPO	N° CICLOS
Desnaturação Inicial	95	3'	1
Desnaturação	95	45''	35
Anelamento	60*	1'	35
Extensão	72	3'	35
Extensão Final	72	5'	1
Fim	16	-	1

*Obs.: A temperatura de anelamento varia conforme a temperatura de *melting* (T_m) dos oligonucleotídeos. A temperatura de anelamento deve ser em torno de 5°C menor que a do T_m mais baixo.

O produto final foi analisado em gel de agarose a 1,2% e sequencialmente purificado pelo *QIAquick Gel Extraction Kit* (Qiagen), como descrito anteriormente.

5.3. Digestão enzimática do inserto purificado e do vetor pET-29a

O inserto purificado no passo anterior e o vetor pET-29a vazio, antes de serem ligados, precisavam ser digeridos pelas enzimas de restrição *NdeI* e *HindIII* (Invitrogen), com a finalidade de deixar as mesmas extremidades coesivas para ambos. A digestão ocorreu

separadamente, a uma temperatura de 37 °C, por 3 horas para o inserto e 4 horas para o vetor, seguindo as proporções dos reagentes a seguir:

	Inserto Purificado	Vecteur pET-TEV
DNA	25 µL	28,5 µL
<i>NdeI</i>	1,5 µL	1,5 µL
<i>HindIII</i>	1,5 µL	1,5 µL
Tampão <i>React 2*</i> 10x	3,5 µL	3,5 µL
Água nuclease free q.s.p.	35 µL	35 µL

*Tampão fornecido pelo fabricante em que ambas as enzimas funcionam em 100% de eficiência.

O inserto digerido foi purificado através do *QIAquick Gel Extraction Kit* (Qiagen), como descrito anteriormente, assim como o vetor digerido foi purificado pelo método *Wizard® Plus Minipreps DNA Purification System* (Promega), também já descrito.

5.4. Clonagem do inserto no vetor pET-29a

Após digestão e purificação, o processo de clonagem do gene foi realizado com auxílio do *DNA ligation KIT* (TAKARA), tampão de ligação (*ligation mix*) e uma proporção 25:25~250 fmol vetor/inserto. O procedimento foi executado de acordo com recomendações do fabricante, a 25°C por 5 minutos:

REAGENTE	QUANTIDADE
Vetor	1 µL
Inserto	5 µL
<i>Ligation Mix</i>	6 µL
TOTAL	12 µL

5.5. Transformação do plasmídeo de expressão em células de *E. coli* BL 21 competentes

Após clonagem, o plasmídeo resultante da ligação (pET-29a_PELa) foi inserido em células de *E. coli* BL21 competentes, segundo protocolo de transformação já descrito, com modificação para o antibiótico de resistência, que no caso do pET-29a é a kanamicina, e o uso dispensado de X-GAL e IPTG. Colônias positivas foram submetidas a validação e posterior sequenciamento, para então prosseguir-se aos experimentos de indução da expressão da proteína recombinante.

6. Expressão recombinante da lectina de sementes de *P. elegans*

Após confirmação de clones positivos de *E. coli* BL21 transformados com o plasmídeo pET29-a contendo a sequência correta para a lectina de *P. elegans* (PELa), foi realizado um experimento piloto de expressão, mediante SDS-PAGE, (LAEMLLI, 1970) com uma comparação dos lisados celulares, para a escolha da melhor condição para indução.

As bactérias transformadas foram pré-cultivadas em 10 mL de meio líquido LB (contendo 30 µg/mL de kanamicina) a 37°C, sob agitação constante (180 rpm), por um período de 16 horas. Em seguida, o material foi totalmente inoculado em 100 mL de meio LB com 30 µg/mL de kanamicina, até ser atingida uma OD_{600nm} entre 0,6-0,8, quando então foi adicionado ao meio o indutor de expressão isopropil-β-tiogalactosidase (IPTG) a uma concentração de 0,5 mM, permanecendo a cultura em incubação a 16, 20, 25, 30 ou 37°C por até 4 horas ou período *overnight*, sendo retirada uma alíquota de 1 mL nos tempos intermediários de cada condição de indução. No final, as células foram sedimentadas por centrifugação a 6000 rpm durante 20 minutos a 4°C, sendo em seguida ressuspensas em tampão TRIS (Tris-HCl 50 mM, NaCl 150 mM, CaCl₂ 5 mM, MnCl₂ 5mM, pH 7,6) e lisadas por ruptura mecânica, utilizando-se para isso pressão de 1700 bars em um *One Shot Cell Disrupter (Constant System)*. A separação das frações solúvel e insolúvel foi realizada mediante centrifugação a 12000 x g por 30 minutos a 4°C. O sobrenadante foi coletado (fração solúvel) e o sedimento (fração insolúvel) solubilizado diretamente em tampão de amostra 5 x concentrado numa proporção 1:5. Procedeu-se então a análise das frações por SDS-PAGE 12% (LAEMLLI, 1970). Após as avaliações iniciais foram

feitas induções maiores para purificação em larga escala a temperatura de 30°C por um período de 4 horas.

7. *Refolding* e purificação da lectina recombinante de *P. elegans*

Como uma maior proporção da lectina recombinante encontrava-se na fração insolúvel para todas as condições de indução testadas, optou-se por estratégia de purificação isolar a proteína a partir de corpos de inclusão.

Corpos de inclusão produzidos em *E. coli* são normalmente compostos de moléculas protéicas desnaturadas empacotadas em forma de partículas. *Refolding* de proteínas a partir de corpos de inclusão em suas formas bioativas pode ser laborioso, resultar em um baixo rendimento e quantidade para um maior custo de uma produção recombinante. Com novas informações disponíveis, sobre estrutura e função de agregados protéicos em corpos de inclusão bacterianos, tem sido possível desenvolver melhores procedimentos de solubilização e *refolding* com uma maior recuperação da proteína bioativa.

A recuperação de proteína ativa solúvel de corpos de inclusão purificados requer a desnaturação do polipeptídeo e então seu *refolding* a uma forma ativa. Muitos exemplos de proteínas recuperadas de corpos de inclusão são bem conhecidos e usados tanto para aplicações comerciais como acadêmicas. Existem métodos bem estabelecidos para purificar corpos de inclusão e solubilizar a proteína agregada por desnaturação. Há, no entanto, nenhum método confiável para prever as condições necessárias para redobrar a proteína. Assim, a identificação das condições necessárias ao adequado *refolding* da proteína continua a ser uma ciência empírica. Portanto, com o objetivo de ajudar a simplificar o processo de identificação da composição do melhor tampão e do método mais adequado para o *refolding* proteína, foi utilizado o TM *Kit Proteína QuickFold Screening* (AthenaES). O kit contém 15 composições de tampões diferentes que permitiram uma rápida identificação dos fatores que têm um efeito importante sobre o *refolding* da proteína. A partir desta informação os experimentos puderam ser realizados para determinar a formulação ideal de tampão. No parágrafo abaixo, a descrição do protocolo final estabelecido como o mais adequado para o *refolding* da proteína.

A partir de um experimento de indução a 30°C por 4 h, as células foram precipitadas e ressuspendidas como descrito anteriormente, numa proporção de 10 mL do tampão de lise Tris-

HCl 50 mM, NaCl 150 mM, CaCl₂ 5 mM, MnCl₂ 5mM, pH 7,6 a cada grama de precipitado celular. Após ruptura celular, foram separadas por centrifugação as frações solúvel (sobrenadante) e insolúvel (precipitado); o sobrenadante foi descartado e o precipitado foi submetido a uma estratégia de lavagem e purificação. O precipitado foi então lavado 2 x em Tris-Cl 50 mM pH 8,0, NaCl 0,5 M, Uréia 4 M, EDTA 1 mM, deoxicolato 1 mg/mL, numa proporção de 5 mL a cada grama de precipitado celular inicial, seguido de centrifugação a 20000 x g por 15 minutos a 4°C. O precipitado final foi então submetido a solubilização em Tris-Cl 50 mM pH 8,0, Urea 8 M, DTT 10 mM, numa proporção de 2 mL por grama de precipitado celular inicial, e mantido a 50°C por cerca de 10-20 minutos. Após centrifugação a 20000 x g, 30 minutos, 4°C, o sobrenadante solubilizado teve sua concentração protéica ajustada para em torno de 1 mg/mL, mensurado por absorvância em 280 nm. Após o ajuste, a solução foi então diluída gota a gota em Tris-HCl 50 mM pH 7,5, NaCl 240 mM, MnCl₂ 2 mM, CaCl₂ 2 mM, KCl 10 mM, DTT 1 mM, numa proporção final de 1:20, sob agitação lenta. Nesse momento o *refolding* começa a ocorrer, podendo-se visualizar uma consistência turva na solução, devido ao material mal enovelado em suspensão. O sistema foi então deixado em repouso por 1h a 10°C e depois submetido a uma filtração em papel de filtro comum para separar eventuais partículas em suspensão.

Logo em seguida à filtração, toda a solução diluída e filtrada foi aplicada em uma coluna de manose-agarose pré-equilibrada com Tris-HCl 50 mM pH 7,5, NaCl 0,15 M, MnCl₂ 2 mM, CaCl₂ 2 mM. Após lavagem da coluna com a mesma solução, a lectina refoldada e ativa foi então eluída com a adição de manose 200 mM. Todas as frações coletadas foram mensuradas e analisadas qualita e quantitativamente através de absorvância a 280nm.

A proteína purificada (fração retida) foi então lavada e concentrada como o desejado para cada análise utilizando um dispositivo de ultrafiltração Vivaspin (cutoff 10 kDa; GE Healthcare), sendo depois armazenada a 4°C.

Todos os passos até a obtenção da lectina recombinante purificada foram analisados em eletroforese em gel de poliacrilamida a 12% em condição desnaturante com uso de β -mercaptoetanol e dodecil sulfato de sódio (SDS-PAGE), seguindo-se o método descrito por Laemmli (1970), adaptado para o uso de placas de vidro verticais, descrito anteriormente.

II. Ferramentas para o estudo de interação lectina-carboidrato

1. Atividade hemaglutinante da lectina PELa

A determinação da atividade hemaglutinante (A.H.), foi feita segundo Moreira & Perrone (1977), com algumas modificações descritas a seguir. O ensaio de atividade hemaglutinante da lectina recombinante foi realizada através de diluições (50 µL) seriadas em duplicatas (1/2, 1/4, 1/8, 1/16, 1/32, etc.) em placas de microtitulação com fundo em V, contendo 50 µL tampão Tris-HCl 50 mM, pH 7,5, com NaCl 0,15 M. A 50 µL de cada diluição foi adicionado igual volume de uma suspensão de eritrócitos de coelho (fornecida por BioMérieux) a 2%, levemente agitados e incubados a 37°C por 30 minutos. A mistura foi deixada em repouso por mais 30 minutos à temperatura ambiente e a aglutinação foi observada macroscopicamente. A menor concentração protéica que apresentou hemaglutinação visível é considerada a concentração correspondente a uma unidade de hemaglutinação (UH), sendo o título expresso em unidades de hemaglutinação (UH/mL).

2. Análise da especificidade de PELa por *Glycan-Array*

Glycan Arrays tornaram-se uma poderosa ferramenta para a elucidação de alto nível das interações de carboidratos de diferentes estruturas, com uma grande variedade de alvos biológicos, incluindo anticorpos, proteínas, vírus e células. Esta técnica é especialmente adequada para estudos glicômicos, pois as matrizes apresentam ligantes de carboidratos de uma maneira que imita as interfaces de interações célula-célula (Figura 12) (LIANG & WU, 2009).

O *Glycan-Array* é baseado no princípio de ELISA (*Enzyme-Linked ImmunoSorbent Assay*). Os primeiros testes compreendiam uma gama de oligossacarídeos biotinilados fixados por um braço espaçador em placas de 96 a 384 poços, recobertos por estreptavidina, permitindo realizar a análise em paralelo sobre uma grande diversidade de açúcares de forma minorizada e rápida (Figura 13). Permitem determinar a especificidade das proteínas que se ligam aos carboidratos, como lectinas. Para esse tipo de ensaio, lectinas devem ser marcadas por um composto fluorescente, ou serem fornecidas com um anticorpo (BLITX *et al.*, 2004).

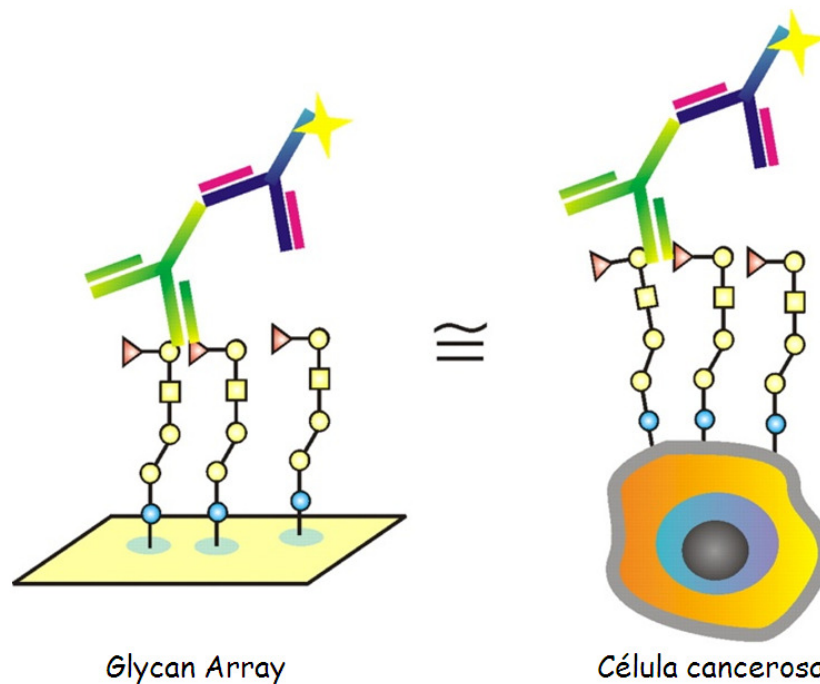


Figura 12. *Glycan array* como uma imitação da superfície celular, expressando glicanos para estudo de interações multivalentes (WANG et al., 2008).

Métodos mais recentes propostos pelo *Consortium for Functional Glycomics* (<http://www.functionalglycomics.org>) e por outros laboratórios baseiam-se no princípio de “*glicochips*” ou *Printed Covalent Glycan Array*. Nesse ensaio, utiliza-se placas de vidro comercial funcionalizadas com N-hydroxysuccimidine (NHS), que é capaz de ligar-se covalente a glicanos ou glicoconjugados, desde que sejam acoplados a um grupo amina (BLIXT *et al.*, 2004). Esse método realiza a análise em paralelo sobre uma grande diversidade de açúcares de forma minorizada e rápida, permitindo determinar a especificidade das proteínas que se ligam aos carboidratos, como lectinas (Figura 13).

Para realização desse ensaio, a lectina é marcada pelo composto fluoroiscenado Alexa Fluor® 488 – TFP (ácido carboxílico – 2,3,5,6 tetrafluorofenil), segundo recomendações do fabricante (Invitrogen). A lectina marcada é separada do excesso do fluoróforo através de filtração em gel em D-Salt Poliacrylamide. A análise é feita pelo *Consortium for Functional Glycomics* (CFG; <http://www.functionalglycomics.org>).

Para realização desse ensaio, PELA foi preparada numa concentração de 1 mg/mL em tampão bicarbonato de sódio 0,1 M pH 9,0. O composto TFP foi solubilizado em DMF (dimetilformamida) numa concentração de 10 mg/mL instantes antes de sua utilização. A mistura

na proporção 1:10 TFP/proteína foi incubada uma hora em temperatura ambiente com agitação. A proteína foi finalmente separada do reagente residual por dessalinização em uma coluna de poliacrilamida D-saltTM (Pierce) pré-equilibrada com Tris-HCl 50 mM pH 7,5, NaCl 0,15 M, MnCl₂ 2 mM, CaCl₂ 2 mM, e a concentração final foi de 1,2 mg/mL de proteína. Após a marcação, a amostra foi enviada para o *Consortium for Functional Glycomics*, CFG, core H, Printed Array Ver 4.1 (<http://www.functionalglycomics.org>).

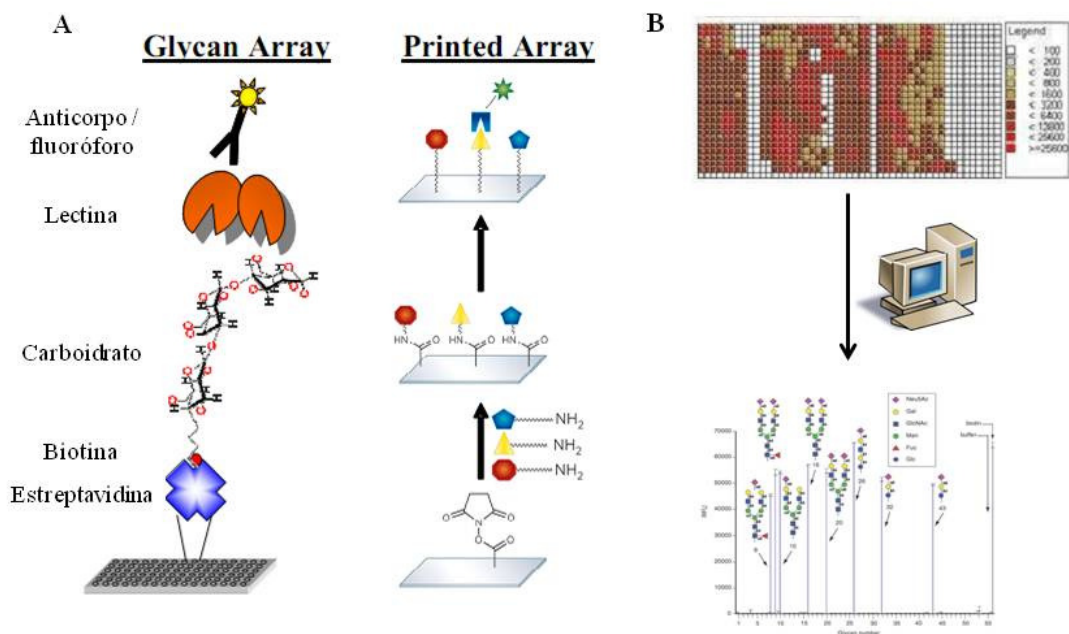


Figura 13. Descrição de *glicochips*. A. Diagrama das duas técnicas utilizadas para fazer *glicochips*. À esquerda é representada o princípio de *glycan array*, onde o açúcar biotilado é fixado a uma placa funcionalizada com estreptavidina. À direita, ligantes naturais ou sintéticos são covalentemente ligados a uma placa de vidro funcionalizada com NHS. Em ambos os casos, a fixação da lectina é revelada pelo uso de um anticorpo fluorescente, ou diretamente através da marcação da lectina por Alexafluor488. B. As intensidades de fluorescência são processadas por um computador e representadas sob a forma de um histograma baseado nos açúcares testados.

3. Ensaios de microcalorimetria- *Isothermal Titration Microcalorimetry* (ITC)

Entre os vários métodos para o estudo da termodinâmica do reconhecimento entre um receptor e seu ligante (açúcar-proteína, proteína-proteína ou outros), a microcalorimetria de titulação isotérmica é a única maneira de obter todos os parâmetros que caracterizam a interação em um único experimento, como a constante de afinidade (K_a), a estequiometria da interação (n) e as contribuições entálpicas (ΔH) e entrópicas ($T\Delta S$). Além do mais, ao contrário de outros

métodos que exigem a imobilização do ligante ou receptor, que por vezes podem interferir na interpretação dos resultados, a microcalorimetria permite que as moléculas estejam livres em solução.

3.1. Princípio da Técnica

O calorímetro é um dispositivo composto por duas células no interior de uma câmara adiabática (VP-ITC Microcal®). A primeira célula, de referência, contém o tampão utilizado no experimento, enquanto que a segunda, chamado de célula de medição, contém a proteína diluída no mesmo tampão que a célula de referência (Figura 14). Durante a titulação, as duas células são mantidas à mesma temperatura. Com uma seringa, pequenas quantidades de ligante são injetados na célula de medição em intervalos regulares. Cada adição de ligante provoca uma troca de calor refletindo a interação macromolécula-ligante. O sinal gerado é proporcional à quantidade de complexo formado. A máquina mede a quantidade de energia necessária para trazer as duas células à mesma temperatura. A medida é a potência elétrica fornecida ao longo do tempo para cada injeção ($\mu\text{cal}/\text{sec}$), e depois ela é integrada.

No início do experimento, o número de sítios disponíveis é superior à quantidade de ligante injetado: o sinal medido portanto é alto. Com o curso da injeção, vai diminuindo o número de sítios disponíveis, resultando em uma diminuição no sinal. As quantidades medidas de calor permitem obter ao final um termograma experimental (Figura 15). Quando as condições experimentais são ótimas, a curva obtida é de forma sigmóide. Um software de processamento de resultados (Origin®) calcula uma curva teórica que se sobrepõe aos pontos experimentais. A partir da curva é possível extrair os parâmetros da interação. As informações sobre o ponto de inflexão indicam a estequiometria, a magnitude da curva informa sobre a entalpia, enquanto que a inclinação da curva indica a afinidade (Figura 15). A partir desses valores e usando as equações da termodinâmica, é possível encontrar outros parâmetros tais como a entropia da interação.

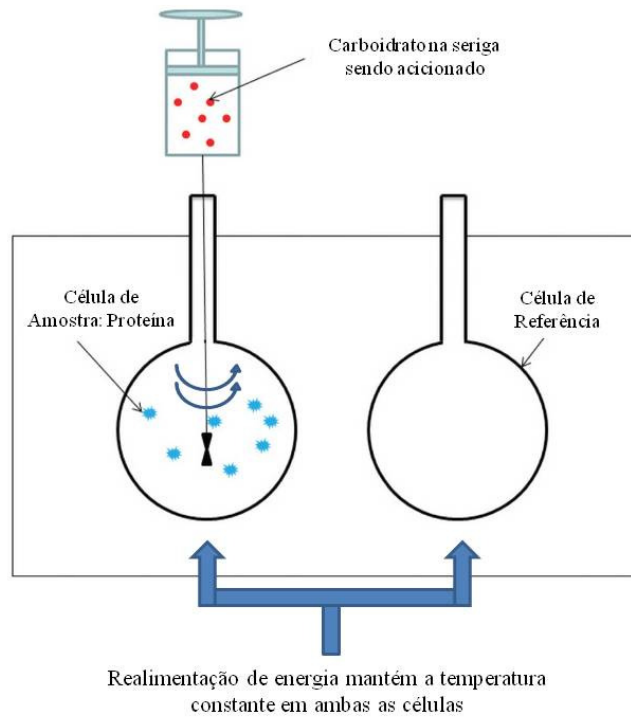


Figura 14. Representação esquemática de um microcalorímetro.

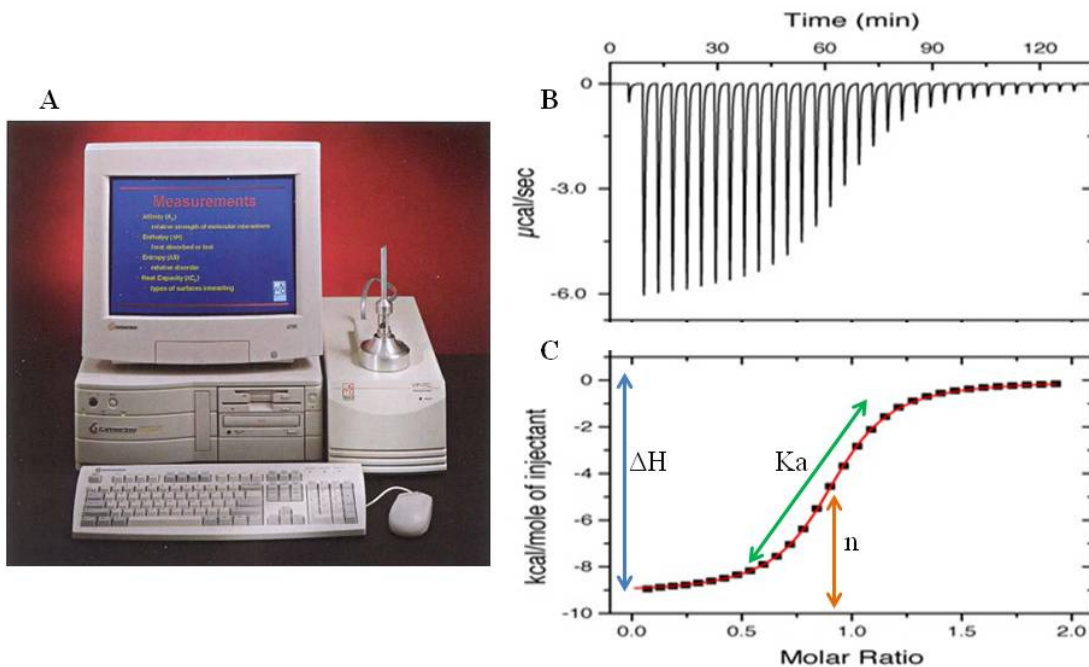


Figura 15. Princípio da microcalorimetria de titulação isotérmica. **A.** Exemplo de um microcalorímetro de titulação isotérmico. **B.** Exemplo de um termograma experimental. O aparelho mede a energia necessária para trazer as duas células à mesma temperatura. Quanto mais os sítios são ocupados, mais a energia diminui. **C.** Superposição da curva teórica dos pontos experimentais para obter as constantes termodinâmicas: estequiometria (n), entalpia (ΔH) e constante de afinidade (K_a).

3.2. Noções de Termodinâmica

A termodinâmica pode ser definida como a ciência que estuda fenômenos que dependem da temperatura e suas alterações. Em uma interação não-covalente entre uma macromolécula (M) e seu ligante (L), há um equilíbrio notável entre a forma complexada (ML) e as formas livres (M e L). Este equilíbrio depende da força de interação entre a macromolécula e o ligante e é definido por sua constante de afinidade K_a (M^{-1}). Quanto maior o valor desta constante, mais alta é a afinidade. Essa constante de associação é inversamente proporcional à constante de dissociação (K_d), que pode ser calculado da seguinte forma:

$$K_a = \frac{[ML]}{[M][L]} = \frac{1}{K_d}$$

A energia livre de Gibbs (G), a uma dada temperatura, pode ser calculado a partir da seguinte equação:

$$\Delta G = -R.T.\ln K_a$$

onde T é a temperatura absoluta em Kelvin (K), R é a constante dos gases perfeitos ($8.314 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$) e K_a é a constante de associação ou afinidade (M^{-1}).

A energia livre resulta da energia final gerada pela reação e permite caracterizar a evolução do sistema. Seu valor diz respeito a espontaneidade de uma reação química. Sempre que a variação de ΔG é negativa, por exemplo, a reação química é exergônica (ou seja, vai fornecer energia ao sistema), sendo portanto espontânea.

Para obter um estudo mais detalhado da termodinâmica é necessário medir a entalpia livre, a fim de quantificar a exata contribuição entálpica e endotérmica ($T\Delta S$) da reação:

$$\Delta G = \Delta H - T\Delta S$$

onde ΔG é a variação de energia livre ($\text{kJ}\cdot\text{mol}^{-1}$ ou $\text{kcal}\cdot\text{mol}^{-1}$), ΔH é a variação de entalpia ($\text{kJ}\cdot\text{mol}^{-1}$ ou $\text{kcal}\cdot\text{mol}^{-1}$) e ΔS é a variação de entropia ($\text{kJ}\cdot\text{mol}^{-1}$ ou $\text{kcal}\cdot\text{mol}^{-1}$).

A entalpia é definida como uma grandeza termodinâmica cuja variação reflete a quantidade de calor posto em jogo durante uma transformação isobárica. Os sistemas buscam ter a energia mais baixa possível; ΔH tende portanto a ser negativo. A entropia, por outro lado, reflete a desordem de um sistema. O aumento da temperatura favorece os movimentos brownianos e aumenta o efeito entrópico da reação, tendo o efeito de influenciar diretamente as variações de energia livre.

Em sistemas biológicos, a entalpia reflete a força das interações geradas pela proteína com seu ligante em comparação com aqueles mantidos com o solvente. Elas são principalmente ligações de hidrogênio ou interações de Van der Waals. Uma entalpia de interação favorável requer a colocação correta dos grupos de doadores e receptores de elétrons no sítio de ligação. A contribuição entrópica, por sua vez, é devido a dois fenômenos. O primeiro é devido ao rearranjo de moléculas de água deslocadas pelo ligante quando este último estiver posicionado em seu local de ligação. Quando a interação envolve contatos hidrofóbicos, a contribuição será, neste caso, mais favorável. O segundo é causado por uma perda de graus de liberdade do ligante ou proteína. Um dissacarídeo em solução, por exemplo, pode adotar várias conformações, devido à rotação em torno do ângulo de torção, mas no momento de interação, será forçado a uma conformação única. Esta contribuição entrópica é geralmente negativa.

A interação entre uma lectina e um açúcar geralmente provoca uma forte entalpia, devido às inúmeras ligações de hidrogênio, e uma contribuição entrópica desfavorável causada pelas mudanças conformacionais.

3.3. Análise da afinidade da lectina de *P. elegans* por oligossacarídeos através de ITC

PELa teve sua afinidade por alguns oligossacarídeos definidos à partir do resultado de *Glycan array* mensurada por microcalorimetria isotérmica de titulação (ITC). Os experimentos foram realizados com um *VP-ITC Isothermal Titration Calorimeter* (Microcal), a $25,0 \pm 0,1^\circ\text{C}$. A concentração da proteína na célula microcalorimétrica (1,4478 mL) foi em torno de 1 mM em Tris-HCl 50 mM pH 7,5, NaCl 0,15 M, MnCl_2 2 mM, CaCl_2 2 mM. Dois heptassacarídeos disponíveis para o projeto (preparados na mesma solução da proteína) a uma concentração de 10 mM (10 vezes superior a de lectina) foram distintamente adicionados em 30 injeções de 10 μL cada. O controle consistiu da adição das soluções de açúcares na ausência da proteína, o que rendeu mudanças insignificantes na entalpia.

A curva integrada do calor liberado foi analisada por regressão não-linear, utilizando-se um modelo de um ou dois sítios de ligação (Origin 7.0). Os dados experimentais foram fixados em uma curva de titulação teórica usando, fornecendo a constante associação (K_a), a entalpia de ligação (ΔH) e o número de sítios de ligação por monômero (n). Os parâmetros termodinâmicos, tais como variação de energia livre (ΔG) e de entropia (ΔS) foram calculados a partir da equação $\Delta G = \Delta H - T\Delta S = -RT \ln K_a$. Todos os experimentos foram realizados com valores de $10^{-6} < c < 10^{-2}$ (WISEMAN *et al.*, 1989).

III. Estudos da estrutura cristalina de PELa através de Cristalografia por Difração em raios X

A cristalografia de raios-X é provavelmente o método de análise estrutural mais poderoso na elucidação de estruturas tridimensionais de moléculas e macromoléculas biológicas. A aplicação da cristalografia de macromoléculas para o estudo estrutural, especialmente de proteínas, deve o seu nascimento aos estudos de Max Perutz e John Kendrew, que receberam o Prêmio Nobel em química (1962) pelo seu trabalho sobre a determinação da estrutura tridimensional da hemoglobina e mioglobina. De 1959 até o presente, a biologia estrutural tem avançado consideravelmente.

O crescente poder dos computadores, a facilidade de acesso à radiação síncrotron e, especialmente, as grandes descobertas em biologia molecular, contribuíram para fazer da cristalografia uma tecnologia de ponta no estudo estrutural de biomoléculas. No entanto, a base teórica deste método, assim como algumas dificuldades encontradas desde o começo, não mudaram, mesmo se hoje a resolução de uma nova estrutura possua muito menos problemas técnicos do que há 30 anos. Em particular, a cristalização de uma macromolécula ainda é baseada em métodos tentativa-erro, que às vezes incluem certas condições experimentais muito difíceis de controlar. Uma vez que a proteína foi cristalizada, o problema de fase é a segunda dificuldade a ser resolvida com diferentes técnicas que foram desenvolvidas.

1. Cristalização – Obtenção dos Cristais de Proteína

Uma das partes mais críticas na resolução da estrutura de uma macromolécula biológica é a cristalização da mesma. Os principais passos envolvidos na resolução da estrutura a nível atômico de uma macromolécula biológica estão mostrados na figura 16. Entre todos estes passos, a cristalização é o menos entendido. Biocristalização é um processo de múltiplos parâmetros, que envolve os três passos clássicos da cristalização: nucleação, crescimento e cessação de crescimento (BOISTELLE & ASTIER, 1988).

Para crescer cristais de qualquer composto, um estado de supersaturação tem que ser alcançado, um estado que é termodinamicamente instável, no qual pode se desenvolver ou numa fase amorfa ou cristalina quando este retornar ao equilíbrio. Supersaturação pode ser alcançada

pela evaporação lenta do solvente ou pela variação de alguns dos parâmetros como temperatura, pH e força iônica. O conhecimento das solubilidades macromoleculares é um pré-requisito para o controle das condições de cristalização.

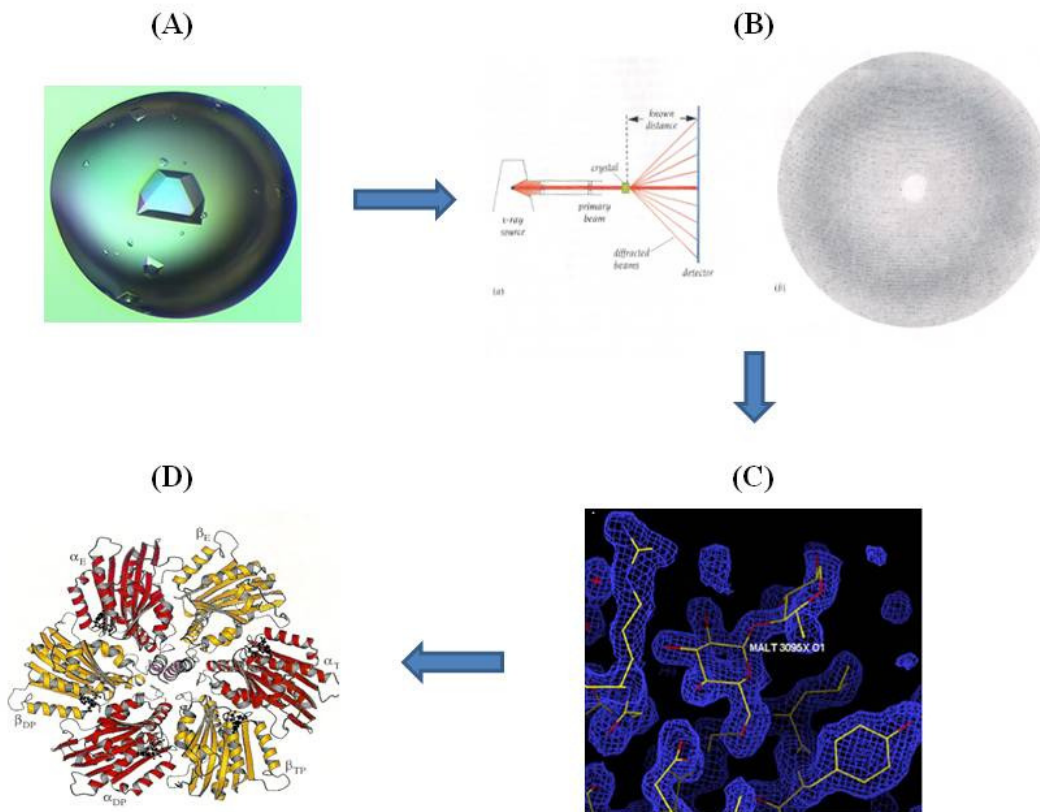


Figura 16. Principais passos na resolução da estrutura de uma biomolécula por cristalografia. A. Cristalização. B. Coleta de dados por difração de raios X. C. Resolução da estrutura. D. Análise da estrutura.

Existem muitas técnicas para a cristalização de macromoléculas biológicas, todas tendo como objetivo trazer a solução a um estado de supersaturação. A fim de obter cristais de proteínas, um método frequentemente utilizado é a difusão de vapor (Figura 17). Esta técnica foi utilizada pela primeira vez na cristalização do tRNA (HAMPEL *et al.*, 1968). Nesta técnica, uma gota contendo poucos microlitros da macromolécula biológica a ser cristalizada em um agente de cristalização, tampão e aditivos, é equilibrada contra um reservatório contendo a solução de cristalização a uma concentração mais alta do que na gota e um volume largamente maior. O equilíbrio prossegue por meio da difusão das espécies voláteis (água e solventes orgânicos) até que a pressão de vapor na gota se iguale àquela do reservatório. Se o equilíbrio ocorre por meio de troca da água (da gota para o reservatório), isto leva a uma diminuição do volume da gota, e

consequentemente a um aumento na concentração de todos os constituintes da gota de cristalização, inclusive da proteína. Para espécies químicas com pressão de vapor maior que a pressão de vapor da água, a troca ocorre do reservatório para a gota. Os mesmos princípios se aplicam para gotas suspensas (*hanging drops*), gotas sentadas (*sitting drops*) e gotas sanduíches (*sandwich drops*).

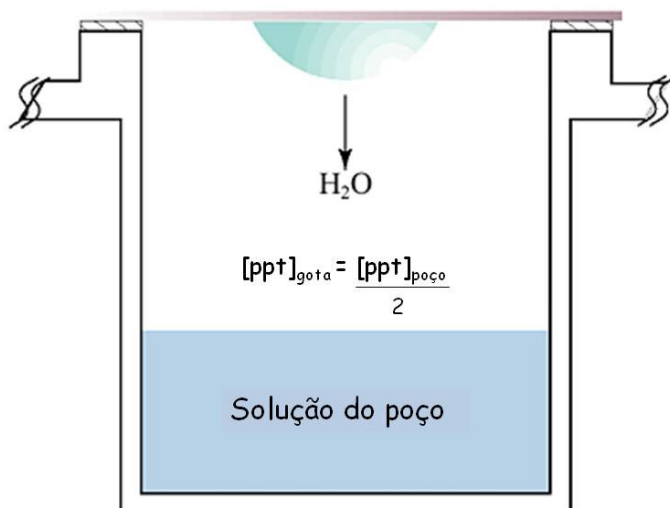


Figura 17. Princípio de cristalização pelo método de difusão de vapor - gota suspensa. A concentração do agente precipitante na gota é metade daquela no reservatório.

Com o aumento do número de macromoléculas biológicas cristalizadas com sucesso, tornou-se óbvio que muitas das condições de cristalização se assemelhavam, ou seja, havia uma concentração de resultados positivos de cristalização que usava um número limitado de precipitantes, tampões e aditivos. Isto levou à proposição de diversos métodos de cristalização (CARTER & CARTER, 1979), onde um número limitado de condições eram testadas, usando-se pequenas quantidades da macromolécula biológica, por volta de poucos miligramas. A partir da observação dos resultados preliminares destes experimentos era possível determinar que tampão, aditivo e agente precipitante seriam os mais favoráveis e a partir daí proceder sucessivos melhoramentos até se conseguir cristais adequados, ou ainda, em casos favoráveis, obter cristais adequados já na primeira tentativa com as condições padrões. Dentro deste raciocínio a Dra. Jaru Jancarick da UC Berkeley propôs o método da matriz esparsa (JANCARICK & KIM, 1991) onde diversas condições diferentes são tentadas para se cristalizar a macromolécula biológica. O número de variáveis que afetam a cristalização, tais como, concentração, temperatura, pH, força

iônica, aditivos específicos e precipitantes, é grande e combinatória; o número total de possíveis condições de cristalização a serem testadas, portanto, é tão grande que uma procura completa com todas as condições em potencial seria inviável. Kits comerciais, tais como *Screens Hampton Research*, triam várias condições, e muitas vezes são o ponto de partida para os ensaios.

Adicionalmente, caracterizar a interação proteína-ligante requer a obtenção de cristais em que a macromolécula é complexada com o ligante. Para isso, proteína e ligante são misturados na gota de modo que ambos interajam e se cristalizem juntos, num processo denominado de co-cristalização. Uma vez o ligante presente, o complexo proteína-ligante pode ainda não cristalizar nas mesmas condições experimentais que o cristal nativo, nem no mesmo grupo espacial. O método de *soaking*, em que um cristal de proteína nativa (não complexada) ou é mergulhado ou tem adicionado à gota uma solução contendo o ligante para que ele se difunda no cristal, é também uma alternativa, desde que a acessibilidade ao sítio de ligação não esteja bloqueada por moléculas vizinhas.

2. Difração em Raios X, coleta e processamento de dados

A coleta e o processamento dos dados à partir de difração do cristal em raios X é o próximo passo no objetivo de se obter a estrutura. O sucesso desta etapa depende da qualidade dos cristais e da fonte de radiação utilizada.

Nesta etapa, os cristais são caracterizados, ou seja, seus parâmetros de rede e suas simetrias são determinados (Figura 18). O conhecimento das amplitudes e das fases permite a reconstituição da densidade eletrônica do cristal. Quando um cristal é colocado em um feixe de raios X, a densidade eletrônica de cada átomo do cristal interage com a radiação, dispersando-a em todas as direções do espaço. É a presença de uma rede cristalina que amplifica a radiação somente em determinadas direções, dando origem ao padrão de difração característico.

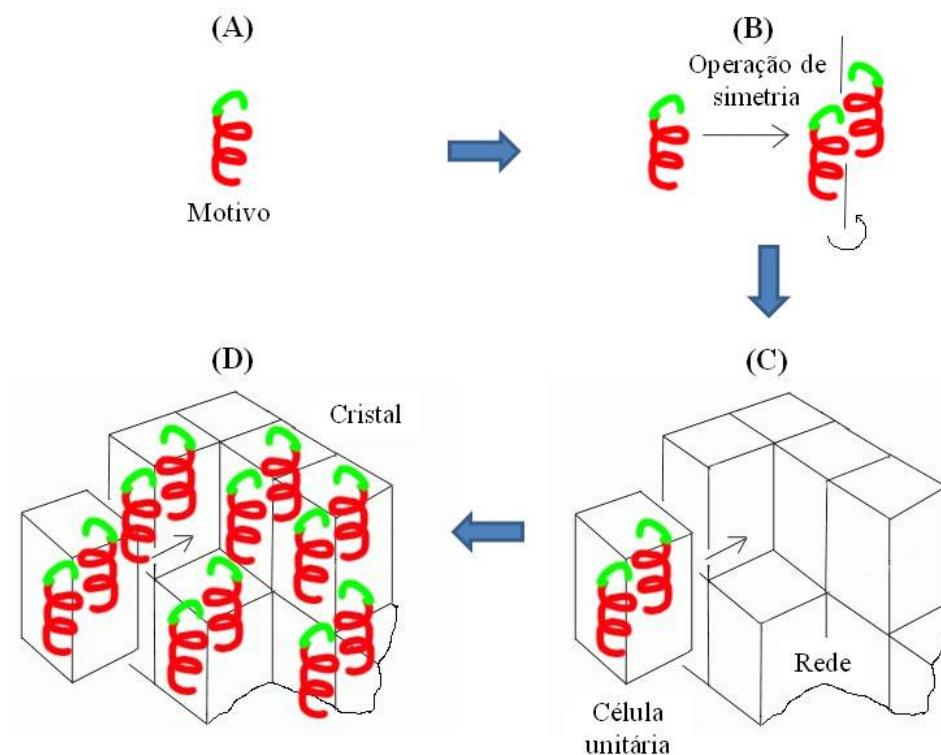


Figura 18. Formação e características de uma rede cristalina. Um cristal é um arranjo periódico de um motivo em uma estrutura (A). Se este motivo se repetir em três dimensões, percebe-se a maneira mais simples para formar um cristal. Muitas vezes o motivo, também conhecido como a unidade assimétrica, é submetido a uma série de operações de simetria, rendendo diferentes cópias em orientação (B). Se não houver operações de simetria adicionais, o conteúdo da célula unitária já está criado (C). O cristal é então construído a partir de células unitárias organizadas em uma rede tridimensional periódica (D) (Adaptado de Rupp, 2006).

As bases matemáticas e geométricas do processo de espalhamento foram feitas por Bragg, Laue e Ewald no início do século (Figura 19 e 20). O cristal é girado em torno de um eixo geralmente perpendicular ao feixe de raios X e os raios difratados são coletados em um detector, normalmente CCD (*charge-coupled device*), capaz de gravar um grande número de pontos (*spots*) em um tempo de aquisição bem curto. Cada imagem corresponde a uma rotação de $0,2-1^\circ$ grau e pode ser considerada uma "fatia" do espaço recíproco (Figura 20). A posição de uma reflexão contém as informações sobre a geometria do grupo espacial e da rede cristalina, enquanto que a intensidade I medida depende da distribuição da densidade eletrônica na unidade assimétrica.

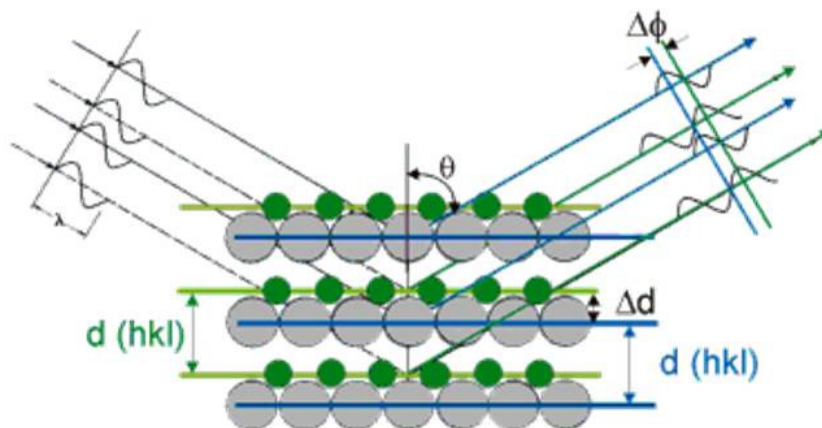


Figura 19. Visão geral do modelo de Bragg, quando se trata de seqüências de planos igualmente espaçados, mas composto por diferentes tipos de átomos, separados por uma distância Δd . Essa separação geométrica cria diferenças de fase dentro de um mesmo feixe difratado que provoca interferências e dão lugar a variações de intensidade (dependendo da direção), o que permite obter informações sobre a estrutura dos átomos que formam o cristal (disponível em <http://www.xtal.iqfr.csic.es/Cristalografia/>).

Cristais de macromoléculas são bastante sensíveis a raios X. Muitas vezes é necessário utilizar vários cristais para se obter um conjunto de dados que seja suficiente para prosseguir com os próximos passos. A coleta deve ser feita preferencialmente com o cristal resfriado instantaneamente num fluxo de nitrogênio gasoso a aproximadamente 100 K (criocristalografia). Isto permite a redução dos efeitos danosos da radiação direta e aumenta o tempo de vida do cristal. A distância do detector ao cristal deve ser regulada de forma a chegar o mais próximo possível da melhor resolução, limitada pela qualidade do cristal.

De posse das primeiras imagens, é possível realizar a indexação e determinar o grupo espacial, dimensões e orientação da célula unitária, além de estimar a mosaicidade. Esse passo é necessário para planejar uma estratégia de coleta de dados eficiente, minimizando o tempo e maximizando os resultados.

Na indexação, ocorre a identificação dos pontos e a determinação do índice de Miller de cada ponto. As imagens são analisadas utilizando programas como o MOSFLM (LESLIE, 1992), que nos permitem dar a cada reflexão seu índices de Miller (h, k, l) e sua intensidade I associado a um erro experimental σI . Todos as reflexões são então integradas pelo programa, as intensidades I são colocadas na mesma escala e convertidas a fatores de estrutura observadas F_{obs} .

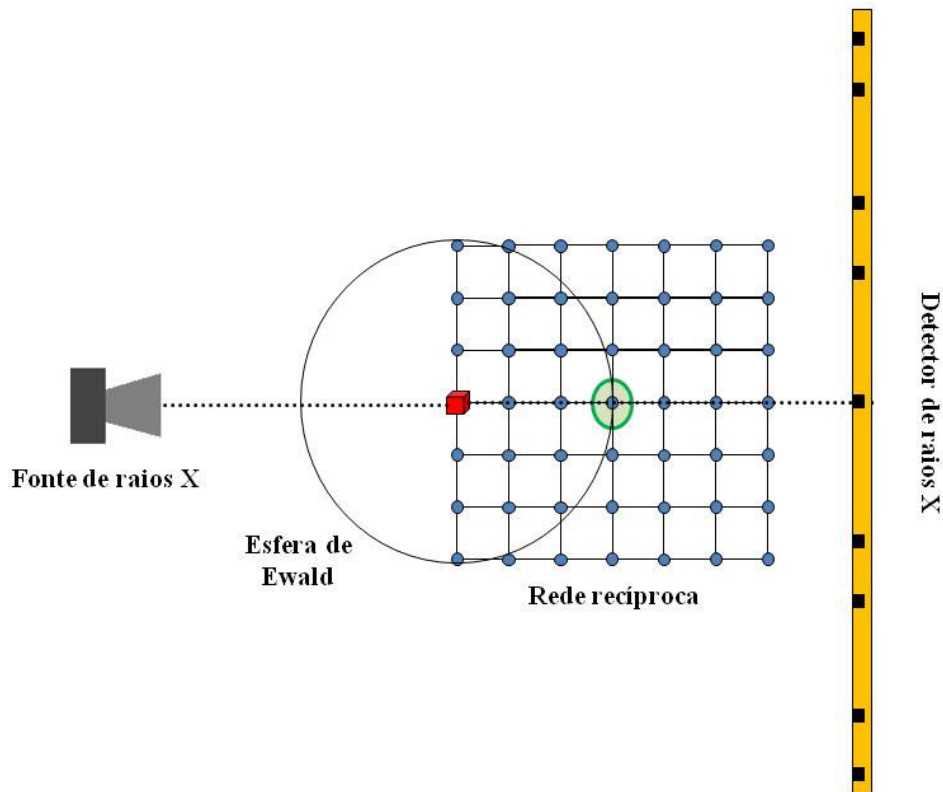


Figura 20. Esfera de Ewald: rede recíproca e a geometria de difração. Num experimento real, o cristal é girado e, portanto, a rede recíproca também gira. Apenas os pontos na rede recíproca que coincidem com a esfera de Ewald são detectáveis na tela. Para coletar a maioria das reflexões é preciso girar o cristal. Nem todas as reflexões precisam ser detectadas devido à geometria de espalhamento em combinação com os parâmetros de rede do cristal.

Devido à presença de operações de simetria no cristal, algumas reflexões parciais serão conectadas por simetria e, assim, terão a mesma intensidade de difração. A concordância entre elas é dada pelo R_{merge} . As informações necessárias sobre a distribuição de densidade eletrônica no cristal estarão contidas na verdade em um número reduzido de reflexões, chamadas únicas. A **completeza** é a porcentagem de reflexões únicas mensuradas. A relação entre as reflexões totais e únicas dão a **multiplicidade**. Um terceiro parâmetro é a razão entre a intensidade de uma reflexão e seu erro, $I/\sigma I$. Essas três grandezas, com a máxima resolução obtida, são os parâmetros mais importantes para avaliar a qualidade de um conjunto de dados experimentais processado.

3. Refinamento e Determinação da Estrutura

O fator de estrutura é caracterizada pela sua amplitude e fase. Experimentalmente, a frequência é conhecida, a amplitude está disponível, é proporcional a $(I)^{1/2}$ (raiz quadrada da medida da intensidade de reflexão), mas a fase é desconhecida.

O conceito central de toda a teoria da difração é que o fator de estrutura é a transformada de Fourier da distribuição de densidade eletrônica no cristal. Pelas propriedades de uma transformada de Fourier, o oposto também é verdadeiro, ou seja, a densidade eletrônica é a transformada de Fourier do fator de estrutura, o que significa que pode-se teoricamente converter dados cristalográficos e fazer a reconstrução do conteúdo espacial da rede cristalina. O fator de estrutura é caracterizado pela sua amplitude e fase. Experimentalmente, a frequência é conhecida, a amplitude está disponível, é proporcional a $(I)^{1/2}$ (raiz quadrada da medida da intensidade de reflexão), mas a fase é desconhecida. Para calcular a densidade eletrônica (ρ) em um ponto do cristal xyz , deve-se somar a contribuição de todas as ondas difratadas pelos planos hkl (cujas amplitudes dependerá do número de elétrons em cada plano), adicionadas da sua relação de fase:

$$\rho(xyz) = \frac{1}{V} \sum_{hkl} |F_{hkl}| \exp(i\alpha_{hkl}) \exp[-2\pi i(hx + ky + lz)]$$

Durante uma coleta de dados de difração, pode-se medir a amplitude de cada reflexão F_{hkl} , mas não a sua fase correspondente α_{hkl} , que é invariavelmente perdida durante o experimento de difração. Na verdade ainda não há nenhuma maneira para recombinar a intensidade e a fase de um difratado de raios X. Diferentes métodos para resolver o problema da fase foram então desenvolvidos.

Existem basicamente três métodos utilizados na resolução da estrutura para solucionar o problema das fases no caso das proteínas. O método da substituição molecular, da substituição isomórfica e da dispersão anômala.

A técnica mais utilizada para resolver o problema da fase é a substituição molecular (*Molecular Replacement* - MR). Esse método é baseado no princípio de que macromoléculas apresentam alta identidade sequencial, tendendo a se enovelar de forma bastante similar, e desta

forma, uma proteína cuja estrutura já tenha sido determinada, serve como um modelo inicial para a obtenção de um conjunto preliminar de fases que pode ser subsequentemente refinado. O sucesso deste método está no fato de hoje em dia existir um grande número de macromoléculas que já tem sua estrutura conhecida, mas o mesmo está condicionado à disponibilidade de uma estrutura (cristalina, RMN ou modelagem) com alta identidade. Em termos de similaridade de sequência entre as duas proteínas, a aplicação da MR tem chance de sucesso se a identidade é, de pelo menos, 25-30%, podendo haver exceções. Em geral, a estrutura utilizada como modelo é primeiro encaixada na nova rede com uma função de rotação (método de Patterson), e uma vez que as orientações favoráveis forem encontradas, são sujeitas a uma função de translação, tentando pôr a molécula corretamente orientada na rede. Os softwares disponíveis atualmente, como MOLREP (CCP4) e PHASER (MCCOY *et al.*, 2005) usam funções de pesquisa automáticas, geralmente requerindo uma intervenção mínima do usuário.

Uma vez que a estrutura foi posta em fase e que um primeiro modelo está disponível, o próximo passo é o refinamento das posições atômicas. Refinamento é o processo de encontrar uma melhor concordância entre o modelo proposto para a estrutura da molécula e a sua estrutura real. Esta concordância deve refletir numa igualdade entre os fatores de estrutura calculados (F_{calc}) e os fatores de estrutura observados (F_{obs}). O acompanhamento da qualidade deste processo é feito por meio do cálculo do fator R (R_{factor}) durante o processo [somatório das diferenças absolutas entre os fatores de estrutura observados (F_{obs}) e os fatores de estrutura calculados (F_{cal}) sobre a somatória dos fatores observados (MCREE, 1999)]. O objetivo é minimizar as diferenças entre o F_{calc} e o F_{obs} , tendo em conta as restrições estruturais e geométricas (comprimentos, ângulos de ligação, torções, etc.) que são inerentes à natureza molecular do modelo.

Para estruturas cristalográficas de proteínas resolvidas à resolução menor de 2,0 Å, espera-se que o R_{factor} atinja valores menores que 20%. Como o R_{factor} pode ser minimizado artificialmente, e levando em consideração a redundância dos dados de raios X, uma pequena porcentagem das reflexões são excluídas do refinamento e utilizadas como um conjunto de teste no cálculo de um novo fator R chamado R_{free} , que é isento de minimização artificial. Normalmente o R_{free} possui um valor um pouco mais alto que o R_{factor} , excedendo-o em até 5% para estruturas refinadas a alta resolução. Na prática, os erros macroscópicos do modelo são

sempre corrigidos "à mão" pelo cristalógrafo, que dispõe para isso de softwares de modelagem molecular, como o COOT (EMSLEY & COWTAN 2004).

Na etapa final do refinamento é feita a inserção das moléculas de água, usando o programa COOT (EMSLEY & COWTAN, 2004). Para isso é gerado um mapa do tipo $F_o - F_c$, que mostra as densidades não explicadas do mapa, que podem ser provenientes de moléculas de água ou de ruídos. Com o objetivo de verificar a veracidade das moléculas de água inseridas, é feito um refinamento do fator de vibração térmica (*b_factor*) (DELATORRE & DE AZEVEDO, 2001), onde todas as moléculas que apresentaram *b_factor* acima de 50,0 são excluídas, pois provavelmente trata-se de picos de ruídos. Do mapa de densidade eletrônica final ($2F_o - F_c$) é verificada a concordância com o modelo, o número de resíduos de aminoácidos que apresentam quebras na densidade eletrônica da cadeia principal, as falhas nas densidades eletrônicas das cadeias laterais, o N-terminal e o C-terminal das cadeias.

A validação do modelo refinado é efetuada com softwares como o PROCHECK (LASKOWSKI *et al.*, 1993), que analisam todos os parâmetros estruturais. Esse programa afere as distâncias ideais entre as ligações, os ângulos de ligação, os ângulos diédricos e os contatos de Van der Waals, usando para isso a comparação com proteínas modelos. Uma das saídas deste programa é o gráfico de Ramachandran, que afere as medidas dos ângulos diédricos Φ e Ψ sendo, portanto, um bom indicador da qualidade do modelo.

4. Resolução da estrutura cristalina de PELa em complexo com oligossacarídeos

4.1. Cristalização

Baseado nesses princípios, a PELa foi submetida a dois *screening* de co-cristalização: um primeiro com o trissacarídeo (α Man1-3(α Man1-6)Man e um segundo com o heptassacarídeo simétrico $\text{GlcNAc}\beta$ 1-2Man α 1-3($\text{GlcNAc}\beta$ 1-2Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAc β -Asn. Para ambos, seguiu-se essa estratégia.

A proteína purificada foi concentrada a 10 mg/mL em Tris-HCl 50 mM pH 7,5, NaCl 0,15 M, MnCl_2 2 mM, CaCl_2 2 mM e o oligossacarídeo foi adicionado a uma concentração final de 1 mM e deixado incubando em temperatura ambiente por 1 hora. A melhor condição (pH, sal e precipitantes) de formação do cristal da proteína foi obtida através da realização de um "screen" de cristalização, que consiste no método da matriz esparsa inicialmente descrito por JANCARIK

& KIM (1991). Para esta etapa foram utilizados os kits *Crystal Screen* I e II da Hampton. O método de cristalização utilizado foi o de difusão de vapor, gota suspensa, utilizando placas de cristalização de 24 poços. Foram colocados em cada poço da placa de cristalização 500 μL da condição do poço e a gota foi composta por 1 μL da solução de proteína e 1 μL da condição do poço. Os poços foram vedados com silicone e deixados em repouso à temperatura de 20°C.

Após a obtenção de cristais numa determinada solução pertencente ao kit do screen inicial, otimizações dessas condições de cristalização puderam ser feitas, variando a concentração de precipitante e o pH da solução, repetindo o método da difusão de vapor. Essa otimização visou melhorar a condição de formação do cristal, produzindo assim, um cristal com características necessárias para que ele possa ser mais bem difratado quando submetido a raios-x.

4.2. Coleta de dados em difração de raios X

A partir dos cristais obtidos, dois experimentos de difração de raios x foram realizados utilizando-se co-cristais da lectina recombinante com o ligante trimanose, crescidos na condição 42 do *Crystal Screen* II Kit Hampton (Tris 0,1 M pH 8,5, sulfato de amônio 1,5 M, glicerol 12% v/v); e com o ligante heptassacarídeo, crescidos na condição 48 do *Crystal Screen* II Kit Hampton (Bicina 100 mM pH 9, PEG 20 K 10%, dioxano 2%).

Os cristais com trimanose foram inicialmente congelados em nitrogênio líquido (100 K), após imersão rápida na solução crioprotetora com concentração final de glicerol 25% (v/v) em solução precipitante. Os cristais foram transportados ao ESRF (European Synchrotron Radiation Facility, Grenoble, França) e postos em linha para serem difratados na estação ID14-1, usando um detector ADSC Q4R CCD (charge-coupled device). Um conjunto de dados de difração foi coletado a 2,1 Å de resolução (1 Å = 0,1 nm), no grupo espacial $P2_12_12_1$ ($a= 42,62$, $b= 80,68$, $c= 148,76$ Å, $\alpha=\beta=\gamma= 90,0^\circ$). Foram coletadas 140 imagens, com a placa de imagens a uma distância de 251 mm, aplicando uma dose com tempo de exposição de 1,5 s e com oscilação de 0,75°. Dados de difração foram indexados, integrados e escalonados usando o programa MOSFLM (LESLIE, 1992) e SCALA (EVANS, 2006), e convertidos em fatores de estrutura usando o conjunto de programas CCP4 (COLLABORATIVE COMPUTATIONAL PROJECT, 1994).

Os cristais com heptassacarídeo foram inicialmente congelados em nitrogênio líquido (100 K) após imersão rápida na solução crioprotetora com concentração final de glicerol 30% (v/v) em

solução precipitante. Os cristais foram transportados ao ESRF (European Synchrotron Radiation Facility, Grenoble, França) e postos em linha para serem difratados na estação BM30A, usando um detector ADSC Q4R CCD (charge-coupled device). Um segundo conjunto de dados de difração foi coletado a 1,6 Å de resolução, no grupo espacial $P2_12_12_1$ ($a= 51,40$, $b= 76,98$, $c= 125,59$ Å, $\alpha=\beta=\gamma= 90^\circ$). Foram coletadas 91 imagens, com a placa de imagens a uma distância de 256,2 mm, aplicando uma dose com tempo de exposição de 10 s e com oscilação de 1° . Dados de difração foram indexados, integrados e escalonados usando o programa MOSFLM (LESLIE, 1992) e SCALA (EVANS, 1993), e convertidos em fatores de estrutura usando o conjunto de programas CCP4 (COLLABORATIVE COMPUTATIONAL PROJECT, 1994).

4.3. Substituição Molecular

Para o primeiro conjunto de dados (PELa_trim), inicialmente foi feito um alinhamento para análise de similaridade e identidade entre a sequência primária da lectina recombinante com o banco de dados de proteínas depositadas no *Nacional Center of Biotechnology Information* (NCBI), utilizando a ferramenta BLAST. Do resultado desta análise, foi escolhida como a proteína com maior identidade e similaridade a lectina de *Pterocarpus angolensis* (PDB : 1N3O). A partir de suas coordenadas como modelo, foi feita a substituição molecular, que consistiu em calcular o posicionamento do modelo com a molécula real através de movimentos de rotação e translação. Esse cálculo foi feito pelo programa PHASER (MCCOY *et al.*, 2005) e a escolha do modelo ideal foi feita pela análise dos valores de *log-likelihood gain*, RFZ e TFZ fornecidos pelo programa. Para o segundo conjunto de dados (PELa_hept), utilizou-se como modelos as coordenadas da estrutura de PELa em complexo com trimanose (PDB: 3ZVX).

4.4. Refinamento e Validação

Após a substituição molecular, a estrutura obtida foi submetida a refinamentos de corpo rígido e refinamentos posicionais para um melhor ajuste da molécula na unidade assimétrica através do uso do programa REFMAC5 (MURSHUDOV *et al.*, 1997), parte integrante do pacote CCP4. A finalidade deste refinamento foi encontrar uma melhor concordância entre o modelo

proposto e a estrutura real da molécula, colocando o modelo na mesma posição que a molécula real ocupa na sua rede cristalina. Para avaliar os resultados dos cálculos do refinamento foram observados os valores de R_{factor} e R_{free} . O fator da estrutura foi calculado de acordo com o espalhamento de onda, em uma determinada direção e intensidade, no espaço recíproco definido pelos índices de Miller. Na obtenção de um modelo após substituição molecular com R_{factor} inicial de 40% um bom refinamento possui valores de R_{factor} em torno de 20% e R_{free} variando até em 5% desse mesmo valor.

Posteriormente foram adicionadas as moléculas de água utilizando-se o programa REFMAC, seguido de um novo refinamento, e cada estrutura foi submetida à validação pelo programa MolProbity (DAVIS *et al.*, 2007), que avalia os ângulos, a estereoquímica e as distâncias interatômicas. A avaliação da posição dos aminoácidos em relação às regiões permissíveis foi feita pelo gráfico de Ramachandran. As coordenadas foram depositadas no *Protein Data Bank* nos códigos 3ZVX e 3ZYR. Todas as figuras foram feitas utilizando os programas PyMol (DELANO, 2002) e LigPlot (WALLACE *et al.*, 1995).

Capítulo 3:

Clonagem e expressão heteróloga de uma lectina de sementes de *Platypodium elegans* (PELa)

Esse capítulo reúne os principais resultados concernente a metodologia desenvolvida para a clonagem e descrição do gene de PELA, sua produção heteróloga no sistema *E. coli*, assim como a obtenção da proteína recombinante purificada e ativa.

1. Extração e análise do RNA total de sementes de *P. elegans*

A extração do RNA é o passo fundamental para análises moleculares do transcriptoma, tanto por meio de bibliotecas de cDNA, RT-PCR como por RTqPCR. É imprescindível o estabelecimento de um protocolo de forma a obter um material com qualidade e em quantidade necessárias para esses estudos. A principal preocupação na extração de RNA consistiu na prevenção de sua degradação por ribonucleases (RNAses), que são enzimas extremamente resistentes a vários tratamentos, inclusive térmicos.

O RNA total extraído de um grama de sementes maduras de *P. elegans* pelo kit *Plant RNA Reagent* (Invitrogen) apresentou um rendimento de 238,6 ng/mL, com um perfil eletroforético mostrado na figura 21, apresentando um pequeno grau de degradação.

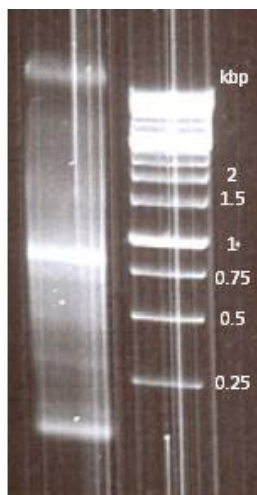


Figura 21. Eletroforese em gel de agarose (1%). Extração do RNA total de sementes de *P. elegans*. 1. RNA total. 2. Marcadores moleculares (kbp).

Apesar da qualidade não ter sido excelente, a partir dessa extração foi possível, por transcrição reversa, amplificar regiões específicas cDNA utilizado em ambas as estratégias RACE.

A tabela 7 reúne todos os iniciadores (*primers*) utilizados nesse estudo:

Tabela 7. Lista dos iniciadores (*primers*) utilizados nesse estudo.

Iniciador (<i>primer</i>)	Tipo	5'- Sequência - 3'*	Sequência (aa) Correspondente	Passo
PLAT1	senso	TN GCN GTN GAR TTY GAY AC	(V/L/I)AVEFDT	3'-RACE
PLAT2	senso	CAY ATH GGN ATH GAY GTN AA	HIGIDVN	3'-RACE
PLAT3	antisenso	TGATATAGAGATCCCTCTGGA	SRGISIS	5'-RACE
AP	antisenso	GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTT	-	3'-RACE
AAP	senso	GGCCACGCGTCGACTAGTACGGGIIIGGGIIGGGIIG	-	5'-RACE
PLAT4 (<i>NdeI</i>)	senso	GGAATTCC <u>CATATGCT</u> ACTGAATAGGGCAAAC	MLLNRA	clonagem pET-29a
PLAT5 (<i>HindIII</i>)	antisenso	GGCA <u>AAGCTT</u> TCACATATCACGTGCAAGATAC	LARDM Stop	clonagem pET-29a

*R, Y, e N indicam (A ou G), (C ou T), e (A, C, G ou T), respectivamente. I, ácido inosídico. As sequências sublinhadas indicam as posições dos sítios para *NdeI* ou *HindIII*.

2. Estrutura, caracterização e clonagem do mRNA de PELa

2.1. Produto 3'RACE

Dois experimentos 3' RACE foram realizados. A reação de RT-PCR usando um dos oligonucleotídios degenerados (PLAT1 ou PLAT2) para a sequência codificadora da lectina em associação com o oligodT (*Adapter prime*) resultou na amplificação do cDNA em diferentes comprimentos. Bandas eletroforéticas de aproximadamente 500 e 550 pb foram observadas para as amplificações feitas na presença de PLAT1 e PLAT2, respectivamente (Figura 22). O resultado foi coerente com o esperado, pois a sequência codificadora deveria possuir em torno de 400 e 350 pb (referente às posições de anelamento de PLAT1 e PLAT2, respectivamente), fora as demais estruturas esperadas em extremidades 3' de mRNA (região 3' não codificadora e cauda poliA).

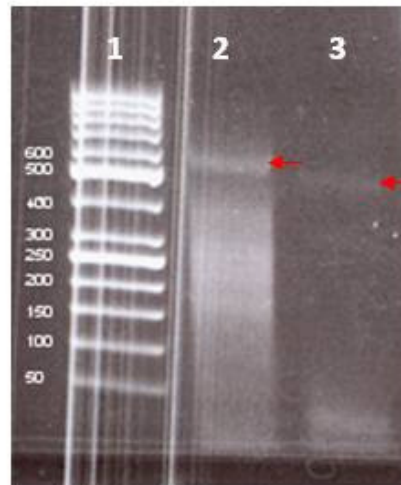


Figura 22. Eletroforese em gel de Agarose (1,2%). Produtos 3'RACE. 1. Marcadores moleculares (pb). 2. Amplificação na presença de PLAT1. 3. Amplificação na presença de PLAT2.

Após purificação da banda eletroforética e reamplificação, somente o amplificado por PLAT2 obteve um bom perfil eletroforético (Figura 23), sendo submetido à clonagem em pGEM-t *easy* e posterior sequenciamento.

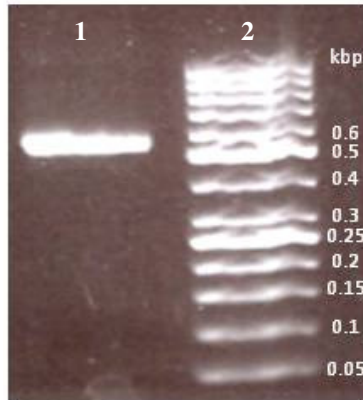


Figura 23. Eletroforese em gel de Agarose (1,2%). 1. Produto 3' RACE purificado e reamplificado. 2. Marcadores moleculares (kpb).

O sequenciamento do produto 3' RACE amplificado pelos oligonucleotídios PLAT2 e oligodT (AP) revelou a presença de uma região codificadora de 321 nucleotídeos terminada por um códon *stop*, uma região 3' não codificadora de 165 nucleotídios e uma cauda poliA com 27 nucleotídios em extensão. O início da sequência da região não codificadora foi utilizado para desenhar um oligonucleotídio específico (PLAT3) a ser usado na estratégia 5' RACE. A análise da sequência completa do mRNA codificador da lectina de *P. elegans* é feita nas figuras 25 e 26.

2.2. Produto 5' RACE

Utilizando-se os oligonucleotídeos PLAT3 e oligodG (AAP), o primeiro se anelando à região 3' correspondente do mRNA e o segundo à cauda poliC extendida na região 5' do cDNA, foi possível amplificar o produto 5' RACE. A análise em eletroforese de gel de agarose (1%) desse produto revelou uma banda em torno de 800 pb (Figura 24).

Após ligação em pGEM-T easy e sequenciamento, o produto 5' RACE revelou a estrutura restante do mRNA, cuja sequência codificadora apresentou 786 nucleotídeos, com um códon de iniciação (*start*) e uma sequência de 15 pb na região 5' não codificadora (localizada logo depois da sequência poliG complementar à cauda poliC extendida no cDNA pelo método). A análise da sequência completa do mRNA codificador da lectina de *P. elegans* é feita nas figuras 25 e 26.

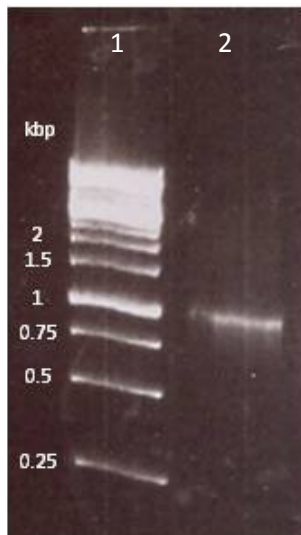


Figura 24. Eletroforese em gel de Agarose (1%). Produto 5' RACE. 1. Marcadores moleculares (kbp). 2. Produto 5' RACE purificado e reamplificado.

2.3. Estrutura e caracterização do mRNA de PELa

A figura 25 reúne a estrutura completa do mRNA codificador da lectina de *P. elegans*, deduzida à partir dos produtos gerados pelas estratégias 3' e 5' RACE: a região codificadora, com 786 nucleotídios, incluindo uma sequência ATG iniciadora (*start*) e uma sequência TGA de parada (*stop*), além das regiões 3' e 5' não codificadoras e da cauda poliA, característica de mRNA.

A partir da determinação desta sequência, foi possível traçar a estratégia de clonagem desse mRNA no vetor de expressão pET-29a, graças ao desenho de iniciadores (PLAT4 e PLAT5) que flanqueassem as extremidades da sequência, adicionada de sítios para enzimas de restrição próprios para a estratégia (*NdeI* e *HindIII*). A sequência nucleotídica codificadora completa está representada na figura 26.

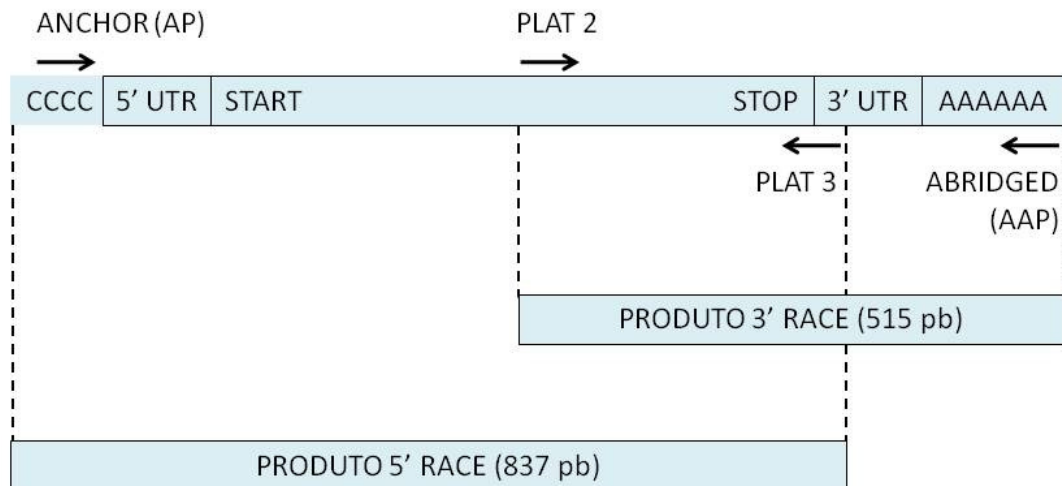


Figura 25. Estrutura completa do mRNA codificante para a lectina de *P. elegans* (região codificadora: 786 nucleotídios), deduzida à partir dos produtos 3' e 5' RACE. UTR: região terminal não codificadora.

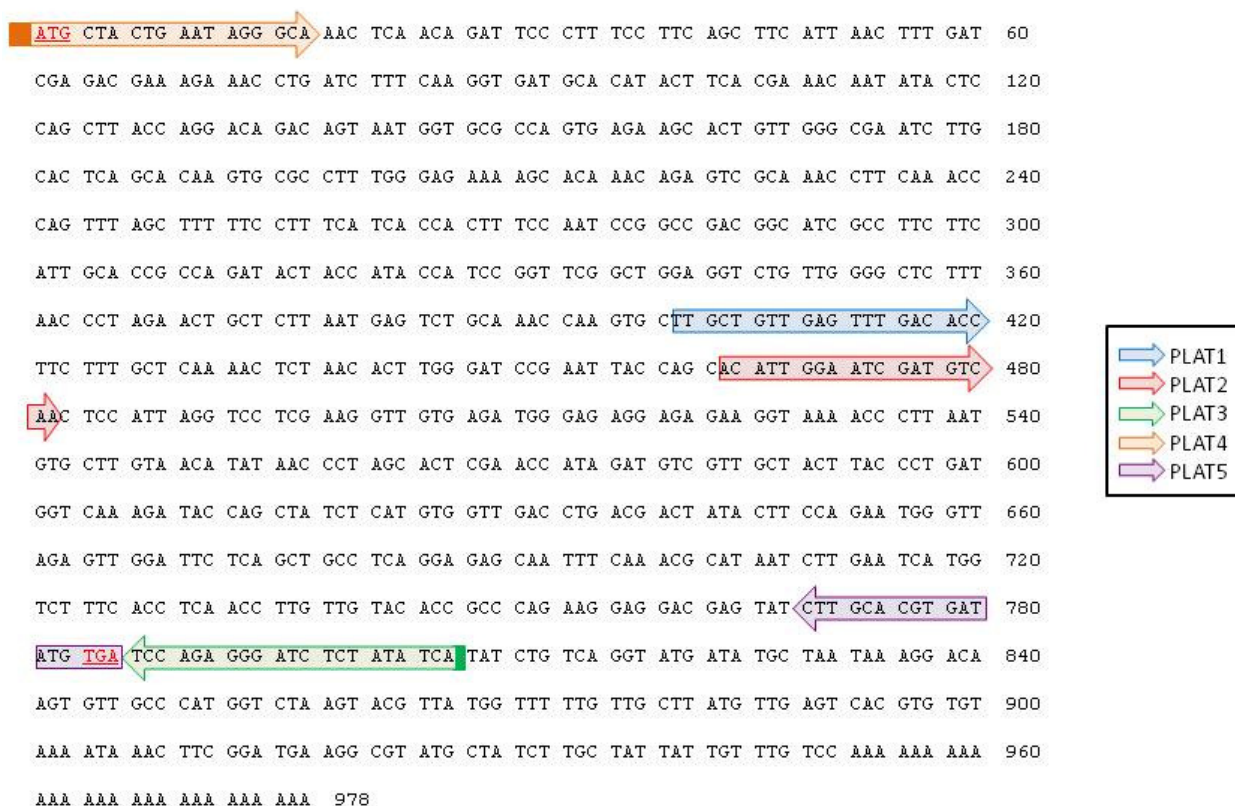


Figura 26. Sequência nucleotídica codificadora do mRNA da lectina de *P. elegans*. A sequência correspondente ao anelamento dos iniciadores PLAT 1, PLAT2, PLAT 3, PLAT4 e PLAT5 estão exibidos nas setas azul, vermelha, verde, laranja e lilás, respectivamente. PLAT4 e PLAT5 contêm sítios de restrição a endonucleases (*NdeI* and *HindIII*, respectivamente). Códon de iniciação (*start*) e de parada (*stop*) evidenciados em vermelho e sublinhados.

3. Determinação da sequência protéica e Alinhamento com lectinas de leguminosas

A partir da sequência de aminoácidos de uma proteína, podem-se fazer predições a cerca de modificações pós-traducionais na sequência protéica, estruturas secundárias, terciárias e até mesmo quaternárias, em interações com outras cadeias. Os resultados gerados guiam o trabalho experimental e fornecem subsídios para perguntas biológicas pertinentes à proteína analisada.

A partir do software *ProtParam* (GASTEIGE *et al.*, 2005), que computa várias propriedades físico-químicas deduzíveis da sequência protéica, obteve-se valores úteis como o pI teórico (ponto isoelétrico), o coeficiente de extinção (soma dos coeficientes de extinção molar dos resíduos de Tirosina, Triptofano e Cisteína), e a massa molecular.

A sequência protéica da lectina de *P. elegans*, deduzida a partir da sequência nucleotídica de seu mRNA, com 261 aminoácidos, massa molecular de 29287,5 Da, pI teórico de 5,73 e coeficiente de extinção molar de $36440 \text{ M}^{-1} \cdot \text{cm}^{-1}$ (280 nm, abs 0,1% = 1,244) está apresentada na figura 27, sendo evidenciado prováveis sítios de *N*-glicosilação, preditos pela sequência.

A *N*-glicosilação consiste na fixação de uma molécula de açúcar no grupo $-\text{NH}_2$ da cadeia lateral de um resíduo de Asn em uma das sequências consenso seguintes: Asn-X-Ser e Asn-X-Thr (e em alguns casos raros, Asn-X-Cys), onde X pode ser qualquer aminoácido, exceto prolina (BAUSE & LEGLER, 1981). Essas sequências são conhecidas como tripeptídeo *sequon*.

Por muitos anos, pensava-se que a *N*-glicosilação estava presente apenas em eucariotas. Mas hoje, proteínas *N*-glicosiladas em procariotos é um fato (SCHAFFER *et al.*, 2001). Em eucariotos, a *N*-glicosilação protéica ocorre no retículo endoplasmático (RE) (HELENIUS & AEBI, 2001). Os três prováveis sítios marcados na sequência de PELa, nas posições 7, 127 e 186, podem apresentar-se glicosilados na lectina sendo expressa em sua fonte natural, uma vez que se encontrem na superfície da lectina. O atual sistema de expressão utilizado, células de *E. coli* BL21, não apresenta aparato para realizar modificações pós-traducionais como a *N*-glicosilação.

Em relação à presença de uma sequência peptídico sinal, após uma análise de previsão pelo software SinalP 4.0 (<http://www.cbs.dtu.dk/services/SignalP/>), não foi prevista nenhuma sequência característica (NIELSEN *et al.*, 1997). Comparando com sequências de outras lectinas de leguminosas, é possível identificar sequências relacionadas ao peptídeo sinal em análises de

genes de PHA, PNA e PAL envolvendo um sítio para clivagem (VOELKER *et al.*, 1986; VAN EIJSDEN *et al.*, 1992; LORIS *et al.*, 2004).

```

ATG CTA CTG AAT AGG GCA AAC TCA ACA GAT TCC CTT TCC TTC AGC TTC ATT AAC TTT GAT 60
M L L N R A N S T D S L S F S F I N F D 20
CGA GAC GAA AGA AAC CTG ATC TTT CAA GGT GAT GCA CAT ACT TCA CGA AAC AAT ATA CTC 120
R D E R N L I F Q G D A H T S R N N I L 40
CAG CTT ACC AGG ACA GAC AGT AAT GGT GCG CCA GTG AGA AGC ACT GTT GGG CGA ATC TTG 180
Q L T R T D S N G A P V R S T V G R I L 60
CAC TCA GCA CAA GTG CGC CTT TGG GAG AAA AGC ACA AAC AGA GTC GCA AAC CTT CAA ACC 240
H S A Q V R L W E K S T N R V A N L Q T 80
CAG TTT AGC TTT TTC CTT TCA TCA CCA CTT TCC AAT CCG GCC GAC GGC ATC GCC TTC TTC 300
Q F S F F L S S P L S N P A D G I A F F 100
ATT GCA CCG CCA GAT ACT ACC ATA CCA TCC GGT TCG GCT GGA GGT CTG TTG GGG CTC TTT 360
I A P P D T T I P S G S A G G L L G L F 120
AAC CCT AGA ACT GCT CTT AAT GAG TCT GCA AAC CAA GTG CTT GCT GTT GAG TTT GAC ACC 420
N P R T A L N E S A N Q V L A V E F D T 140
TTC TTT GCT CAA AAC TCT AAC ACT TGG GAT CCG AAT TAC CAG CAC ATT GGA ATC GAT GTC 480
F F A Q N S N T W D P N Y Q H I G I D V 160
AAC TCC ATT AGG TCC TCG AAG GTT GTG AGA TGG GAG AGG AGA GAA GGT AAA ACC CTT AAT 540
N S I R S S K V V R W E R R E G K T L N 180
GTG CTT GTA ACA TAT AAC CCT AGC ACT CGA ACC ATA GAT GTC GTT GCT ACT TAC CCT GAT 600
V L V T Y N P S T R T I D V V A T Y P D 200
GGT CAA AGA TAC CAG CTA TCT CAT GTG GTT GAC CTG ACG ACT ATA CTT CCA GAA TGG GTT 660
G Q R Y Q L S H V V D L T T I L P E W V 220
AGA GTT GGA TTC TCA GCT GCC TCA GGA GAG CAA TTT CAA ACG CAT AAT CTT GAA TCA TGG 720
R V G F S A A S G E Q F Q T H N L E S W 240
TCT TTC ACC TCA ACC TTG TTG TAC ACC GCC CAG AAG GAG GAC GAG TAT CTT GCA CGT GAT 780
S F T S T L L Y T A Q K E D E Y L A R D 260
ATG TGA 786
M * 261

```

Figura 27. Sequência polipeptídica da lectina de *P. elegans* deduzida a partir do mRNA correspondente. Prováveis sítios de *N*-glicosilação previstos pela sequência evidenciados em quadro vermelho.

A sequência protéica de PELa apresentou similaridades com outras lectinas de leguminosas. PELa teve o maior nível de identidade de sequência (82%) com a lectina Man/Glc específica de *Pterocarpus rotundifolium* (número de acesso: AAT57665), seguida das lectinas Man/Glc específicas de *Pterocarpus angolensis* (número de acesso: CAD19804), e de *Arachis hypogae* (número de acesso: AAA74572), com identidades de sequência de 76 e 70%, respectivamente. Após o alinhamento com lectinas de leguminosas relacionadas que apresentam

estruturas cristalinas determinadas (Figura 28), foi possível observar na sequência da lectina de *P. elegans* regiões bastante conservadas em relação ao sítio de ligação a metais e sítio de ligação a carboidratos, assim como as cinco voltas descritas como integrantes do sítio de ligação a carboidratos (LORIS *et al.*, 1998; SHARMA & SUROLIA, 1997).

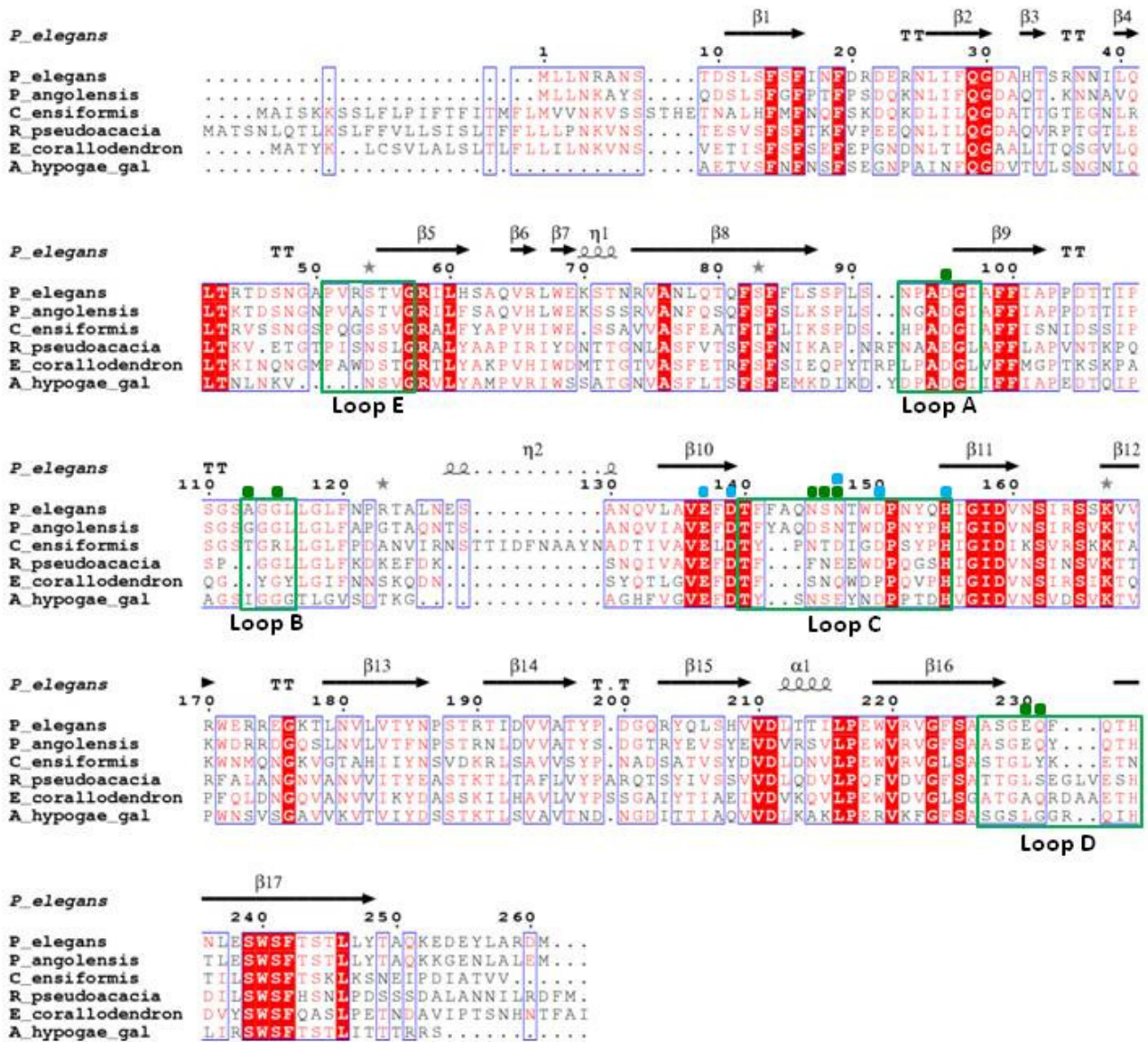


Figura 28. Alinhamento de seqüências polipeptídicas de lectinas de leguminosas. O alinhamento e a predição de estruturas secundárias foi adquirido utilizando o programa webESPrpt (GOUET *et al.*, 1999). Os dados de seqüência foram obtidos pelo Entrez (*cross-database search*, banco de dados para seqüências protéicas). Números ID: AEK69351 (*Platypodium elegans* lectin), CAD19804 (*Pterocarpus angolensis* lectin), CAA25787 (*Canavalia ensiformis* lectin), BAA36415 (*Robinia pseudoacacia* lectin) e CAA36986 (*Erythrina collarodendrom* lectin) e 2_PEL_A (*Arachis hypogaea* lectin). Pontos verdes evidenciam resíduos envolvidos no sítio de ligação a metais e pontos azuis os resíduos do sítio de ligação a carboidrato em PELA; em blocos verdes, as cinco voltas envolvidas no sítio de ligação a carboidrato de PELA a oligossacarídeo.

Ambas as sequências nucleotídica e polipeptídica da lectina de *P. elegans* foram depositadas no *Genbank*, podendo ser acessadas através dos códigos JN133278 e AEK69351, respectivamente.

4. Clonagem no vetor de expressão pET-29a e transformação em células de *E. coli* BL21

A amplificação do produto 5' RACE com os iniciadores PLAT4 e PLA5, após purificação, procedimento de clonagem em pET-29a e subsequente sequenciamento, teve sua sequência avaliada e confirmada como correta. O plasmídeo recombinante foi então utilizado para transformar células competentes de *E. coli* BL21, cujos clones positivos foram selecionados e submetidos à extração do plasmídeo para validação da transformação.

Durante o procedimento de validação da transformação em células de *E. coli* BL21, a digestão do plasmídeo com as enzimas *Nde*I e *Hind*III resultou na liberação de um fragmento específico ao tamanho esperado, em torno de 750 pb (Figura 29), confirmando a eficiência da transformação.

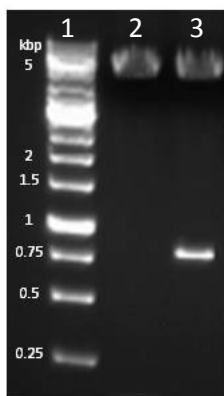


Figura 29. Validação da transformação de células de *E. coli* com o plasmídeo pET-29a_PELa. 1. Marcadores moleculares (kbp). 2. Controle negativo. 3. Plasmídeo após digestão com *Nde*I e *Hind*III.

Após a validação, clones positivos foram submetidos à indução da expressão da proteína.

5. Indução e Expressão de PELa

A figura 30 apresenta uma eletroforese em 12% das proteínas totais extraídas das células de *E. coli* BL21 transformadas com o plasmídeo pET-29a_PELa, após indução da expressão a 30°C com IPTG 0,5 mM por um período de 4 horas. O aumento da expressão protéica ao longo do período de indução é evidente, ocorrendo um aumento específico de uma banda em torno de 29 kDa, tamanho referente ao esperado, excluindo a possibilidade da mesma ter sofrido algum processamento proteolítico.

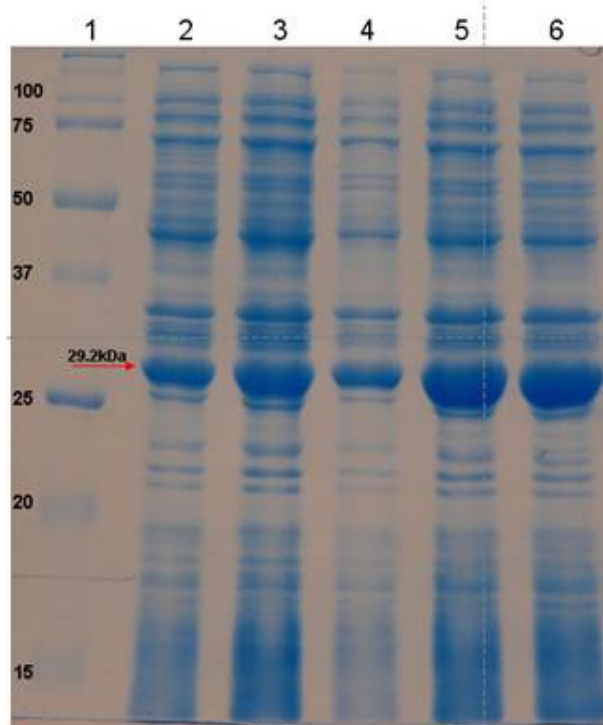


Figura 30. SDS-PAGE 12% de extrato de células de *E. coli* BL21 induzidas com IPTG 0,5 mM. 1. Marcadores moleculares (kDa). 2. Extrato celular após 0 hora de indução. 3. Após 1 hora. 4. Após 2 horas. 5. Após 3 horas. 6. Após 4 horas.

Após expressar a proteína, testamos a sua solubilidade. Após lise celular, a presença de PELa foi constatada na fração insolúvel dos extratos de bactérias induzidas por IPTG (Figura 31). É possível visualizar uma banda preponderante presente na fração insolúvel (corpos de inclusão), correspondente àquela presente em extrato celular induzido.

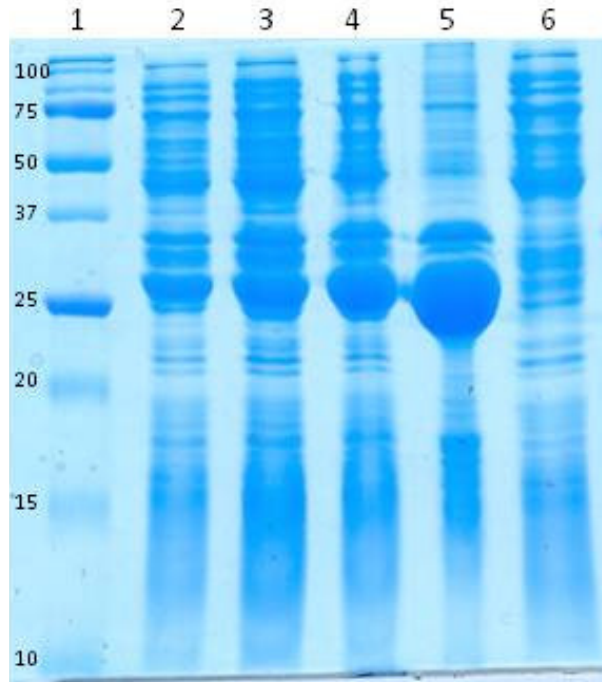


Figura 31. SDS-PAGE 12% do extrato celular de *E. coli* BL21 após indução e ruptura. 1. Marcadores moleculares (kDa). 2, 3 e 4. Extrato celular após 0, 1 e 4 horas de indução, respectivamente. 5. Fração celular insolúvel (corpos de inclusão). 6. Fração celular solúvel

Não se pode prever, com exatidão, quais proteínas serão expressas de forma solúvel em *E. coli*. A formação de corpos de inclusão pode ser influenciada por vários fatores, incluindo a natureza da proteína, a célula hospedeira, as condições de crescimento e indução e o nível de expressão resultante da escolha do vetor. Mesmo com outras temperaturas e concentrações de indutor testadas no experimento, não foi possível obter a proteína recombinante na fração solúvel do extrato celular. Seguiu-se, portanto, a estratégia de purificá-la a partir de corpos de inclusão, pós-*refolding* (renaturação).

6. *Refolding* e Purificação de PELa

O teste de solubilidade foi realizado para verificar se PELa estava presente na fração solúvel ou insolúvel do extrato das proteínas totais bacterianas. Com o tampão de lise utilizado, a proteína recombinante permaneceu no sedimento, no qual estavam as proteínas insolúveis. Este experimento permitiu definir que a purificação de PELa deveria ser sob condições desnaturantes.

Foram utilizados 2 L de cultura bacteriana contendo a construção pET-29a_PELa para a expressão da proteína com a finalidade de purificá-la. Como não houve alteração nos níveis de expressão da proteína recombinante após 4h, a indução com IPTG foi feita durante este período. Ao fim da indução, as células foram precipitadas e cada 1 grama de precipitado celular foi aliquotado individualmente (armazenagem a -20 °C).

Para a purificação dos corpos de inclusão, posteriormente, cada 1 grama de bactérias precipitadas foi lisado e submetido a duas centrifugações/lavagens sucessivas na presença de Tris-HCl 50 mM pH 8,0, NaCl 0,5 M, Uréia 4 M, , EDTA 1 mM, deoxicolato 1 mg/mL. Esse passo foi útil para remover restos celulares, proteínas (principalmente proteases) adsorvidas, contaminantes em geral, já que nessa condição ainda não havia proteína recombinante solubilizada.

Os corpos de inclusão foram em seguida solubilizados em Tris-HCl 50 mM pH 8,0, Uréia 8 M, DTT 10 mM e, após ajuste da concentração protéica para 1 mg/mL, submetido ao processo de renaturação (*refolding*). A renaturação foi obtida diluindo-se 20 vezes a preparação protéica de forma lenta e em agitação no tampão Tris-HCl 50 mM pH 7,5, NaCl 240 mM, MnCl₂ 2 mM, CaCl₂ 2 mM, KCl 10 mM, DTT 1 mM, incubando a suspensão formada a 10 °C. Após 1 hora, a fração renaturada encontrava-se solubilizada, enquanto que uma fração não corretamente renaturada insolúvel formava uma suspensão turva, a qual foi descartada após uma filtração simples.

A purificação da proteína recombinante renaturada foi feita através de uma coluna de afinidade manose-agarose (Sigma). O cromatograma obtido representa a passagem de toda a amostra renaturada pela coluna de afinidade, onde é possível visualizar a fração referente à eluição na presença de manose 200 mM e sua absorbância correspondente a 280 nm (Figura 32). Foi possível visualizar uma fração não retida à matriz cromatográfica, correspondente a provável proteína pós-renaturação ainda má enovelada, não apresentando sítio de ligação a carboidrato em conformação ativa.

Todas as frações obtidas foram analisadas em SDS-PAGE (Figura 33), juntamente com todos os passos do reenovelamento. A análise por SDS-PAGE indicou uma massa molecular em torno de 29 kDa, bem correspondente à massa molecular teórica calculada à partir da sequência primária. A partir de cada litro de cultura bacteriana foi possível obter cerca de 5 mg de proteína recombinante. As frações obtidas dessa purificação foram concentradas e/ou dessalinizadas para

as condições requeridas de cada análise posterior por ultrafiltração (Vivaspin, cutoff 10 kDa; GE Healthcare), e depois devidamente armazenadas a 4 °C.

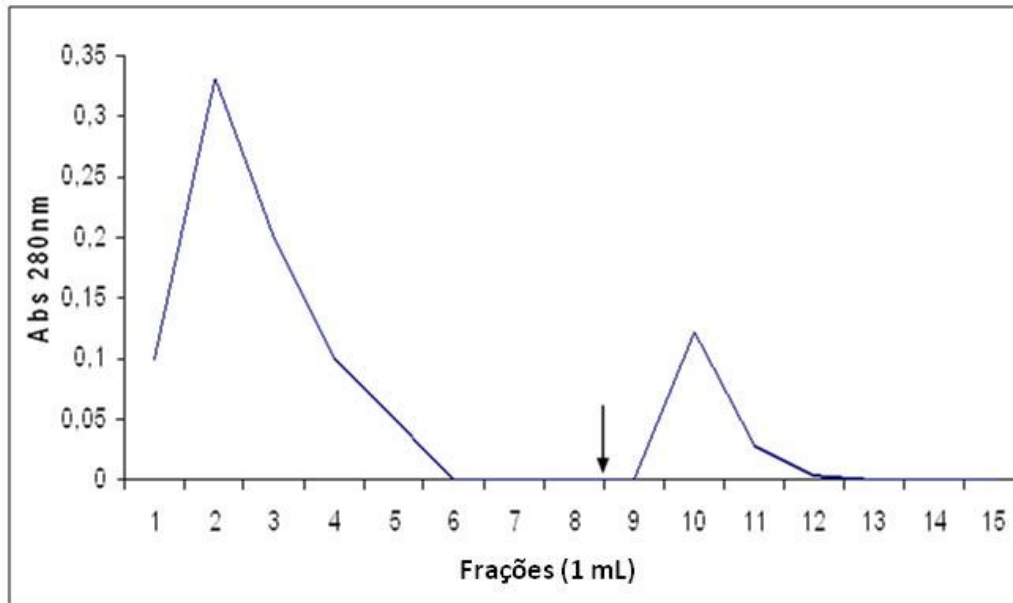


Figura 32. Cromatografia de afinidade em matriz de Manose-Agarose: Purificação de PELa pós-*refolding*. Proteína pós-*refolding* à partir de 1 grama de precipitado celular foi aplicado em 1 mL de gel pré-equilibrado com Tris -HCl 50 mM pH 7,5, NaCl 240 mM, MnCl₂ 2 mM, CaCl₂ 2 mM. Fração retida foi eluída com a adição de manose 200 mM, representada pela seta. Frações coletadas (1 mL) foram monitoradas por absorvância a 280 nm.

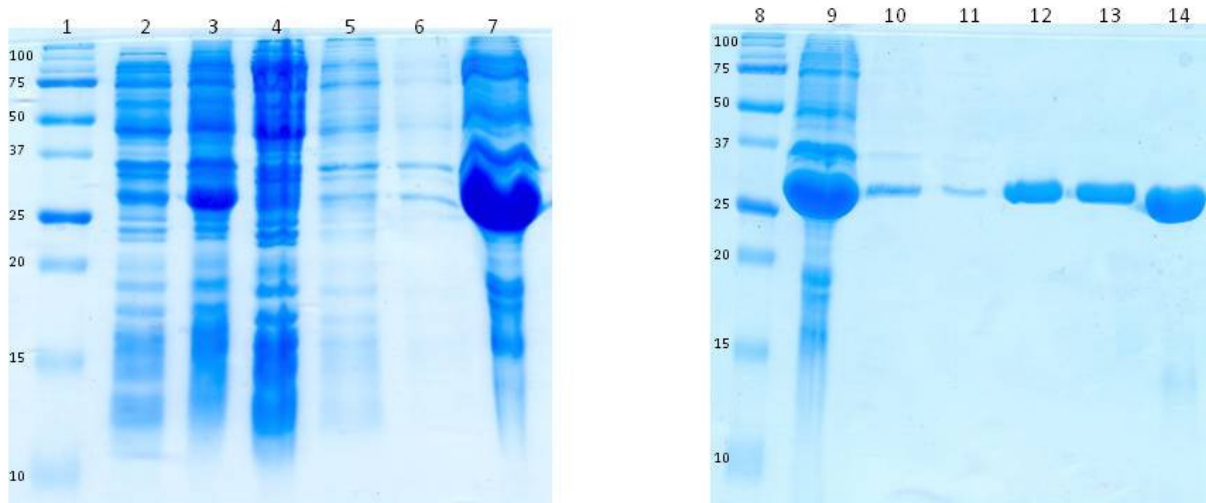


Figura 33. SDS-PAGE 12%: *refolding* e purificação de PELa. 1, 8. Marcadores moleculares (kDa). 2. Extrato celular antes da indução com IPTG. 3. Extrato celular após 4 horas de indução com IPTG. 4. Fração solúvel pós lise. 5, 6. Lavagens corpos de inclusão com uréia 4 M. 7, 9. Fração insolúvel (corpos de inclusão) solubilizada em uréia 8 M. 10. PELa pós-*refolding*. 11. Fração não retida à matriz de Manose-agarose. 12, 13. Frações retidas eluídas com manose 200 mM. 14. PELa purificada e concentrada a 3 mg/mL.

Capítulo 4:

Estudo das propriedades hemaglutinantes, especificidade e afinidade da lectina PELa

Os resultados agrupados nesse capítulo descrevem a funcionalidade, assim como a especificidade de PELa e sua afinidade frente a alguns oligossacarídeos, fazendo sempre que possível uma comparação com outras lectinas da mesma família analisadas de forma semelhante.

1. Propriedades hemaglutinantes de PELa

Para confirmar o correto enovelamento e funcionalidade de PELa obtida pós-reenovelamento e purificação, ensaios de hemaglutinação com uma suspensão de hemácias de coelho a 2% em solução salina foram realizados como um experimento rápido, de fácil visualização e conclusivo.

A figura 34 mostra as propriedades hemaglutinantes de PELa purificada e concentrada a partir de 1,5 mg/mL. A partir de diluições seriadas, foi possível calcular a quantidade de unidades de hemaglutinação em solução (UH/mL), em detrimento da concentração protéica. Considerou-se o equivalente a 1 UH/mL a última diluição que ainda apresentava hemaglutinação nítida macroscopicamente ($2^7 = 128$ UH/mL), em relação ao controle negativo (-). A proteína recombinante apresentou, portanto, 1 UH/mL para 23 µg/mL de proteína (atividade específica de 42 UH/mgP).

Os resultados mostram que PELa possui propriedades hemaglutinantes, confirmando o enovelamento correto e o estado funcional da proteína.

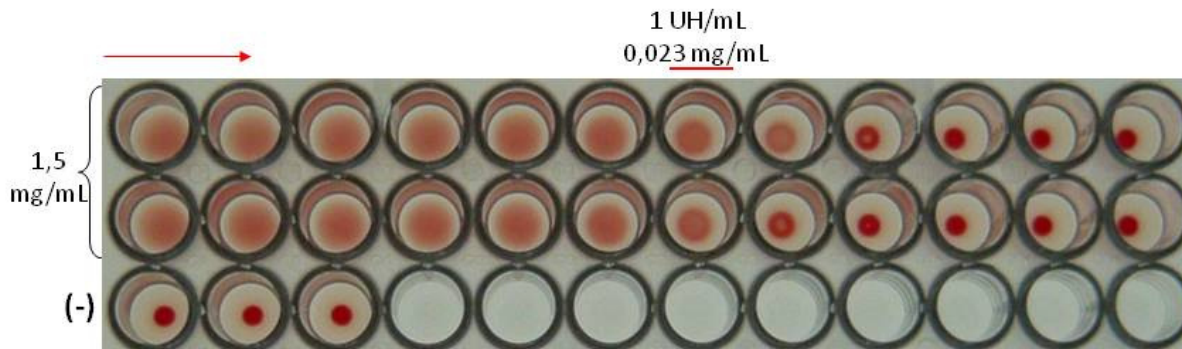


Figura 34. Hemaglutinação de hemácias de coelho por PELA. (-) Controle negativo, sem lectina.

A proteína recombinante mostrou-se ativa somente após a purificação em manose-agarose. Com esse procedimento, foi possível confirmar o estado ativo de todas as preparações provenientes de reenovelamento, purificação e ultrafiltração, para serem utilizadas e analisadas nos demais experimentos.

2. Estudos de especificidade e afinidade na interação lectina-carboidrato em PELA

Lectinas de vegetais apresentam a capacidade de reconhecer especificamente mono e oligossacarídeos inexistente em seus vegetais de origem, mas abundantes em outros organismos, o que juntamente com a resistência destacada à proteólise e estabilidade mediante variações consideráveis de temperatura e pH, reforçam a idéia do envolvimento de tais proteínas na função de defesa do vegetal. Portanto, a especificidade encontrada em lectinas vegetais, encontra-se direcionada para carboidratos presentes em potenciais patógenos, como bactérias, vírus e fungos, e parasitas, como insetos, animais, etc., cujas membranas celulares e/ou estruturas externas possuem expostos ligantes em potencial às lectinas vegetais. Nesse contexto, o conhecimento dos receptores oligossacarídicos torna-se de suma importância para a compreensão do mecanismo de interação com esses ligantes naturais, e para a concepção de glicomiméticos que permitam manipular essa interação.

As técnicas de *glycan array* e microcalorimetria isotérmica de titulação (ITC) foram utilizadas nesse estudo respectivamente para identificar a especificidade fina e caracterizar a afinidade da lectina recombinante de *P. elegans* para vários oligossacarídeos de interesse.

2.1. Estudos de especificidade por *Glycan Array*

Os métodos de análise de especificidade disponível no *Consortium for Functional Glycomics* (CFC) permitiram testar um grande número de oligossacarídeos. PELA foi marcada com o fluoróforo Alexa Fluor 488 e enviada para análise. Mais de 450 oligossacarídeos naturais foram testados pelo método de *Glycan Array*. Três concentrações diferentes de lectina foram testadas 50, 10 e 1 $\mu\text{g/mL}$. Os resultados obtidos para PELA foram de ótima qualidade, mesmo a uma baixa concentração (10 $\mu\text{g/mL}$). A comparação entre os resultados obtidos nas três concentrações diferente permite uma análise semi-quantitativa para os oligossacarídeos mais interessantes (Figura 35).

Os dados observados mostraram uma especificidade marcante da lectina por glicanos complexos biantenados, o que agrupa quase todos os 14 principais compostos (A-N) em destaque na figura 35. Os tipos oligomanose (A a D) e biantenados e simétricos (G a L) são bem reconhecidos, mas uma preferência nata surge por glicanos assimétricos, apresentando um braço curto com apenas uma manose na ramificação $\alpha 1-6$ e um braço longo com pelo menos uma manose e um GlcNAc na ramificação $\alpha 1-3$ (compostos E e F), facilmente observado pela fluorescência relativa alta medida no experimento com esses compostos. A preferência pelos glicanos assimétricos é conservada nas três concentrações testadas da lectina (50, 10 e 1 $\mu\text{g/mL}$). A influência do ácido siálico terminal no braço 1-6 parece ser moderada, uma vez que os oligossacarídeos E e F possuem respostas equivalente na concentração mais baixa testada. A comparação da fluorescência obtida pelos oligossacarídeos E e M, assim como F e N, demonstram claramente que a adição de um GlcNAc no braço 1-6 diminui significativamente a afinidade pela lectina.

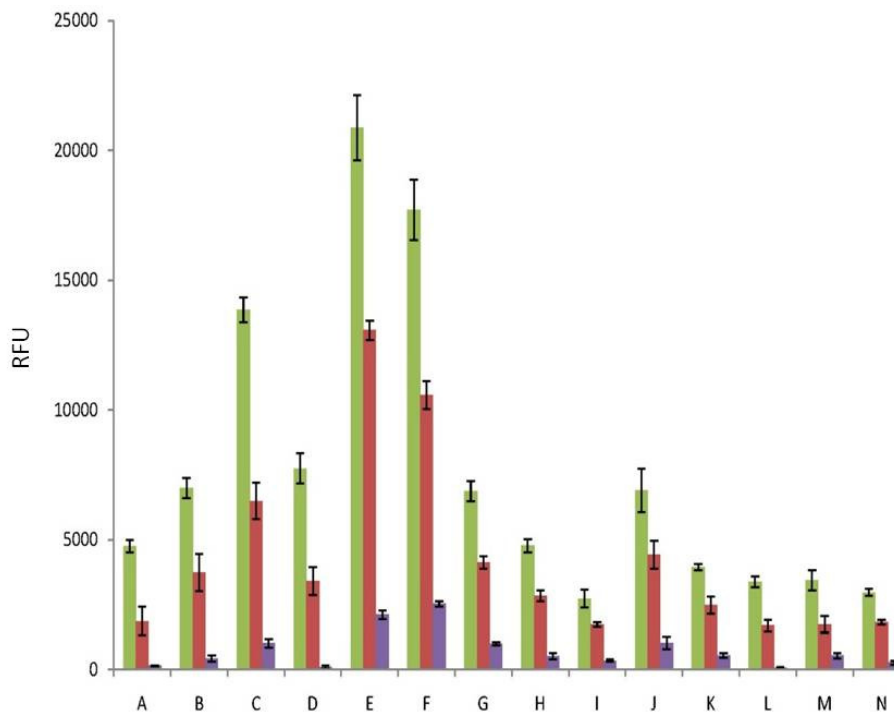
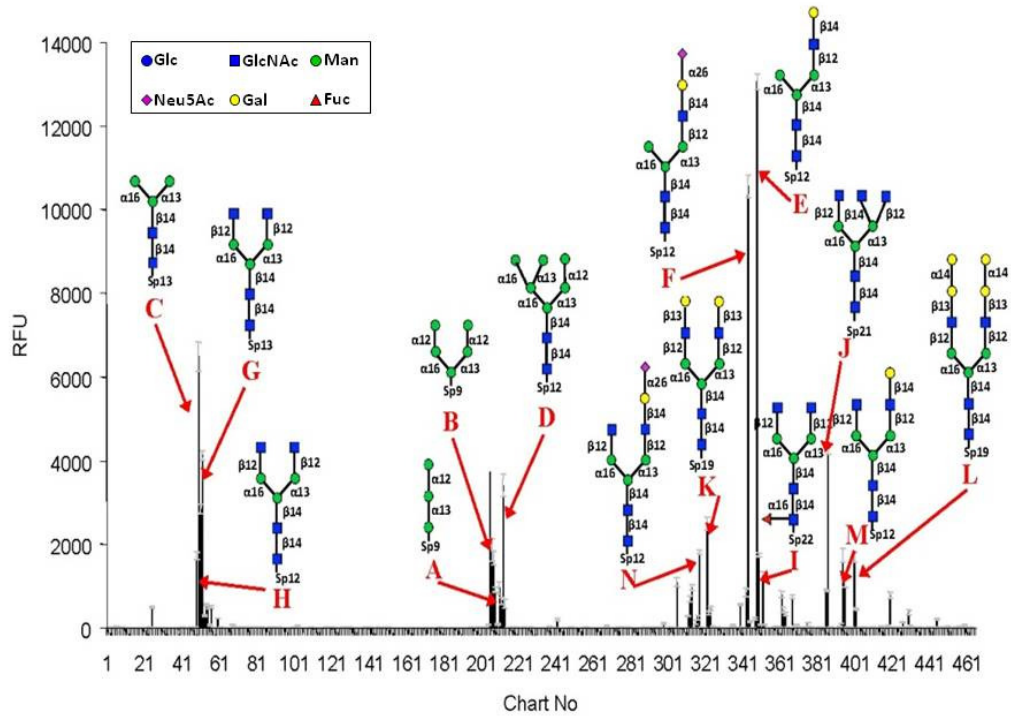


Figura 35. Glycan array. Superior: ligação de PELA marcada (10 μ g/mL) a todos os carboidratos disponíveis do *glycan array* (1 a 461) versus RFU (unidades relativas de fluorescência). Inferior: comparação da ligação em três diferentes concentrações da lectina (verde 50 μ g/mL, laranja 10 μ g/mL, violeta 1 μ g/mL). Resultado completo incluindo a lista de oligossacarídeos compreendidos disponível a partir do *Consortium of Functional Glycomics* (<http://www.functionalglycomics.org/>).

Outras lectinas de leguminosas também tiveram suas especificidades analisadas por experimentos de *Glycan Array*, também disponíveis em <http://www.functionalglycomics.org/>. Entre elas, encontram-se ConA, PSA, DGL e LCA, todas Man/Glc específicas, que apresentaram uma preferência maior a *N*-glicanos do tipo oligomanose, seguida de uma especificidade intermediária a glicanos biantenados, e depois, mais fracamente, para glicanos assimétricos.

Esses dados obtidos permitiram descrever PELa como uma lectina que apresenta uma especificidade rara não-usual por glicanos assimétricos, claramente diferente de outras lectinas de leguminosas Man/Glc específicas da família da ConA.

Os experimentos de microcalorimetria e de cristalografia envolveram análises tanto de compostos usuais como de não-usuais, para tentar explicar as bases moleculares e termodinâmicas para esse comportamento.

2.2. Estudos de afinidade por microcalorimetria (ITC):

A fim de melhor caracterizar as interações entre PELa e oligossacarídeos, os valores de K_a e os parâmetros termodinâmicos de ligação foram determinados utilizando microcalorimetria de titulação, um método que é adequado para a caracterização das interações proteína-carboidrato (DAM & BREWER, 2002).

A microcalorimetria isotérmica de titulação (ITC) foi utilizada para melhor caracterizar as interações de duas moléculas mais interessantes de dois grupos de especificidade com a lectina: um heptassacarídeo biantenado simétrico e um assimétrico (compostos H e E, respectivamente).

Os oligossacarídeos H e E, respectivamente um simétrico apresentando dois braços terminando com GlcNAc, e um assimétrico apresentando uma manose isolado no braço 1-6 e um trissacarídeo β Gal1-4 β GlcNAc1-2 α Man no braço 1-3, foram fornecidos em uma forma ligada à asparagina pelo colaborador do trabalho Carlo Unverzagt (Univ. Baireut, Alemanha). A porção oligossacarídica desses compostos correspondentes a H e E está representada na figura 36. A afinidade entre PELa e os oligossacarídeos foi estimada por microcalorimetria de titulação através da medição do calor liberado em cada injeção de alíquotas do ligante na célula que contém a proteína.

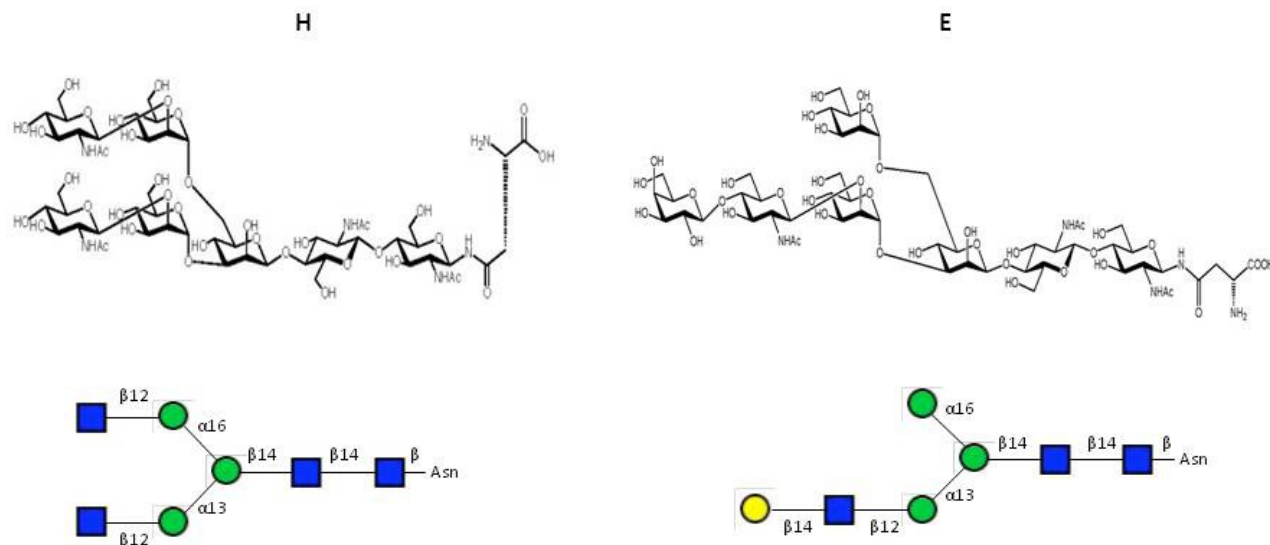


Figura 36. Compostos H (heptassacarídeo simétrico) e E (heptassacarídeo assimétrico) em suas formas ligadas à Asn, utilizados nos experimentos de microcalorimetria.

Nesse experimento, utilizando um modelo simples de um oligossacarídeo por monômero, foi possível determinar que os dois oligossacarídeos se ligam à PELA com afinidades similares, com um K_d em torno de $5 \mu\text{M}$, que é considerada uma afinidade forte para uma interação proteína-carboidrato. Para o composto H, uma melhor curva é obtida utilizando-se um modelo de dois sítios (figura 37), o que indica que cada braço do heptassacarídeo simétrico pode se ligar à lectina, mas com afinidades distintas, embora não tenha sido possível determinar de modo quantitativo as afinidades dos dois sítios.

O estudo cristalográfico da interação entre PELA e o heptassacarídeo simétrico (composto H), descrito no próximo capítulo em detalhes, mostra que qualquer manose (ligadas à manose central nas posições 3 ou 6) podem vir a ocupar o sítio principal, o que pode explicar o modelo de dois sítios.

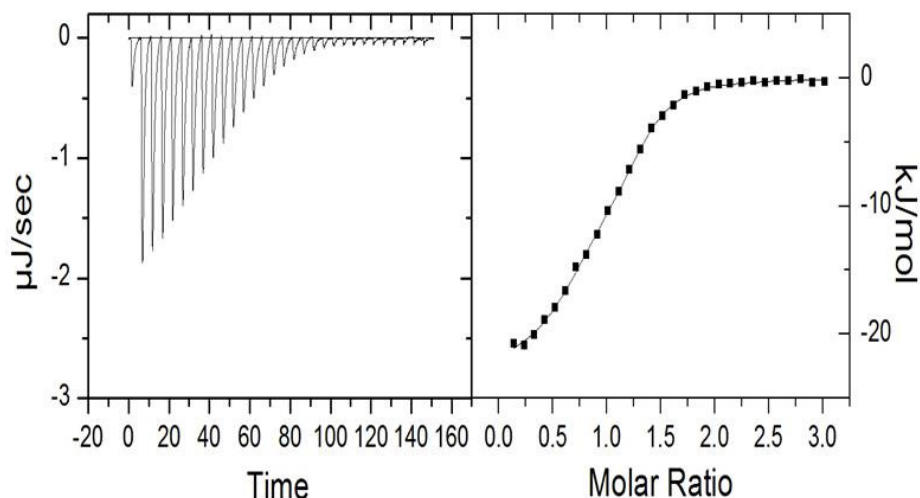


Figura 37. Curva de microcalorimetria obtida para PELa (0,06 mM) com o heptassacarídeo simétrico H (0,775 mM). Termograma à esquerda e integração e ajustamento ao modelo com dois sítios à direita.

Curvas de titulação comparativas foram obtidas para a ligação de PELa aos composto E e H (Figura 38) e os dados termodinâmicos relativos a cada interação estão reunidos na Tabela 8, utilizando-se o modelo de um sítio.

Tabela 8. Dados de ITC para a ligação dos oligossacarídeos a PELa. Todos os valores mensurados são dados como uma média de pelo menos dois experimentos. Desvio-padrão < 0,03 para valores de n e < 5% para demais valores.

Ligante	K_a ($10^6 M^{-1}$)	K_d (μM)	n	ΔG (kJ/mol)	$-\Delta H$ (kJ/mol)	$-T\Delta S$ (kJ/mol)
Composto H	0,22	4,4 (0,3)	1,06 (0,6)	-30,6	23,0 (0,3)	-7,6 (0,4)
Composto E	0,16	6,0 (0,2)	1	-29,8	59,3 (1,8)	29,4 (1,8)

A afinidade de PELa pelos dois oligossacarídeos testados (K_a) é de uma ordem de grandeza elevada (μM) quando comparada a outras interações lectina-açúcar em leguminosas. As únicas atividades lectínicas de afinidade comparáveis foram obtidas com oligossacarídeos, envolvendo as lectinas PAL e ConA (GARCIA-PINO *et al* 2007;. MANDAL *et al.*, 1994, BUTS *et al.*, 2006), mostrando afinidades de ordem micromolar por oligossacarídeos (ConA-trimanose : $K_d = 30 \mu M$; PAL-trimanose : $K_d = 16 \mu M$; PAL-biantenne : $K_d = 16 \mu M$).

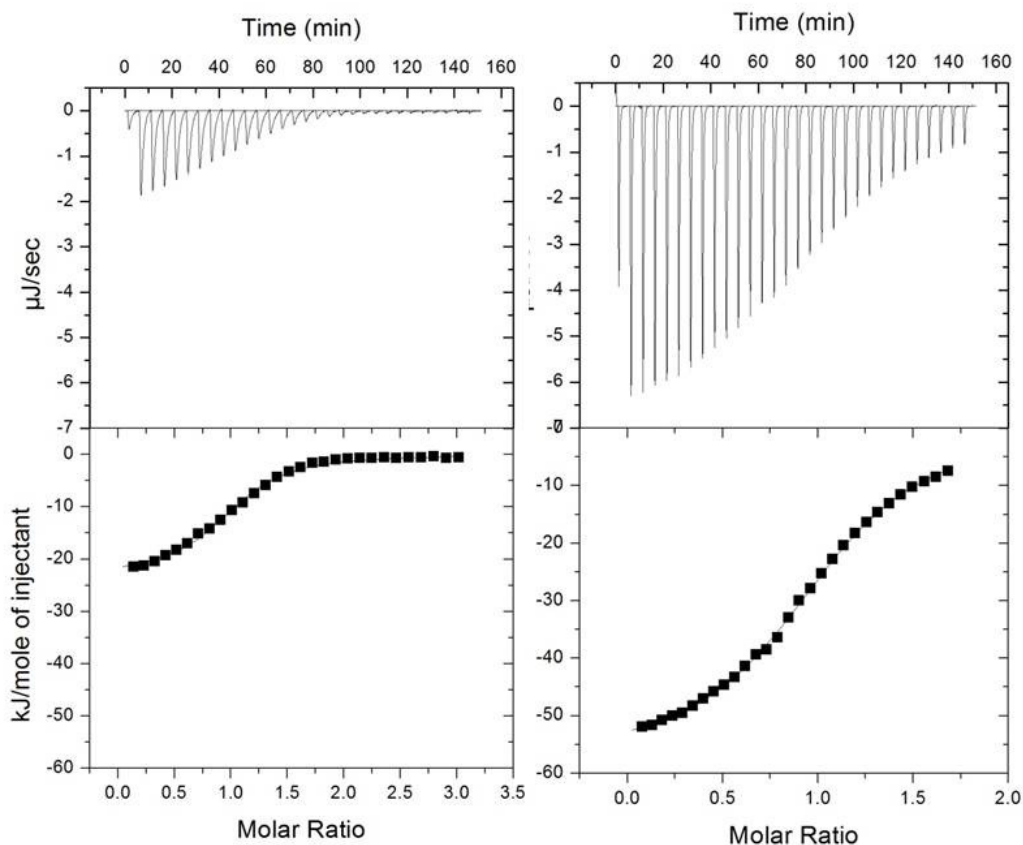


Figura 38. Curvas de microcalorimetria obtidas para PELa (0,6 mM) e os oligossacarídeos H (0,775 mM, à esquerda) e E (0,45 mM, à direita) em tampão Tris (pH 7,6) a 25 °C. Acima, termogramas obtidos de 30 injeções automáticas de 10 μ l de cada açúcar na célula contendo PELa. Abaixo, a integração do calor total liberado em função da titulação do ligante. A linha contínua representa o melhor ajuste para os dados obtidos para um modelo de um sítio.

A afinidade entre o oligossacarídeo H e a lectina PELa é ainda 1,36 vezes superior à afinidade apresentada pelo composto E.

A afinidade de PELa pelos dois oligossacarídeos é próxima, assim como a variação de energia livre (ΔG), apresentando grandezas termodinâmicas similares. A análise da variação de energia livre de ambas as interações, entretanto, indicam que os modos de interações são diferentes, por estratégias distintas. A interação de PELa com o oligossacarídeo H é dirigida por uma contribuição entálpica forte ($\Delta H = -23$ kJ/mol), com uma pequena contribuição entrópica favorável ($-T\Delta S = -7,6$ kJ/mol), perfazendo cerca de 20% da energia de ligação, pouco habitual para interações envolvendo oligossacarídeos. Já a interação de PELa com o oligossacarídeo E apresenta uma contribuição entálpica enorme ($\Delta H = -59,3$ kJ/mol), compensada por uma contribuição entrópica extremamente desfavorável ($-T\Delta S = 29,4$ kJ/mol). A presença de um

GlcNAc adicional no braço 1-6 parece estar correlacionado com a perda de contribuição entálpica. Geralmente, essas associações tipificadas por um termo entálpico favorável são devido ao alto número de ligações de hidrogênio e contatos de Van der Waals entre faces hidrofóbicas do açúcar e aminoácidos aromáticos, que é contrabalanceada por uma contribuição entrópica desfavorável, que tem sido atribuída tanto ao rearranjo de moléculas de solvente, como pela perda de flexibilidade conformacional / mobilidade do ligante (DAM & BREWER, 2002). Essas observações também são confirmadas por Buts *et al.* (2006) que relata que em geral a interação de uma lectina com monossacarídeos é dirigida essencialmente pela entalpia, e que para a maioria dos oligossacarídeos a contribuição entrópica torna-se desfavorável.

Capítulo 5:

Estudos Estruturais da forma cristalina de PELa em complexo com oligossacarídeos

Esse capítulo trata da obtenção e a análise das estruturas cristalinas da lectina PELa determinadas em complexo com dois glicanos distintos, os oligossacarídeos A (trimanose) e H (heptassacarídeo simétrico), representados na Figura 39.

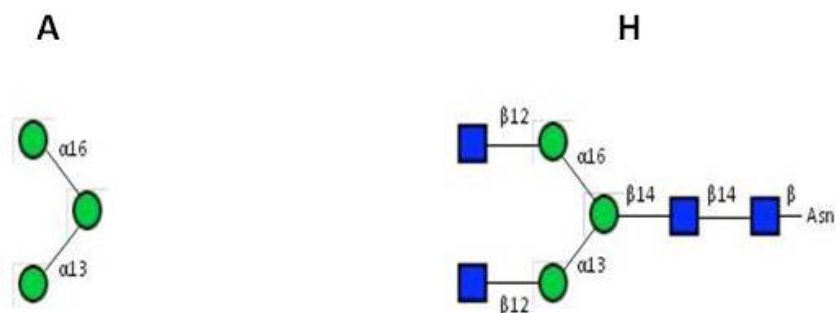


Figura 39. Representação dos Compostos A (trimanose) e H (heptassacarídeo simétrico), utilizados nos experimentos de cristalografia.

O primeiro conjunto (PELa_trim) foi resolvido a uma resolução de 2,1 Å e o segundo a 1,6 Å (PELa_hept). O conjunto PELa_trim foi primeiramente obtido, portanto à partir dele foi descrita as principais análises estruturais da lectina.

1. Estrutura cristalina de PELa em complexo com trimanose

1.1. Cristalização

PELa foi cristalizada a partir de uma solução em torno de 10 mg/mL de proteína solúvel em Tris-HCl 50 mM pH 7,5, NaCl 0,15 M, MnCl₂ 2 mM, CaCl₂ 2 mM, previamente incubada com 1 mM de trimanose (Composto A) por uma hora antes da montagem da placa de cristalização. O trimanose, em detrimento da proteína nativa, foi utilizada para dar mais estabilidade ao sítio de ligação a carboidrato e ao restante da molécula, atuando na estabilidade do empacotamento cristalino. Cristais foram obtidos através do método de difusão de vapor em gota suspensa após 1 mês, com a placa sendo deixada em repouso à temperatura constante de 20°C. A condição que obteve crescimento do cristal foi a 42 do *Crystal Screen II Kit Hampton*, composta por Tris 0,1 M pH 8,5, sulfato de amônio 1,5 M, Glicerol 12% v/v (Figura 40).



Figura 40. Cristais de PELa obtidos na presença de trimanose (α Man1-3(α Man1-6)Man). Condições de cristalização (Hampton *Screen II*, condição 42): Tris 0,1 M pH 8,5, Sulfato de amônio 1,5 M, Glicerol 12% v/v.

1.2. Coleta de dados

Após congelados em solução crioprotetora (condição de cristalização com glicerol a 25% v/v), dados de difração de raios X (comprimento de onda = 0,93340 nm) de cristais de PELa com trimanose foram coletados (expostos a um fluxo contínuo de nitrogênio a 100 K) utilizando como fonte de radiação de luz síncrotron a estação ID14-1, no ESRF (Grenoble, França), usando

um detector ADSC Q4R CCD (Charge-Coupled Device). Foram coletadas 140 imagens de cada cristal, com um ângulo de oscilação igual a $0,75^\circ$ e tempo de exposição ao feixe de radiação de 1,5 segundos por imagem. Os dados foram processados a partir do programa MOSFLM (LESLIE, 1992) e escalonados utilizando o programa SCALA (EVANS, 2006).

Os dados de difração de raios X dos cristais da proteína PELa em complexo com trimanose foram coletados a $2,1 \text{ \AA}$ de resolução. Após indexação dos dados, determinou-se como grupo espacial o primitivo ortorrômbico $P2_12_12_1$ (ângulos $\alpha = \beta = \gamma = 90^\circ$), com parâmetros de célula $a = 42,62 \text{ \AA}$, $b = 80,66 \text{ \AA}$ e $c = 148,75 \text{ \AA}$. Os valores finais foram obtidos após o refinamento destes parâmetros e utilizando as 140 imagens coletadas. O Coeficiente de Matthews ($2,19 \text{ \AA}^3/\text{Da}$) indicou a presença de dois monômeros na unidade assimétrica e a presença de 43,55% de solvente. Segue abaixo a tabela 9 com estatísticas detalhadas da coleta e do processamento de dados.

Tabela 9. Estatística da coleta e do processamento de dados de difração de raios X da lectina PELa em complexo com trimanose.

Parâmetros	Valores
<i>Dados de Coleta</i>	
Linha / Comprimento de onda	Id14EH1 / $0,933 \text{ \AA}$
Grupo espacial	$P2_12_12_1$
Mosaicidade estimada	0,97
<i>Parâmetros de célula unitária</i>	
a, b, c (\AA)	42,62; 80,66; 148,75
Ângulos de célula	$\alpha=\beta=\gamma=90^\circ$
Número de reflexões totais/únicas	99535 / 28894
Quantidade de Solvente (%)	43,55
Monômeros por unidade assimétrica	2
Limite de resolução (\AA)	2,10 (2,2-2,1) ^a
R_{merge} (%)	0,082 (0,330)
Completeza (%)	93,7 (70,1)
Multiplicidade	3,4 (2,4)

I/σI (média)	9,3 (2,0)
--------------	-----------

^a Valores em parêntese representam a camada de mais alta resolução.

1.3. Substituição molecular e refinamento da estrutura

Para resolver o problema de fase foi escolhido o modelo de uma lectina semelhante para fazer a substituição molecular. A estrutura utilizada como modelo de busca para substituição molecular foi a da lectina *P. angolensis* (PDB : 1N3O), com a qual a lectina PELA possui 76% de indentidade de sequência. Para substituição molecular foi usado o programa PHASER (MCCOY *et al.*, 2005), integrante do pacote CCP4 (COLLABORATIVE COMPUTATIONAL PROJECT N° 4; CCP4, 1994). A partir de parâmetros como *log-likelihood gain* de 1243,47 e RFZ de 14,2 e TFZ de 28,7, que demonstram a qualidade do refinamento feito pela molécula modelo escolhida, foi comprovada a eficácia na escolha do modelo 1N3O.

Após a substituição molecular, a estrutura obtida foi submetida a refinamentos de corpo rígido e refinamentos posicionais para um melhor ajuste da molécula na unidade assimétrica através do uso do programa REFMAC5 (MURSHUDOV *et al.*, 1997). A finalidade deste refinamento foi encontrar uma melhor concordância entre o modelo proposto e a estrutura real da molécula, colocando o modelo na mesma posição que a molécula real ocupa na sua rede cristalina. Para avaliar os resultados dos cálculos do refinamento foram observados os valores de R_{factor} e R_{free} .

Os aminoácidos foram posicionados corretamente um a um utilizando o programa COOT (EMSLEY & COWTAN, 2004), assim como a molécula de trimanose, de acordo com a densidade eletrônica obtida, e posteriormente foram adicionadas 326 moléculas de água. Após o último refinamento a estrutura obteve resolução máxima de 2,1 Å, com valores de R_{factor} de 19,5% e R_{free} de 25,2%. Os dados estatísticos de substituição molecular assim como os de refinamento são mostrados na tabela 10.

A qualidade estereoquímica da estrutura foi comprovada a partir do gráfico de Ramachandran, gerado pelo programa Molprobity (CHEN *et al.*, 2010), que analisa as torções dos ângulos Φ e Ψ . O gráfico mostrou um único resíduo em regiões não permitidas (0,21%), e 4,7% dos resíduos presentes em regiões adicionalmente permitidas.

Tabela 10. Estatística dos dados de substituição molecular e refinamento da estrutura de PELA em complexo com trimanose.

Parâmetros	Valores
Faixa de resolução utilizada (Å)	42,60 - 2,10
Nº de reflexões utilizadas no refinamento	27309
Wilson B (Å ²)	24,7
R_{cryst}/R_{free}	19,5/25,2
$R_{msd} bonds$ (Å)	0,015
$R_{msd} angles$ (°)	1,544
$R_{msd} chiral$ (Å ³)	0,106
Número de átomos de proteína (A/B)	1935/1926
Bfac (Å ²)	23,5/23,7
Número de moléculas de água (A/B)	177/149
Bfac (Å ²)	29,0/28,3
Número de átomos de ligante	34
Bfac (Å ²)	22,6
Heteroátomos (A/B)	18/8
Bfac (Å ²)	32,4/33,5
Coefficiente de correlação final $F_{obs} - F_{calc}$ (%)	94,5
Coefficiente de correlação livre (%)	91,0
Código PDB	3ZVX

1.4. Análise da estrutura PELA em complexo com trimanose

A lectina PELA, complexada com o ligante trimanose (Man α 1-3Man α 1-6Man), foi refinada observando-se duas moléculas de PELA e uma molécula do ligante na unidade assimétrica. Cada monômero tem o *fold* clássico de lectinas de leguminosas (sanduíche β), detalhes os quais já são bem conhecidos (CHANDRA *et al.*, 2001), incluindo um íon de cálcio e um manganês no sítio de ligação. PELA apresentou-se formando um dímero canônico, também observado em muitas outras lectinas de leguminosas diméricas e tetraméricas (MANOJ & SUGUNA, 2001;

SRINIVAS *et al.*, 2001). A estrutura tridimensional dimérica completa da PELa complexada com trimanose é mostrada na figura 41 e a estrutura do monômero na figura 42. É possível observar a molécula de trimanose localizada no sítio de ligação a carboidrato, assim como os metais Ca^{2+} e Mn^{2+} situados no sítio de ligação a metais. Dois potenciais sítios de glicosilação (Asn7 e Asn127) encontram-se situados respectivamente numa região N-terminal desordenada e sobre uma pequena hélice- α , ambos na superfície da proteína, sendo, portanto susceptíveis a estarem glicosilados na planta.

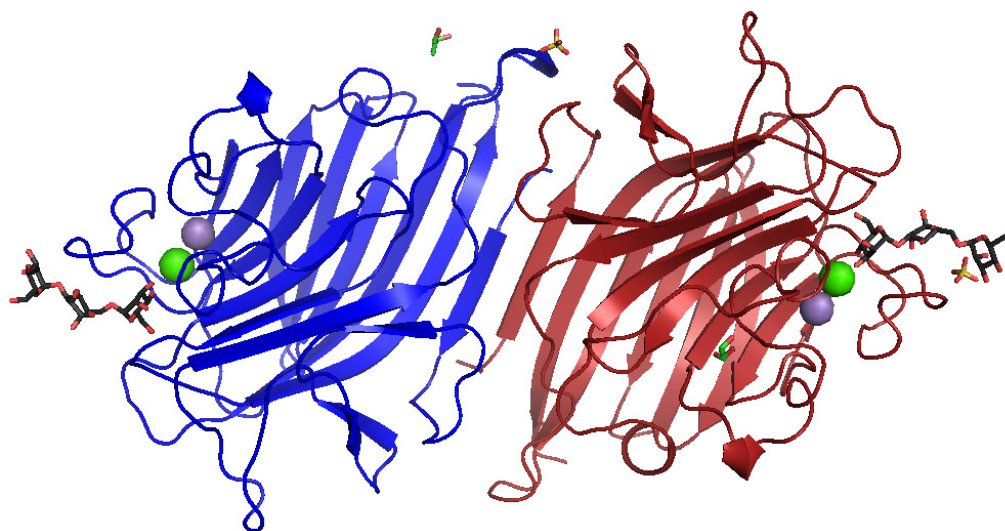


Figura 41. Visão geral da estrutura dimérica de PELa em complexo com trimanose, formada pelos monômeros A em vermelho e monômero B em azul. Presença dos metais Ca^{2+} e Mn^{2+} (esferas verde e roxa, respectivamente), do ligante trimanose ($\alpha\text{Man1-3}(\alpha\text{Man1-6})\text{Man}$) (preto), de duas moléculas de glicerol (verde) e de dois íons sulfato (amarelo).

Os monômeros são compostos por 14 fitas beta e 2 pequenas hélices alfa, sendo interligadas por voltas e alças formando uma estrutura de forma achatada. Esse tipo de enovelamento é conhecido como “ β -sanduíche”, por ser composto majoritariamente de folhas betas antiparalelas (cerca de 60%). A orientação antiparalela das folhas beta, e a mudança entre uma face e outra, forma uma estrutura rígida forte, semelhante a um sanduíche. As duas hélices alfa identificadas em cada monômero contêm uma única volta, formada pelos resíduos Asn₁₂₇ Glu₁₂₈ Ser₁₂₉ Ala₁₃₀ Asn₁₃₁ e Asp₂₁₁ Leu₂₁₂ Thr₂₁₃ Thr₂₁₄ Ile₂₁₅.



Figura 42. Estrutura monomérica de PELA. Ligante trissacarídeo em preto e íons cálcio e manganês em esferas verde e azul, respectivamente.

A análise do primeiro mapa de densidade eletrônica revelou de forma clara a posição do trissacarídeo, formando uma ponte entre dímeros no cristal (Figura 43). Como uma consequência do empacotamento cristalino, os sítios de ligação a carboidrato das subunidades A e B na unidade assimétrica não são equivalentes.

Como o açúcar trissacarídeo está posicionado formando uma ponte entre monômeros A e B de diferentes dímeros (Figura 43), o sítio de ligação a carboidrato de cada monômero compartilha porções de um mesmo ligante. A figura 44 estabelece todas as distâncias das ligações de hidrogênio presentes no sítio de ligação dos monômeros A e B da lectina, assim como a porção do ligante em que cada monômero encontra-se complexado.

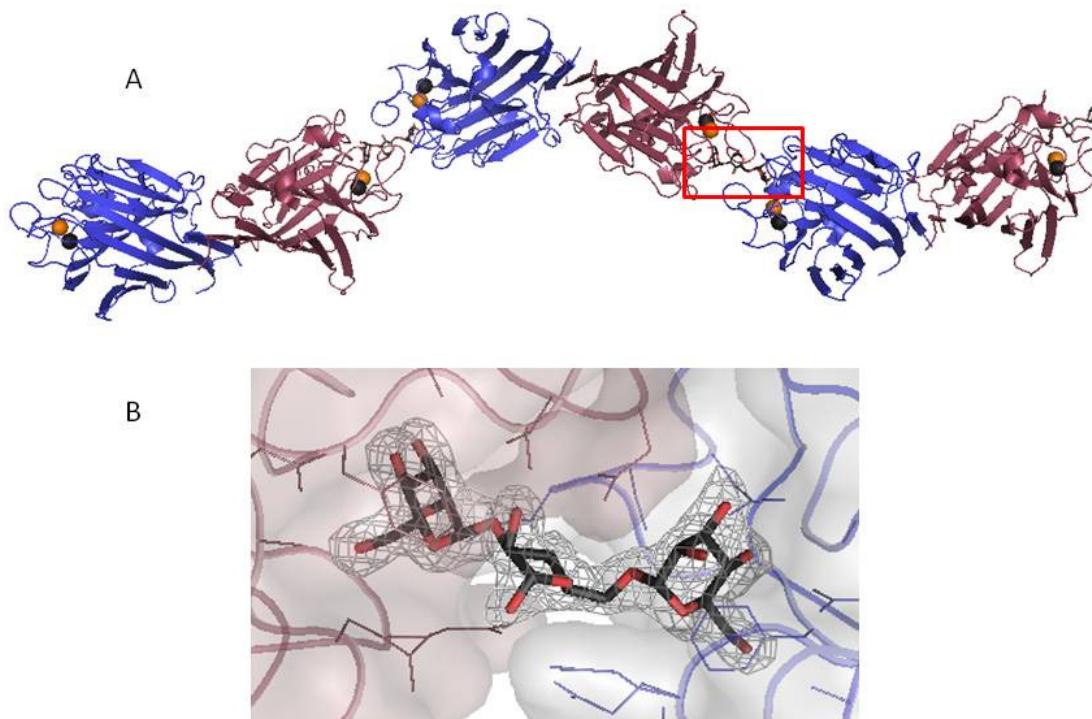


Figura 43. Estrutura cristalina de PELa em complexo com trimanose. A. Ponte entre dímeros unidos por uma molécula de trimanose. B. Densidade eletrônica experimental do ligante trimanose entre monômeros A (vermelho) e B (azul) de dímeros distintos.

Os monômeros A e B fixam, respectivamente, os α Man ligados à manose central nas posições 1-3 e 1-6. Em cada sítio, os aminoácidos implicados são aqueles que se encontram habitualmente nas interações lectina de leguminosa-manose (Asp95, Ala113, Gly115, Asn145, Ser146, Asn147, Gln231). O modo de ligação dos dois resíduos de manose são muito similares, cada uma estabelecendo 5 ligações de hidrogênio diretas com cadeias laterais de Asp95, Asn 147 e o nitrogênio da cadeia principal de Gly115 e Glu230. Contatos hidrofóbicos envolvendo Ala94 e Phe141 também estão presentes. Essa interação corresponde ao modo de ligação clássico a manose observado em PAL e ConA (LORIS *et al.*, 2004; DEREWENDA *et al.*, 1989). Uma interação com uma molécula de água conservada é também observada em ambos os sítios entre O2 e o oxigênio da carbonila de Ala113. O resíduo de manose central estabelece interações com ambos os monômeros, envolvendo Asn145, Ser146 e Gln231 da cadeia A e Glu230 da cadeia B. A ponte do trimanose resulta em um contato mais próximo entre as faces dos dois monômeros protéicos. Contatos proteína-proteína são criados com a participação de ligações de hidrogênio entre a cadeia lateral de Arg123B e as cadeias principais de Thr148A e Asp150A. O resíduo Asn145 de cada monômero interfere com o sítio de ligação do outro monômero, e participam de

uma ligação de hidrogênio entre O2 da manose e Asn145B ou no contato hidrofóbico de C2 com Asn145A (figura 45).

Os dois átomos cálcio e manganês estão ligados de forma reversível à proteína dentro de um sítio próximo ao de reconhecimento a açúcar. Esses metais estão localizados a uma distância de 4.21Å um do outro e estão coordenados por cadeias laterais de seis aminoácidos (Glu137, Asp139, Phe141, Asn147, Asp150, His155) e por quatro moléculas de água. O cálcio mantém a conformação *cis* do Asp95 do sítio de ligação ao açúcar que é conservado em todas as lectinas de leguminosas.

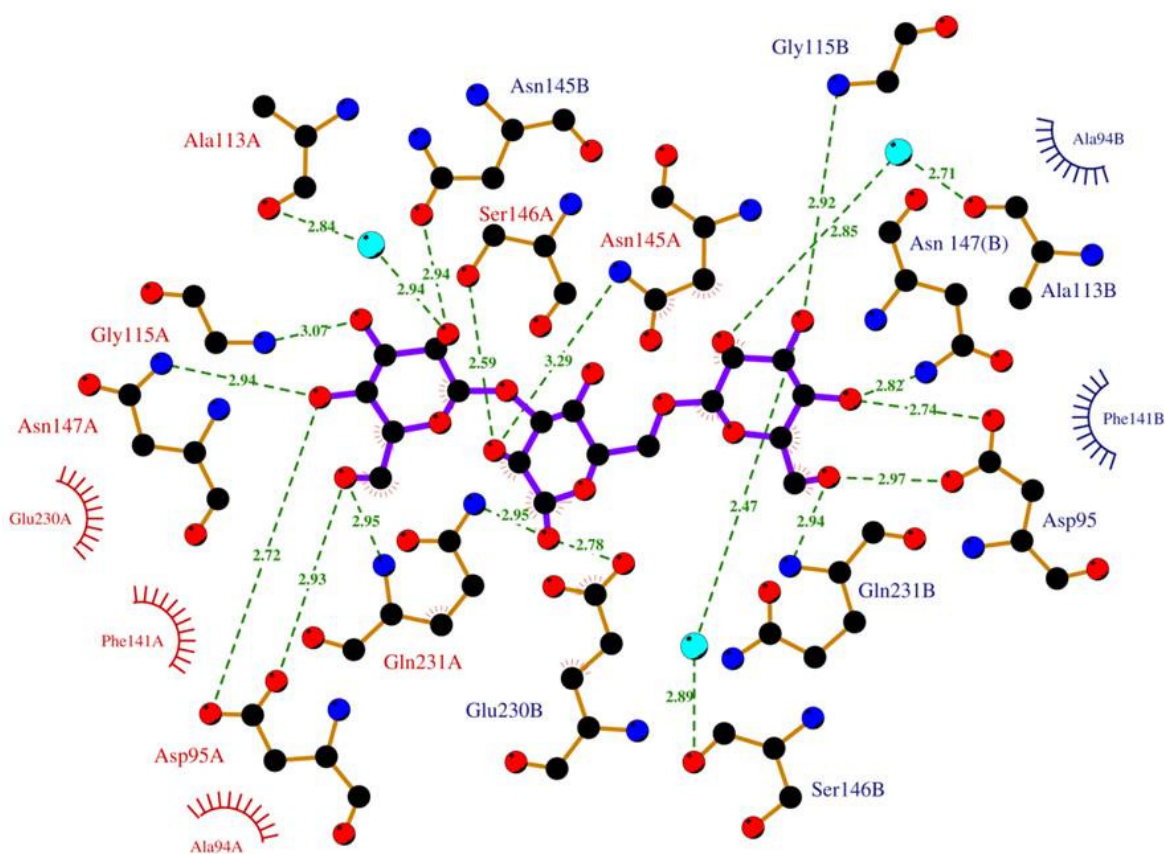


Figura 44. Representação esquemática bidimensional (LigPlot) da rede de pontes de hidrogênio entre os resíduos chave envolvidos na interação com o ligante trimanose, compartilhado entre os monômeros A (resíduos em vermelho) e B (resíduos em azul) de dímeros distintos de PELA. Comprimento das ligações em Å.

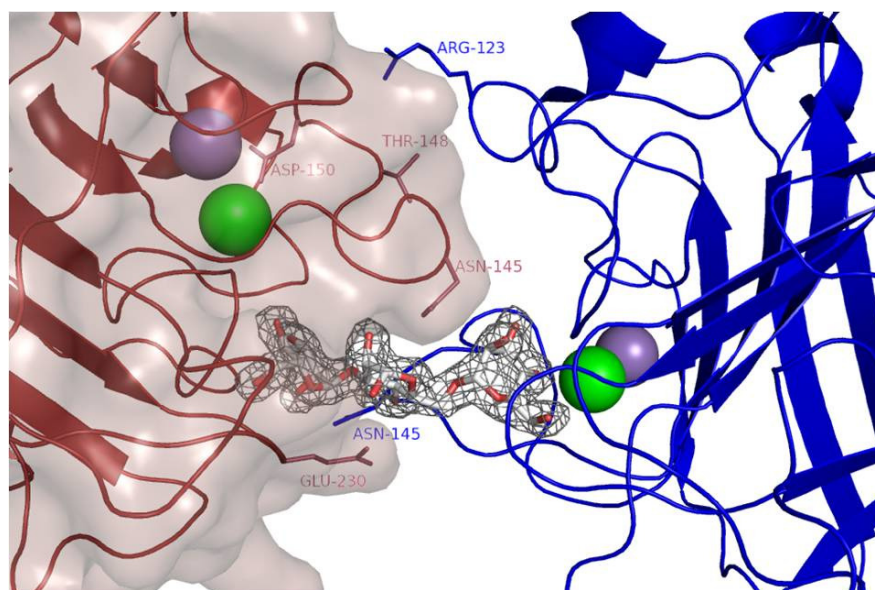


Figura 45. Estrutura cristalina do complexo PELA/trimanose. Visão da interface entre as cadeias A e B de PELA em ponte com o trimanosídeo representado junto com sua densidade eletrônica. Cadeia A representada por superfície (vermelho) e cadeia B como *cartoon* azul.

Analisando o sítio de ligação a carboidrato e o de ligação a metal de PELA, é possível fazer algumas comparações em relação a outras estruturas de lectinas de leguminosas. O sítio de ligação a carboidrato de todas as lectinas de leguminosas consiste em resíduos pertencentes a até cinco voltas polipeptídicas, nomeados de A a E (SHARMA & SUROLIA, 1997), as quais variam em diferentes graus entre lectinas com diferentes especificidades. A figura 46 traz a representação dessas voltas em PELA e seu envolvimento no complexo com trimanose para cada monômero.

As voltas A e B, em geral, contêm um resíduo Asp essencial (invariavelmente precedido por uma ligação cis-peptídica) e um grupo -NH suporte (usualmente de um resíduo Gly), respectivamente. Em PELA, foi possível identificar o resíduo Asp95 na volta A e a Gly115 na volta B, concordando com essa arquitetura, típica do reconhecimento a manose. O resíduo Ala145 do sítio de ligação de PELA a trimanose também participa da volta B. As conformações desses dois trechos não varia muito entre as diferentes lectinas de leguminosas.

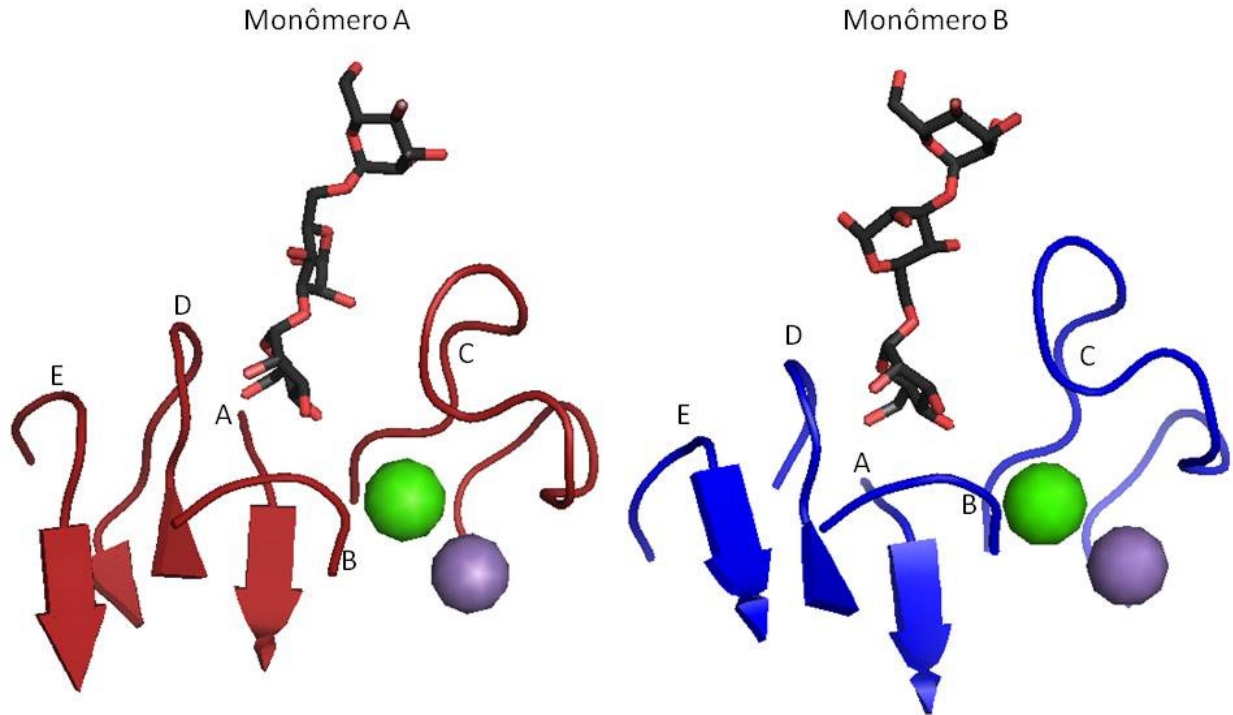


Figura 46. Representação da arquitetura do sítio de ligação a carboidrato de PELA complexada com trimanose, evidenciando as 5 voltas (A-E) típicas de lectinas de leguminosas. Nessa interação, somente resíduos pertencentes às voltas A, B, C e D participaram da ligação ao carboidrato.

A volta C é a de ligação a metal, que envolve os íons cálcio e manganês estruturalmente importantes. Em estruturas cristalinas conhecidas, cinco diferentes tamanhos de volta (12-16 resíduos) adotando oito diferentes conformações são observadas (LORIS *et al.*, 2004). Não existe uma relação simples entre tipo de especificidade a monossacarídeos e o tamanho e conformação da volta C. Em PELA, essa volta tem o comprimento de 16 resíduos, similar em comprimento e conformação ao da lectina de *P. angolensis* (PAL) e de das lectinas *Maackia amurensis* e *Ulex europaeus*, sialolactose e quitobiase específicas (LORIS *et al.*, 2004; IMBERTY *et al.*, 2000; LORIS *et al.*, 2000). Outras lectinas Man/Glc específicas com estruturas cristalinas disponíveis têm comprimentos de volta de 14 resíduos (ConA e outras relacionadas) ou 15 (LOL e outras relacionadas). Cadeias laterais específicas nessa volta influenciam na natureza do açúcar que pode ser acomodado no sítio de ligação (LORIS *et al.*, 2004). Entre os resíduos descritos no sítio de ligação a trimanose de PELA, encontram-se nessa volta os resíduos Asn145, Ser146 e Asn147.

Em contraste às voltas A-C, a D não interage diretamente com o íon cálcio estrutural. Em sua grande variabilidade de comprimento, conformação e sequência, é referida comumente como

a volta determinante da especificidade ao monossacarídeo, assim como do oligossacarídeo (LORIS *et al.*, 1998). Em PELa, a conformação adotada por essa volta de 9 resíduos é idêntica às encontradas em estruturas de lectinas Man/Glc específicas (LORIS *et al.*, 2004; HAMELRYCK *et al.*, 1988; SANZ-APARICIO *et al.*, 1997; WAH *et al.*, 2001; DEREWENDA *et al.*, 1989; BOURNE *et al.*, 1990; RINI *et al.*, 1993), suportando a idéia. Nessa volta encontramos os resíduos Glu230 e Gln231, participantes do sítio de ligação a trimanose.

Finalmente, e em poucos casos, a volta E é encontrada interagindo com o carboidrato ligado, geralmente entre oligossacarídeos e lectinas, agindo como um subsítio, como na leucoaglutinina de *M. amurensis* (MAL) em complexo com sialolactose (LORIS *et al.*, 2000), a lectina IV de *Griffonia simplicifolia* (GS-IV) em complexo com um tetrassacarídeo (DELBAERE *et al.*, 1993) e ConA e PAL em complexo com Man(α 1-2)Man (MOOTHOO *et al.*, 1999, LORIS *et al.*, 2004). Em PELa, essa volta não se encontra envolvida na ligação ao carboidrato trimanose.

Em razão de sua interação com os dois monômeros, a forma do trimanosídeo em PELa é estendida e diferente do que é observado pelo mesmo ligante em outras estruturas de lectinas de leguminosas tais como PAL (LORIS, 2004), ConA (LORIS, 1996) e DGL (ROZWARSKI *et al.*, 1998). A comparação entre PELa e PAL confirmam a similaridade dos sítios de ligação com conformação idêntica para as cinco voltas nomeadas A a E (Figura 47). PAL se liga ao trimanosídeo com o braço 1-6 dentro do sítio de ligação principal da mesma forma que o monômero B de PELa. Entretanto, a conformação do oligossacarídeo observada em PAL é bem diferente, já que o trimanosídeo encontra-se curvado em torno do resíduo fenil do sítio de ligação e a manose-3 estabelece uma ligação de hidrogênio com um resíduo Asn da volta A.

Uma comparação com lectinas da tribo Dioclae é também exibida na figura 47. As voltas do sítio de ligação são bem conservadas entre Dioclae (ConA e DGL) e Dalbergiae (PAL e PELa), com exceção da volta C, que é mais curta em ConA/DGL que em PAL/PELa (14 aminoácidos no lugar de 16, respectivamente), resultando em um sítio de ligação mais aberto que o descrito previamente (LORIS *et al.*, 2004). Como consequência, a conformação do trimanosídeo é diferente em ConA e DGL é voltado sobre a superfície da proteína e a manose 1-3 está em contato com aminoácidos suplementares da volta C. A conformação observada em ConA não é possível em PELa, em razão da conformação diferente da volta C, que implicaria em um impedimento estérico.

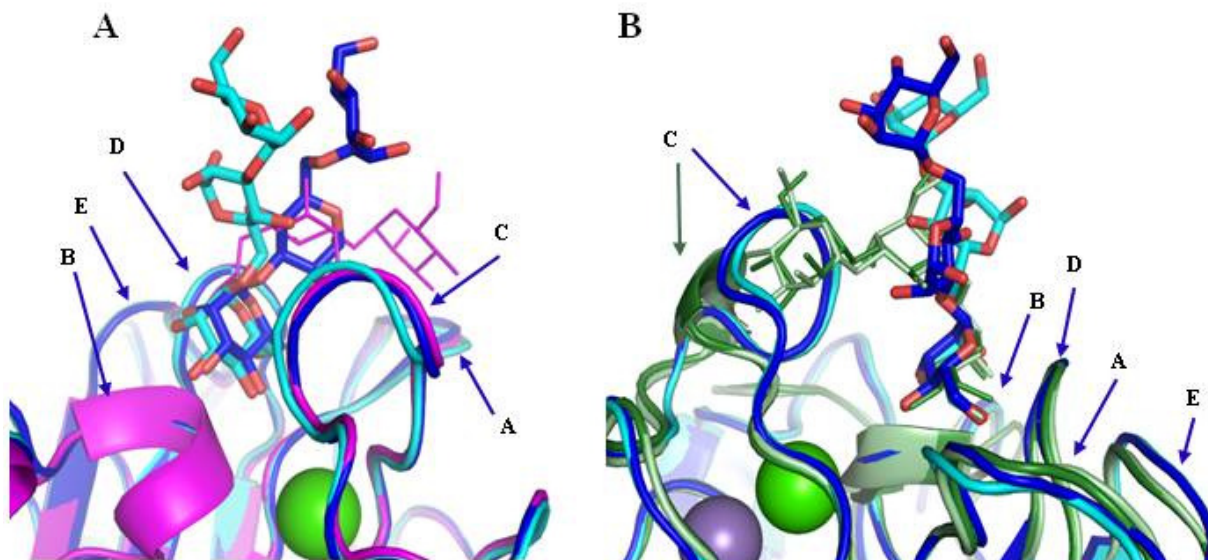


Figura 47. Comparação do reconhecimento do trissacarídeo (Man) por PELA (A) e outras lectinas de leguminosas. A. Sítios de ligação superpostos dos monômeros A (azul) e B (ciano) do complexo PELA/trimanose com o complexo PAL/trimanose (magenta, código PDB: 1Q8V). B. Vista dos monômeros A e B de PELA com as lectinas ConA (verde escuro, código PDB :1ONA) e DGL (verde claro, código PDB : 1DGL). As voltas A-E do sítio de ligação estão evidenciadas.

A cristalização de PELA em complexo com trimanose resultou em uma cadeia infinita de lectinas no cristal com um carboidrato ligando fazendo uma ponte entre dois monômeros através de seu resíduo de manose terminal não-redutor. Mesmo o trimanosídeo tendo sido observado em estruturas cristalinas de diversas lectinas, esse estudo reporta o primeiro caso de cadeia infinita por ligação ao trimanose observado em cristalografia de lectinas de leguminosas. A conformação resultante do trimanosídeo é extremamente estendida e drasticamente diferente da forma curvada previamente observada em outros complexos com lectinas de leguminosas (LORIS *et al.*, 2004; LORIS *et al.*, 1996, ROZWARSKI *et al.*, 1998).

2. Estrutura cristalina de PELA em complexo com heptassacarídeo simétrico

2.1. Cristalização

Uma segunda condição cristalina apresentou PELA complexada ao heptassacarídeo biantenado simétrico $\text{GlcNAc}\beta 1\text{-2Man}\alpha 1\text{-3(GlcNAc}\beta 1\text{-2Man}\alpha 1\text{-6)Man}\beta 1\text{-4GlcNAc}\beta 1\text{-}$

4GlcNAc β -Asn (composto H, figura 39). PELa foi cristalizada a 10 mg/mL em Tris-HCl 50 mM pH 7,5, NaCl 0,15 M, MnCl₂ 2 mM, CaCl₂ 2 mM, tendo sido previamente incubada com 1mM do heptassacarídeo por uma hora antes da montagem da placa de cristalização. A condição que obteve crescimento de cristais foi a 48 do *Crystal Screen II Kit Hampton*, composta por Bicina 100 mM pH 9, PEG 20K 10% e dioxano 2% (Figura 48).

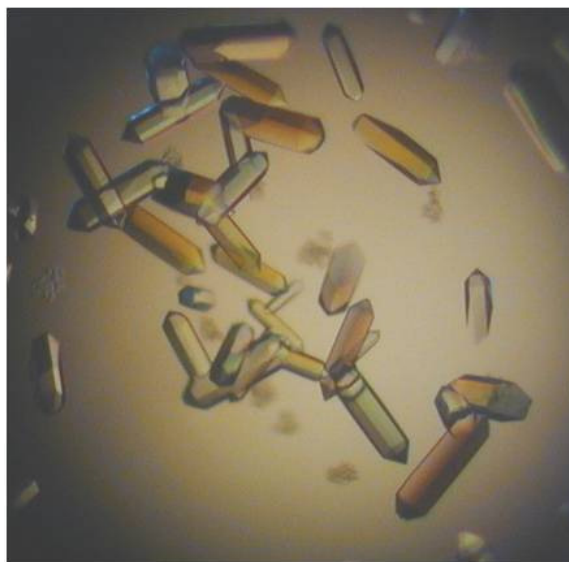


Figura 48. Cristais de PELa obtidos na presença de heptassacarídeo simétrico (GlcNAc β 1-2Man α 1-3(GlcNAc β 1-2Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAc β -Asn). Condições de cristalização (Hampton *Screen II*, condição 48): Bicina 100 mM pH 9, PEG 20K 10% e dioxano 2%.

2.2. Coleta de dados

Após congelados em solução crioprotetora (condição de cristalização com glicerol a 30% v/v), dados de difração de raios X de cristais de PELa com heptassacarídeo foram coletados a 1,6 Å (expostos a um fluxo contínuo de nitrogênio a 100 K), utilizando como fonte de radiação de luz síncrotron a estação BM30A, no ESRF (Grenoble, França), usando um detector ADSC Q4R CCD (*charge-coupled device*) e uma estratégia similar ao conjunto anterior. Os dados foram processados a partir do programa Mosflm (LESLIE, 1992) e escalonados utilizando o programa SCALA (EVANS, 2006).

Após indexação dos dados, determinou-se como grupo espacial o P2₁2₁2₁ (ângulos $\alpha = \beta = \gamma = 90^\circ$), com parâmetros de célula $a = 51,4 \text{ \AA}$, $b = 76,98 \text{ \AA}$ e $c = 125,59 \text{ \AA}$. Os valores finais foram obtidos após o refinamento destes parâmetros e o coeficiente de Matthews ($2,29 \text{ \AA}^3/\text{Da}$)

indicou a presença de dois monômeros na unidade assimétrica e a presença de 43,65% de solvente. Segue abaixo a tabela 11 com estatísticas detalhadas da coleta e do processamento de dados.

Tabela 11. Estatística da coleta e do processamento de dados de difração de raios X da lectina PELa em complexo com heptassacarídeo simétrico.

Parâmetros	Valores
<i>Dados de Coleta</i>	
Linha / Comprimento de onda	BD30A / 0,9797 Å
Grupo espacial	P2 ₁ 2 ₁ 2 ₁
<i>Parâmetros de célula unitária (Å)</i>	
a, b, c	51,4; 76,98; 125,59
Ângulos de célula	$\alpha=\beta=\gamma=90^\circ$
Número de reflexões totais/únicas	186416/55022
Quantidade de Solvente (%)	43,65
Moléculas por unidade assimétrica	2
Limite de resolução (Å)	1,65 (1,74-1,65)
Rmerge ^b (%)	0,054 (0,250) ^a
Completeza (%)	91,4 (65,8)
Multiplicidade	3,4 (2,6)
I/ σ I (média)	13,6 (3,9)

^a Valores em parênteses representam a camada de mais alta resolução.

2.3. Substituição molecular e refinamento da estrutura

Para resolver o problema da fase por substituição molecular desse conjunto de dados, utilizou-se como modelo as coordenadas da estrutura de PELa em complexo com trimanose (PDB: 3ZVX). Para substituição molecular foi usado o programa PHASER (MCCOY *et al.*, 2005), integrante do pacote CCP4 (COLLABORATIVE COMPUTATIONAL PROJECT N° 4; CCP4, 1994). Após a substituição molecular, a estrutura obtida foi submetida a refinamentos de corpo rígido e posicionais para um melhor ajuste da molécula na unidade assimétrica através do

uso do programa REFMAC5 (MURSHUDOV *et al.*, 1997). Para avaliar os resultados dos cálculos do refinamento foram observados os valores de R_{factor} e R_{free} .

A observação do heptassacarídeo no complexo cristalino foi feita pelo programa COOT (EMSLEY & COWTAN, 2004), de acordo com a densidade eletrônica experimental obtida claramente, e posteriormente foram adicionadas 518 moléculas de água. Após o último refinamento, a estrutura obteve resolução máxima de 1,65 Å, com valores de R_{factor} de 15,5% e R_{free} de 18,8%. Os dados estatísticos de substituição molecular assim como os de refinamento são mostrados na tabela 12.

Tabela 12. Estatística dos dados de substituição molecular e refinamento da estrutura de PELa em complexo com heptassacarídeo simétrico.

Parâmetros	Valores
Faixa de resolução utilizada (Å)	47,57-1,65 (1,74-1,65) ^a
Nº de reflexões utilizadas no refinamento	52116
Wilson B (Å ²)	14,00
$R_{\text{crist}}/R_{\text{free}}$	15,5/18,8%
$R_{\text{msd bonds}}$ (Å)	0,015
$R_{\text{msd angles}}$ (°)	1,59
$R_{\text{msd chiral}}$ (Å ³)	0,098
Número de átomos de proteína (A/B)	2004/1986
Bfac (Å ²)	12,4/14,0
Número de moléculas de água (A/B)	295/223
Bfac (Å ²)	25,7/26,6
Número de átomos de ligante (A/B)	89/92
Bfac (Å ²)	14,2/16,6
Heteroátomos (A/B)	30/12
Bfac (Å ²)	24,7/25,7
Coefficiente de correlação final $F_{\text{obs}} - F_{\text{calc}}$ (%)	96
Coefficiente de correlação livre (%)	94,8
Código PDB	3ZYR

^a Valores em parêntese representam a camada de mais alta resolução.

A qualidade estereoquímica da estrutura foi comprovada a partir do gráfico de Ramachandran, gerado pelo programa Molprobity (CHEN *et al.*, 2010). O gráfico mostrou nenhum resíduo em regiões não permitidas, e 2,93% dos resíduos presentes em regiões permitidas.

2.4. Análise da estrutura de PELa em complexo com heptassacarídeo

A estrutura cristalina de PELa em complexo com heptassacarídeo simétrico foi determinada em alta resolução (1,65 Å), com o cristal contendo um dímero de PELa em sua subunidade assimétrica (monômeros A e B) e duas moléculas do heptassacarídeo, uma em cada monômero. A estrutura geral mostra-se inalterada na presença do complexo, apresentando características similares àquela já descrita em complexo com trimanose. Foi possível observar em cada monômero A e B uma molécula de heptassacarídeo localizada no sítio de ligação a carboidrato, assim como os metais Ca^{2+} e Mn^{2+} (Figura 49). A figura 50 mostra a densidade eletrônica experimental sobre o heptassacarídeo no sítio de ligação de PELa, próximo aos íons Ca^{2+} e Mn^{2+} . Os contatos do oligossacarídeo com a superfície protéica encontram-se bem mais estendidos que no complexo com o trimanosídeo. A maior parte das interações se formam sobre o trimanosídeo central e o braço 1-3 (GlcNAc β 1-4Man α 1-3)Man, confirmando também as preferências observadas nos experimentos de *glycan array*. Os oligossacarídeos na cadeia A e B são muito semelhantes em termos de conformação, e toda a descrição abaixo se refere ao monômero B, salvo indicação contrária.

O sítio de ligação de PELa ao heptassacarídeo simétrico (GlcNAc β 1-2Man α 1-3(GlcNAc β 1-2Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAcB-Asn) envolve principalmente 13 resíduos de aminoácidos que interagem com a porção biantenada do ligante da porção glicídica, sobretudo às porções manose central e braço 1-3 GlcNAc β 1-2Man α 1-3. A figura 51 descreve todas as distâncias da rede de ligações de hidrogênio presentes no sítio de ligação da lectina PELa com a molécula de heptassacarídeo.

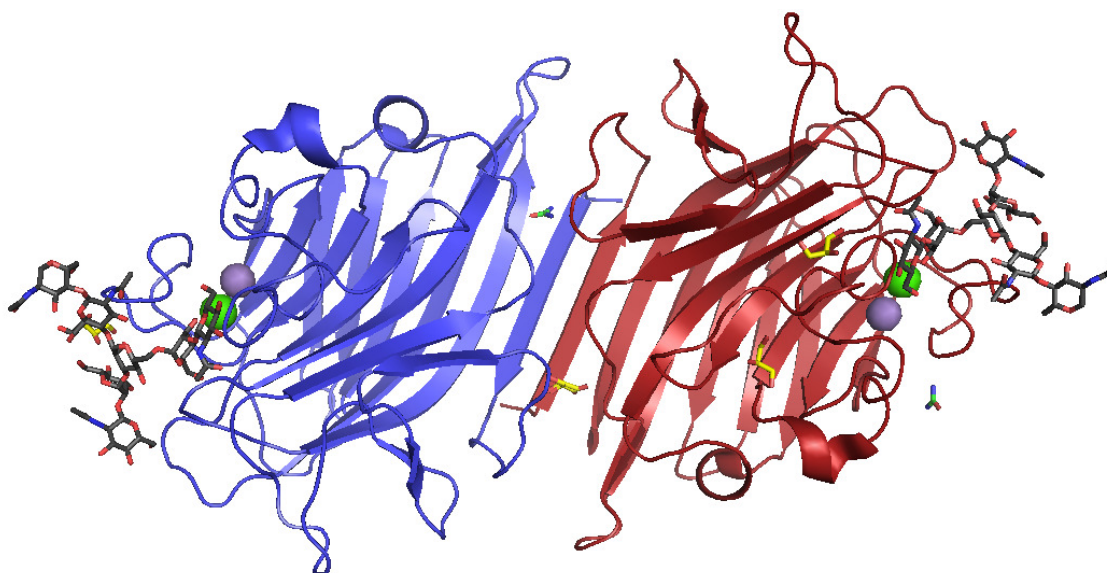


Figura 49. Visão geral da estrutura dimérica de PELA em complexo com o heptassacarídeo simétrico, formada pelos monômeros A em vermelho e B em azul. Presença dos metais Ca^{2+} e Mn^{2+} (esferas verde e roxa, respectivamente), do ligante heptassacarídeo (preto), de quatro moléculas de glicerol (amarelo) e de duas de uréia (verde).

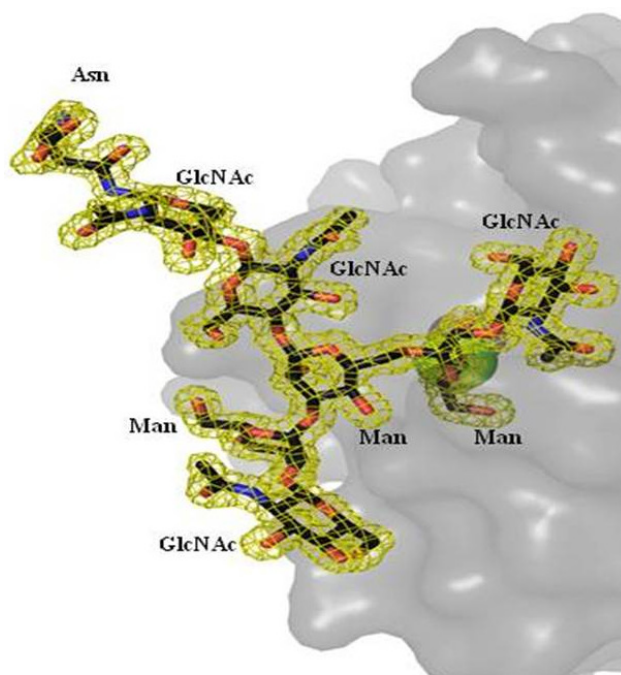


Figura 50. Representação da superfície do sítio de ligação de PELA em complexo com o heptassacarídeo. Densidade eletrônica experimental do ligante heptassacarídeo na superfície do sítio de ligação de PELA.

Entre os resíduos de aminoácidos envolvidos no sítio, estão inclusos os oito resíduos já descritos no sítio de reconhecimento ao carboidrato evidenciado no complexo monômero B com trimanose (Asp95, Ala113, Gly115, Asn145, Ser146, Asn147, Glu230, Gln231) e outros cinco

resíduos adicionais (Ser54, Ser91, Asn92, Gly111, Gly229). Adicionalmente, cinco moléculas de água estão representadas na interação, intermediando resíduos e oligossacarídeo.

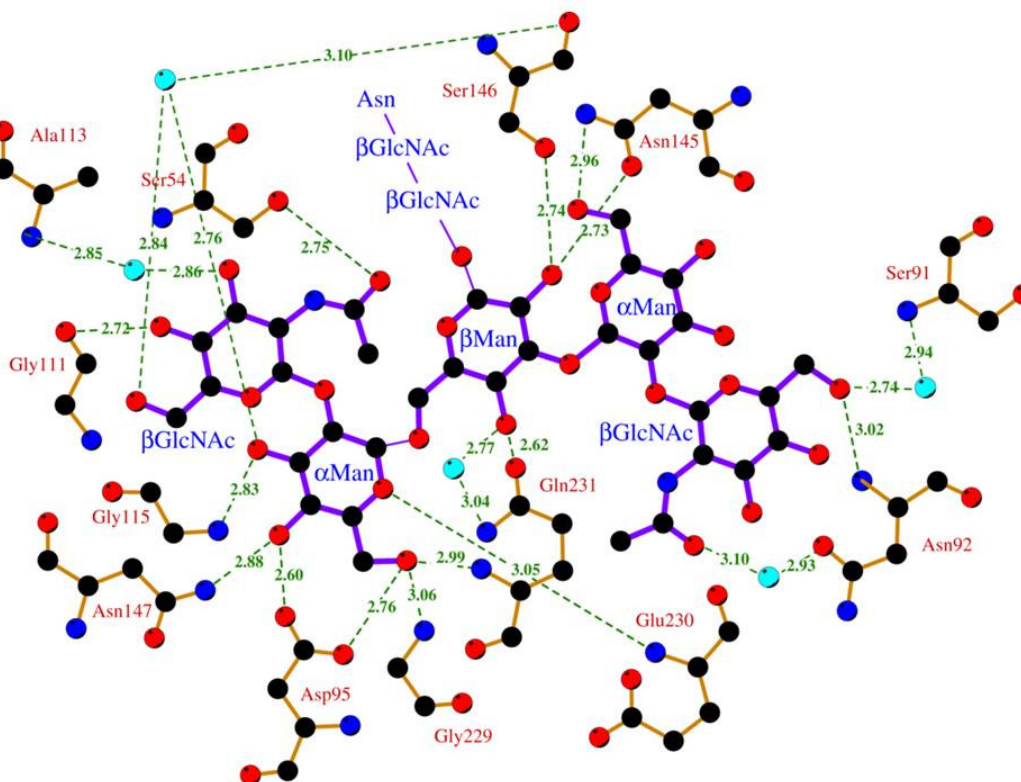


Figura 51. Representação esquemática bidimensional (LigPlot) da rede de ligações de hidrogênio entre os principais resíduos envolvidos na interação do sítio de ligação de PELA com o ligante heptassacarídeo. Comprimento das ligações em Å.

O número de ligações de hidrogênio descritas envolvendo o braço 1-3 apresenta-se em uma quantidade maior do que aquelas envolvendo o braço 1-6.

A manose do braço 1-6 está localizado no sítio de ligação principal (subsítio M) com a mesma orientação e rede de ligações de hidrogênio descritas acima para o ligante trimanosídeo. Com exceção da asparagina e do primeiro GlcNAc ligado a ela, todos os resíduos estão interagindo com a superfície da proteína, resultando em uma superfície maior de contato (Figura 52). O segundo GlcNAc do núcleo quitobiose interage com a volta C e estabelece uma ligação de hidrogênio mediada por uma molécula de água com o oxigênio da cadeia principal de Asn145. Os inúmeros outros contatos envolvem ambos os braços do glicanos. O GlcNAc na braço 1-6 estabelece três ligações de hidrogênio com aminoácidos do sítio -1 (Figura 53). A manose central no sítio 2 e o braço 1-3 nos sítios 2 e 3 também apresentam um grande número de contatos. Mais

de 15 ligações de hidrogênio diretas e 6 ligações mediadas por moléculas de água são observadas, racionalizando assim a grande afinidade medida em experimentos de microcalorimetria.

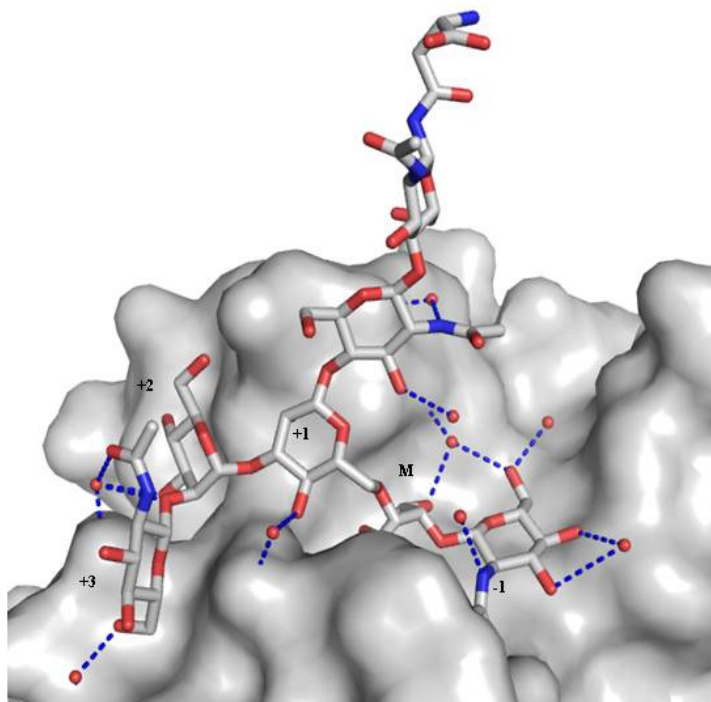


Figura 52. Estrutura cristalina do complexo PELA/heptassacarídeo. Representação do oligossacarídeo sobre a superfície da proteína.

Analisando o envolvimento das cinco voltas (A-E) descritas na arquitetura de sítios de ligação a carboidratos em lectinas de leguminosas (SHARMA & SUROLIA, 1997), o complexo PELA/heptassacarídeo apresentou diferenças em relação à interação com trimanose. A figura 53 traz a representação dessas voltas em PELA e seu envolvimento no complexo com heptassacarídeo. A participação do resíduo Ser54 nessa interação inclui a participação da volta E, até então não envolvida no sítio de reconhecimento a carboidrato de PELA (no complexo com trimanosídeo).

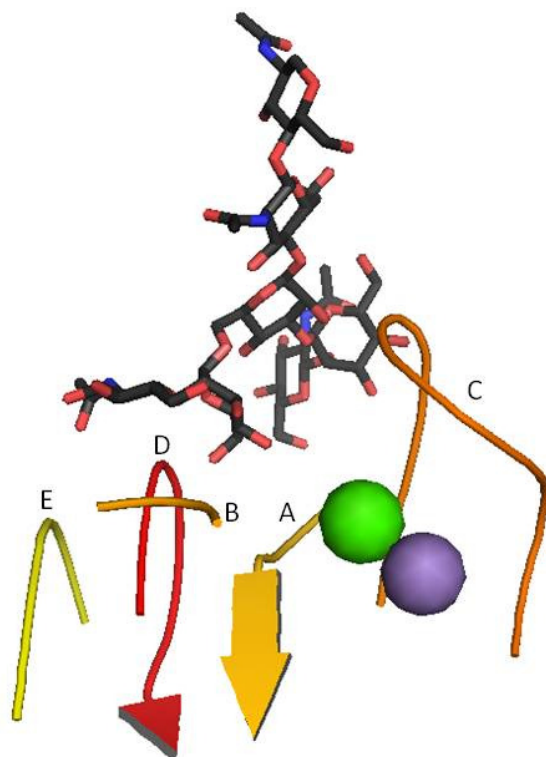


Figura 53. Representação da arquitetura do sítio de ligação a carboidrato de PELA complexada com heptassacarídeo, evidenciando as 5 voltas (A-E) típicas de lectinas de leguminosas. Nessa interação, resíduos pertencentes às cinco voltas A, B, C, D e E participaram da ligação ao carboidrato.

A volta E é encontrado interagindo com o carboidrato ligado em poucos exemplos, geralmente agindo como um subsítio. Entre lectinas de leguminosas Man/Glc ligantes, tem-se descrito a participação da volta E em complexos das lectinas ConA e PAL com Man(α 1-2)Man (MOOTHOO *et al.*, 1999, LORIS *et al.*, 2004), onde a molécula de manose redutora ocupa o sítio primário de ligação a carboidrato, e a segunda molécula de glicose interage com um subsítio, além do sítio primário, estabilizado por ligações de hidrogênio com resíduos da volta E. Em PELA, a Ser54 encontra-se estabilizando o GlcNAc da extremidade do braço 1-3, o que pode justificar uma parte da baixa entropia que dirige a interação da lectina com esse açúcar.

A conformação do glicano biantenado e a rede de contatos são muito semelhantes ao que têm sido descrito previamente para o complexo PAL/pentassacarídeo (BUTS *et al.*, 2006), embora com pequenas diferenças nas extremidades dos braços (Figura 54). O mesmo pentassacarídeo complexado com ConA rendeu a mesma posição do braço 1-6 nos subsítios M e -1, mas uma orientação diferente dos braço 1-3, novamente por causa da diferença importante na orientação conformacional da volta C (MOOTHOO & NAISMITH, 1998).

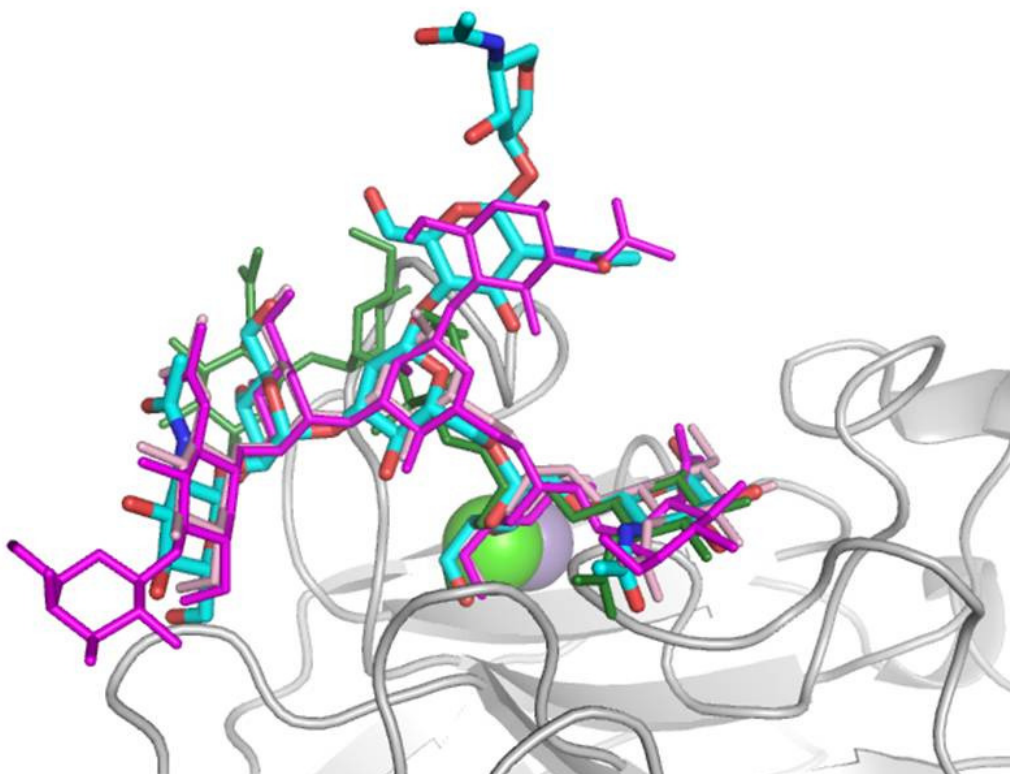


Figura 54. Comparação do complexo PELA/heptassacarídeo com outros de lectinas de leguminosas relacionadas. Superposição da cadeia B (ciano) de PELA em complexo com heptassacarídeo com complexos oligossacarídeos cristalinos de PAL (magenta: 2ARX, salmão: 2A6X) e ConA (verde: 1TEI).

Por último foi feita uma comparação da estrutura de PELA complexada com trimanose e com heptassacarídeo, através da sobreposição das duas estruturas (PDBs: 3ZVX e 3ZYR, respectivamente) (Figura 55). Nessa comparação, fica evidente que a manose do braço 1-6 do heptassacarídeo ocupa o sítio primário de ligação a carboidrato de PELA numa forma similar àquela visualizada para trimanosídeo, com uma arquitetura global de ligação semelhante. A principal diferença vem da orientação do segundo braço que aponta para o exterior no caso do trimanosídeo (para o outro monômero) e que se curva na superfície da proteína para o heptassacarídeo, estabelecendo novos contatos com o braço 1-3 α GlcNAc1-2Man.

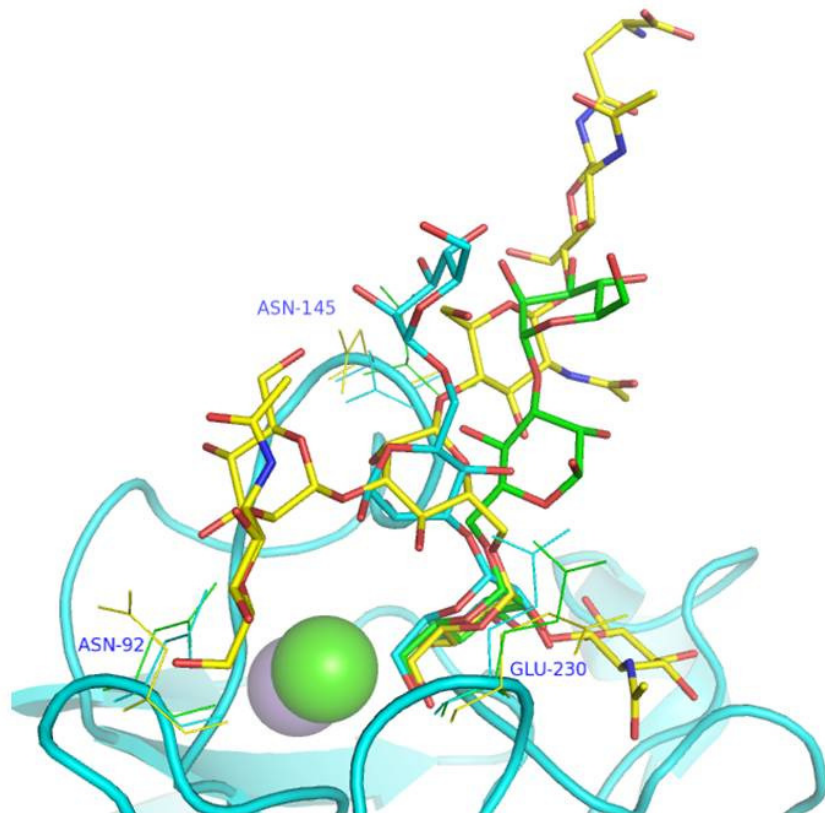


Figura 55. Comparação do reconhecimento dos ligantes trimanose e heptassacarídeo pela lectina PELA. Comparação entre os três modos de ligação estudados: trimanose ligado aos resíduos α Man1-3 (ciano) e pelos resíduos α Man1-6 (verde) e o heptassacarídeo (amarelo).

A preferência para *N*-glicanos assimétricos com um braço 1-6 curto foi demonstrada por análises semi-quantitativas de experimentos de *Glycan array* em concentrações diversas da lectina. Não foi possível obter cristais do complexo PELA e o composto assimétrico E, mas o complexo com o heptassacarídeo simétrico H fornece alguns indícios sobre essa preferência. O GlcNAc do braço 1-3 estabiliza contatos nos sítios estendidos +2 e +3 enquanto mantém uma conformação de baixa energia nas ligações envolvidas. Por outro lado, a ligação do GlcNAc no braço 1-6 no sítio -1 induz uma forte distorção conformacional com custo energético. Em adição, seu oxigênio O4 não está acessível para elongação por galactose. Oligossacarídeos com braço 1-3 longo e braço 1-6 curto seriam, portanto, energeticamente favoráveis para ligação, levando a um excelente acordo entre dados cristalográficos e *Glycan array*. Entretanto, os dados em solução de ITC não concordam perfeitamente, desde que ambos oligossacarídeos simétricos e assimétricos se ligam com uma afinidade igualmente forte à PELA. Todavia, os perfis entálpicos são diferentes; o oligossacarídeo assimétrico apresentou uma entalpia de ligação maior, o que

está de acordo com a hipótese de que um GlcNac extra no braço 1-6 não é favorável. A diferença na contribuição entrópica entre os dois oligossacarídeos é mais difícil de explicar. Entretanto, fica claro que a contribuição entrópica seria bem diferente para um oligossacarídeo apresentando seu Asn ancorado em um chip (*Glycan array*) ou glicoproteína em relação ao mesmo oligossacarídeo em solução (ITC). Esse trabalho, portanto, exemplifica um dos casos onde o estado físico-químico do ligante, ou seja, se o mesmo se apresenta em solução ou ancorado a uma superfície, pode influenciar sua afinidade pelo receptor.

CONCLUSÕES E PERSPECTIVAS

A concepção e execução desse trabalho de tese permitiu a descrição completa da clonagem e produção heteróloga do gene de uma lectina de sementes da espécie leguminosa *Platypodium elegans*, possibilitando a análise completa da especificidade dessa lectina frente a oligossacarídeos diversos, além dos aspectos estruturais e termodinâmicos envolvidos nas interações, contribuindo para o entendimento de como se dá o reconhecimento de carboidratos envolvidos nos mais distintos processos biológicos por lectinas. Este foi o primeiro relato de purificação e estudos estrutura-especificidade de uma lectina recombinante de uma espécie da subtribo Dalbergiae.

A lectina recombinante A de *P. elegans* (PELa) Man/Glc ligante, mostrou ter uma especificidade não-usual à oligossacarídeos complexos do tipo biantenado, com uma preferência sobretudo ao tipo assimétrico, com um braço 1-6 curto de manose. Esta foi a primeira caracterização de uma lectina dessa subtribo por *Glycan array*. Além disso, PELA apresentou uma afinidade considerável a esse ligante assimétrico, quando comparada às demais lectinas da mesma família, remetando sobre a relação de proximidade que essa classe de oligossacarídeo deva ter com o ligante natural da lectina em seu modo de ação (endógeno ou exógeno). Os resultados presentes nesse trabalho tornaram possível a redação de um manuscrito em língua inglesa para publicação que está sendo submetido ao *Journal of Biological Chemistry*.

O estudo dessa lectina presente em sementes de *P. elegans* e sua caracterização bioquímica fornece, portanto, informações estruturais importantes sobre as interações proteína-açúcar necessárias para compreender os mecanismos de ação, o reconhecimento, especificidade e afinidade a oligossacarídeos complexos a nível molecular. Uma vez em posse dessas informações, a sua capacidade de reconhecer glicanos envolvidos em eventos biológicos interessantes, desde a investigação de sua função no organismo de origem, como em aplicações biotecnológicas, poderá ser entendida, amplificada ou revertida.

O maior desafio agora será tentar responder à pergunta biológica sobre a importância dessa interação para a planta de origem: que eventos estão relacionados, em que etapas da vida e como pode ser melhor explorado. O conhecimento do gene e a forma recombinante da lectina, aliados ao conhecimento de como se dá sua interação a oligossacarídeos complexos, são importantes ferramentas de partida para essa abordagem.

RESUMO DE OUTROS TRABALHOS REALIZADOS NO PERÍODO DO DOUTORADO:

O trabalho envolvendo a análise estrutural, de especificidade e de afinidade da forma recombinante de PELa foi utilizado para redigir esse trabalho de tese, por reunir dados desenvolvidas pela colaboração das equipes brasileira e francesa por ocasião do estágio co-tulela, mas durante o período do doutorado, outros projetos foram executados, rendendo resultados agrupados a seguir, em ambas as universidades.

Durante todo o período do doutorado, outros projetos envolvendo lectinas vegetais foram executados em paralelo, na Universidade Federal do Ceará, gerando no total sete publicações em co-autoria. Dentre os trabalhos em fase de conclusão, encontram-se os estudos estruturais, de especificidade e investigação de efeitos biológicos em lectina de *Canavalia brasiliensis* (ConBr), lectinas manose e lactose específicas de *Cymboseman roseum* (CRLI e CRLII), lectina de *Dioclea rostrata* (DrosL) e lectina de *Dioclea wilsonii* (DwiiL). Em relação a esses sujeitos, foram geradas as seguintes publicações, disponíveis em revistas científicas indexadas, segundo descrito abaixo, além de outras seis publicações, resultado de trabalhos em co-autoria durante o período de estágio em iniciação científica e mestrado:

1. PIRES, A. F., ASSREUY A. M., LOPES, E. A., CELEDÔNIO, N. R., SOARES, C. E., RODRIGUES, N. V., SOUSA, P. L., BENEVIDES, R. G. , NAGANO, C. S., CAVADA, B. S., LEAL-CARDOSO, J. H., COELHO-DE-SOUZA, A. N., SANTOS, C. F. Opioid-like antinociceptive effects of oral administration of a lectin purified from the seeds of <i>Canavalia brasiliensis</i> . <i>Fundam. Clin. Pharmacol.</i> (Print), 2011
2. BEZERRA, E. H. SA., ROCHA, B. A. M., NAGANO, C. S., BEZERRA, G. A., MOURA, T. R., BEZERRA, M. J. B., BENEVIDES, R. G. , SAMPAIO, A. H., DELATORRE, P., CAVADA, B. S. Structural analysis of ConBr reveals molecular correlation between the carbohydrate recognition domain and nitric oxide release from endothelial cells. <i>Biochemical and Biophysical Research Communications</i> , 408 (4): 566 - 570, 2011.
3. DAM, T. K., CAVADA, B. S., NAGANO, C. S., ROCHA, B. A., BENEVIDES, R. G. , NASCIMENTO, K. S., DE SOUSA, L. A., OSCARSON, S., BREWER, C. F. Fine specificities of two lectins from <i>Cymbosema roseum</i> seeds: a lectin specific for high-mannose oligosaccharides and a lectin specific for blood group H type II trisaccharide. <i>Glycobiology</i> . 2011 Jul;21(7):925-33.
4. ROCHA, B. A. M., DELATORRE, P., OLIVEIRA, T. M., BENEVIDES, R. G. , PIRES, A. F., SOUSA, A. A. S., SOUZA, L. A. G., ASSREUY, A. M. S., DEBRAY, H., AZEVEDO, W. F., SAMPAIO, A. H., CAVADA, B. S. Structural basis for both pro- and anti-inflammatory response induced by mannose-specific legume lectin from <i>Cymbosema roseum</i> . <i>Biochimie</i> (Paris), 93: 806 - 816, 2011.
5. FIGUEIREDO, JOZI G., B., FLÁVIO S., MOTA, M. R. L., SILVESTRE, P. P., AGUIAR, C. N., BENEVIDES, R. G. , NASCIMENTO, K. S., MOURA, T. R., DAL-SECCO, D., ASSREUY, A. M. S., CAVADA, B. S. Pharmacological analysis of the neutrophil migration induced by <i>D. rostrata</i> lectin: Involvement of cytokines and nitric oxide. <i>Toxicon</i> (Oxford), 54: 736 - 744, 2009.
6. ROCHA, B. A. M., MORENO, F. B. M. B., DELATORRE, P., SOUZA, E. P., MARINHO, E. S., BENEVIDES, R. G. , RUSTIGUEL, J. K. R., SOUZA, L. A. G., NAGANO, C. S., DEBRAY, H., SAMPAIO, A. H., AZEVEDO,

W. F., CAVADA, B. S. Purification, Characterization, and Preliminary X-Ray Diffraction Analysis of a Lactose-Specific Lectin from <i>Cymbosema roseum</i> Seeds. <i>Applied Biochemistry and Biotechnology</i> , 152: 383 - 393, 2009.
7. OLIVEIRA, T. M., DELATORRE, P., ROCHA, B. A. M., SOUZA, E. P. , NASCIMENTO, K. S., BEZERRA, G. A. , MOURA, T. R., BENEVIDES, R. G. , BEZERRA, E. H. S., MORENO, F. B., CAVADA, B. S. Crystal structure of Dioclea rostrata lectin: Insights into understanding the pH-dependent dimer-tetramer equilibrium and the structural basis for carbohydrate recognition in Diocleinae lectins. <i>Journal of Structural Biology</i> , 164: 177 - 182, 2008.
8. BEZERRA, G. A., OLIVEIRA, T. M., MORENO, F. B. M. B., SOUZA, E. P., ROCHA, B. A. M., BENEVIDES, R. G. , DELATORRE, P., AZEVEDO, W. F., CAVADA, B. S. Structural analysis of Canavalia maritima and Canavalia gladiata lectins complexed with different dimannosides: new insights into the understanding of the structure-biological activity relationship in legume lectins.. <i>Journal of Structural Biology</i> . , 160: 168 - 176, 2007.
9. ALENCAR, N. M., ASSREUY, A. M., HAVT, A., BENEVIDES, R. G. , MOURA, T. R., SOUSA, R. B., RIBEIRO, R. A., CUNHA, F. Q., CAVADA, B. S. Vatairea macrocarpa (Leguminosae) lectin activates cultured macrophages to release chemotactic mediators. <i>Naunyn-Schmiedeberg's Archives of Pharmacology</i> . , 374: 275 - 282, 2007.
10. DELATORRE, P., NASCIMENTO, K. S., MELO, L. M., SOUZA, E. P., ROCHA, B. A. M., BENEVIDES, R. G. , OLIVEIRA, T. M., BEZERRA, G. A., BEZERRA, M. J. B., CUNHA, R. M. S., CUNHA, F. A. B., FREIRE, V. N., CAVADA, B. S. Crystallization and preliminary X-ray diffraction analysis of the lectin from Dioclea rostrata Benth seeds. <i>Acta Crystallographica. Series F</i> . , 62: 166/62 - 168, 2006.
11. MOTA, M. R., CRIDDLE, D. N., ALENCAR, N. M., GOMES, R. C., MEIRELES, A. V., SANTI-GADELHA, T., GADELHA, C. A., OLIVEIRA, C. C., BENEVIDES, R. G. , CAVADA, B. S., ASSREUY, A. M. Modulation of acute inflammation by a chitin-binding lectin from Araucaria angustifolia seeds via mast cells. <i>Naunyn-Schmiedeberg's Archives of Pharmacology</i> , 374: 1 - 10, 2006.
12. MORENO, F. B. M. B., BEZERRA, G. A., OLIVEIRA, T. M., SOUZA, E. P., ROCHA, B. A. M., BENEVIDES, R. G. , DELATORRE, P., CAVADA, B. S., AZEVEDO, W. F. New crystal forms of Diocleinae lectins in the presence of different dimannosides. <i>Acta Crystallographica. Series F</i> , 62: 1100 - 1103, 2006.
13. CAVADA, B. S., MARINHO, E. S., SOUZA, E. P., BENEVIDES, R. G. , DELATORRE, P., SOUZA, L. A. G., NASCIMENTO, K. S., SAMPAIO, A. H., MORENO, F. B. M. B., RUSTIGUEL, J. K. R., CANDURI, F., AZEVEDO, W. F., DEBRAY, H. Purification, partial characterization and preliminary X-ray diffraction analysis of a mannose-specific lectin from <i>Cymbosema roseum</i> seeds. <i>Acta Crystallographica. Section F</i> , 62: 235 - 237, 2006.

Durante o período de estágio em co-tutela no exterior, outro projeto em paralelo também foi executado no CERMAV-UJF, envolvendo estudos estruturais e de especificidade da lectina B de *Burkholderia cenocepacia* (BclB), um estudo envolvendo a análise individual de cada uma das duas subunidades que a compõem (denominadas N e C-terminal). Um resumo do trabalho que está em curso de conclusão e redação:

B. cepacia é um complexo (BCC) de bactérias gram-negativas que compreende várias espécies bastante similares fenotipicamente, descritas como um patógeno humano oportunista responsável por epidemias, em particular em pacientes com fibrose cística. Os gliconjugados têm um papel importante em fenômenos de reconhecimento molecular hospedeiro-patógeno. O arsenal que esses microorganismos utilizam para reconhecer, aderir e infectar seu hospedeiro compreende diversos tipos de proteínas, dentre as quais algumas reconhecem gliconjugados: as lectinas. Certas bactérias do complexo BCC, possuem em seu genoma quatro sequências

homólogas ao gene *lecA* codificado pela lectina PA-IIL em *Pseudomonas aeruginosa*. O estudo de uma dessas lectinas solúveis, BclB, foram objeto desse trabalho, com o objetivo de compreender o funcionamento e o papel dessas proteínas na infecção por esses microorganismos. A análise da sequência do genoma de *B. cenocepacia* J2315 mostra que essa proteína de 243 resíduos (25,4 kDa) é composta de dois domínios: um domínio N-terminal de 128 resíduos e de função desconhecida e um domínio C-terminal similar à lectina PA-IIL. Um trabalho de tese anterior a esse, descreveu a clonagem do domínio C-terminal e o caracterizou em relação a sua especificidade por fucose, sobretudo por trissacarídeos como Lewis a e Lewis x, de forma similar a PA-IIL, e também por H tipo 1 e H tipo 3, diferenciando-a de PA-IIL. Estudos envolvendo a clonagem do domínio N-terminal ou a proteína inteira não chegaram a ter uma boa caracterização. Nesse contexto, dando prosseguimento ao trabalho, o domínio C terminal de BclB foi cristalizado na presença dos ligantes Lewis a, H tipo 1 e H tipo 2, obtendo-se dados cristalográficos por difração em raios x de cristais em complexo com H tipos 1 e 2, que estão atualmente sendo processados e analisados. Em relação ao domínio N terminal, também foi clonado separadamente, em uma construção excluindo-se uma região C-terminal codificante para uma volta, com o objetivo facilitar a cristalização. Como se tratava de um domínio de função desconhecida, sem qualquer similaridade de sequência disponível, a proteína foi expressa em um meio rico em Selênio-Met, de tal forma que esse resíduo foi incorporado na sequência do domínio. O domínio foi cristalizado, obtendo-se dados cristalográficos por difração em raios X que permitiram verificar a incorporação do selênio na proteína. Após processamento, esse domínio pareceu apresentar um motivo similar àquele presente na cadeia leve do complexo dineína, uma proteína motora encontrada em microtúbulos, cílios ou flagelos.

Dois outros projetos foram iniciados enquadrados na colaboração, mas devida a algumas dificuldades técnicas encontradas na purificação das proteínas e/ou obtenção de material genético à partir das amostras biológicas não permitiram ir além como com PELA:

Proteína	Justificativa	Objetivo	Situação/problemas
Lectina da alga <i>Codium isthmocladum</i>	Uma pequena quantidade de lectinas puderam ser isoladas e caracterizadas à partir de algas. O laboratório brasileiro evidenciou uma atividade lectínica em extratos da alga verde <i>C. isthmocladum</i> .	-Clonagem e expressão heteróloga -Cristalização em complexo com metais pesados	Após diversas tentativas, preparou-se DNA genômico da alga. As tentativas por identificar o gene da lectina à partir de iniciadores correspondentes à sequência N-terminal se revelaram infrutuosas.
Lectina de sementes da mimosoideae <i>Albizia lebbek</i>	As mimosoideaes contém lectinas da família das jacalinas. Estudos brasileiros preliminares mostraram que extratos de sementes de <i>A. lebbek</i> (ALL) possuem uma atividade lectínica específica por galactose.	-Estudo da especificidade -Clonagem e expressão heteróloga	A lectina foi enviada ao CFC para ser testada em <i>Glycan Array</i> após marcação. Alguma interação significativa pôde ser observada, provavelmente devido a uma deterioração da lectina durante a marcação ou transporte. A clonagem foi tentada a partir de mRNA de sementes. Entretanto, não foi possível identificar o gene dessa lectina a partir dos iniciadores preparados por estratégias diversas.

REFERÊNCIAS BIBLIOGRÁFICAS

ADAMS, S. M., BLAKESLEY, R. Linear amplification DNA sequencing. *Focus*®, 13, 56, 1991.

ADAR R., STREICHER H., ROZENBLATT S., SHARON N. Synthesis of soybean agglutinin in bacterial and mammalian cells. *Eur J Biochem.* 249 (3), 684-689, 1997.

ALENCAR, N. M. N., ASSREUY, A. M. S., ALENCAR, V. B. M., MELO, S. C., RAMOS, M. V., CAVADA, B. S., CUNHA, F. Q., RIBEIRO, R. A. The galactosebinding lectin from *Vatairea macrocarpa* seeds induces in vivo neutrophil migration by indirect mechanism. *Int J Biochem Cell Biol.*, 35, 1674–1681, 2003.

ALENCAR, N. M. N., ASSREUY, A. M. S., BINDÁ, A.H., BENEVIDES, R. G., MOURA, T. R., SOUSA, R. B., RIBEIRO, R. A., CUNHA, F. Q., CAVADA, B. S. *Vatairea Macrocarpa* (Leguminosae) lectin activates cultured macrophages to release chemotactic mediators. *Naunyn-Schmiedeberg's Archives of Pharmacology*, 374, 275-282, 2005.

ALENCAR, N. M. N., ASSREUY, A. M. S., CRIDDLE, D. N., SOUZA, E. P., SOARES, P. M. G., HAVT, A., ARAGÃO, K. S., BEZERRA, D. P., RIBEIRO, R. A., CAVADA, B. S. *Vatairea macrocarpa* lectin induces paw edema with leukocyte infiltration. *Protein Peptide Letters*, 11, 195–200, 2004.

ALENCAR, N. M., CAVALCANTE, C. F., VASCONCELOS, M. P., LEITE, K. B., ARAGÃO, K. S., ASSREUY, A. M., NOGUEIRA, N. A., CAVADA, B. S., VALE, M. R. Anti-inflammatory and antimicrobial effect of lectin from *Lonchocarpus sericeus* seeds in an experimental rat model of infectious peritonitis. *J Pharm Pharmacol.* 57 (7), 919-22, 2005.

ALENCAR, N. M., TEIXEIRA, E. H., ASSREUY, A. M. S., CAVADA, B. S., FLORES, C. A., RIBEIRO, R. A. Leguminous lectins as tools for studying the role of sugar residues in leukocyte recruitment. *Mediators. Inflamm.* 8, 107–13, 1999.

ALENCAR, N.M., ASSREUY, A.M., HAVT, A., BENEVIDES, R.G., DE MOURA, T.R., DE SOUSA, R.B., RIBEIRO, R.A., CUNHA, F.Q., CAVADA, B.S. *Vatairea macrocarpa* (Leguminosae) lectin activates cultured macrophages to release chemotactic mediators. *Naunyn Schmiedebergs Arch Pharmacol.* 374 (4), 275-82, 2007.

AMBROSI, M., CAMERON, N. R., AND DAVIS, B. G. Lectins: tools for the molecular understanding of the glycode. *Org. Biomol. Chem.*, 3, 1593-1608, 2005.

ARANGO, R., ADAR, R., ROZENBLATT, S., SHARON, N. Expression of erythrina corallodendron lectin in escherichia coli. *Eur. J. Biochem.* 205, 575–581, 1992.

ARAÚJO-FILHO, J. H., VASCONCELOS, I. M., MARTINS-MIRANDA, A. S., GONDIM, D. M., OLIVEIRA, J. T. A cona-like lectin from *dioclea guianensis* benth. has antifungal activity

against colletotrichum gloeosporioides, unlike its homologues, conm and cona. *J Agric Food Chem.* 58 (7), 4090-4096, 2010.

ASSREUY, A. M. S., SHIBUYA, M. D., MARTINS, G. J., DE SOUZA, M. L. P., CAVADA, B. S., MOREIRA, R. A., OLIVEIRA, J. A., RIBEIRO, R. A., FLORES, C. A. Anti-Inflammatory Effect Of Glucose-Mannose Binding Lectins Isolated From Brazilian Beans. *Mediators of Inflammation, Inglaterra*, 6, 201-210, 1997.

ASSREUY, A. M., FONTENELE, S. R., PIRES, A. F., FERNANDES, D. C., RODRIGUES, N. V., BEZERRA, E. H., MOURA, T. R., DO NASCIMENTO, K. S., CAVADA, B. S. Vasodilator effects of diocleinae lectins from the canavalia genus. *Naunyn Schmiedebergs Arch Pharmacol.* 380 (6), 509-21, 2009.

AUSUBEL, F.M. Current Protocols in Molecular Biology, Vol. 2, *John Wiley & Sons*, New York, 1989.

BANERJEE, R., DAS, K., RAVISHANKAR, R., SUGUNA, K., SUROLIA, A., VIJAYAN, M. Conformation, protein-carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin-lactose complex. *J. Mol. Biol.*, 259, 281-296, 1996.

BARAUNA, S. C. KASTER, M. P., HECKERT, B. T., DO NASCIMENTO, K. S., ROSSI, F. M., TEIXEIRA, E. H., CAVADA, B. S., RODRIGUES, A. L., LEAL, R. B. Antidepressant-like effect of lectin from *Canavalia brasiliensis* (ConBr) administered centrally in mice. *Pharmacol. Biochem. Behavior*, 85, 160-169, 2006

BARBOSA, T., ARRUDA, S., CAVADA, B., GRANGEIRO, T. B., FREITAS, L. A. R., BARRAL-NETTO, M. *In Vivo* lymphocyte activation and apoptosis by lectins of The Diocleinae Subtribe. *Memórias do Instituto Oswaldo Cruz.* 96 (5), 673-678, 2001.

BAUSE, E., LEGLER, G. the role of the hydroxy amino acid in the triplet sequence asn-xaa-thr(ser) for the n-glycosylation step during glycoprotein biosynthesis. *Biochem J.* 195, 639-44, 1981.

BENETEAU, J., RENARD, D., MARCHÉ, L., DOUVILLE, E., LAVENANT, L., RAHBÉ, Y., DUPONT, D., VILAINE, F., DINANT, S. Binding properties of the n-acetylglucosamine and high-mannose n-glycan pp2-a1 phloem lectin in arabidopsis. *Plant Physiol.* 153 (3), 1345-61, 2010.

BENEVIDES, R. G. Avaliação do Potencial Fungicida e Termiticida de uma Fração Protéica Lectínica de Sementes de *Platypodium elegans* Vogel e Obtenção da Lectina Purificada. *Dissertação* (Mestrado em Bioquímica) – Centro de Ciências, Universidade Federal do Ceará, Fortaleza, 151 p., 2008.

BENJAMIN, C. F., FIGUEIREDO, R. C., HENRIQUES, M. G.M. O., BARJA-FIDALGO, C. Inflammatory and anti-inflammatory effects of soybean agglutinin. *Brazilian Journal of Medical and Biological Research.* 30, 873-881, 1997.

- BERCHTOLD, S., T. JONES, P. MUHL-ZURBES, D. SHEER, G. SCHULER, AND A. STEINKASSERER. The human dendritic cell marker CD83 maps to chromosome 6p23. *Ann. Hum. Genet.*, 63:181, 1999.
- BERMAN, H. M., BHAT TN, BOURNE, P. E., FENG, Z., GILLILAND, G., WEISSIG, H., WESTBROOK, J. The Protein Data Bank And The Challenge Of Structural Genomics. *Nat Struct Biol.* 7, 957-9, 2000.
- BEZERRA, D. P., LEITE, K. B., NUNES, E. P., MARINHO, E. S., CAVADA, B. S. Purificação de uma lectina presente em sementes de *Machaerium acutifolium*. *Anais da XXVI Reunião Nordestina de Botânica*, In: XXVI Reunião Nordestina de Botânica, Fortaleza, 2003.
- BEZERRA, W. M., CARVALHO, C. P. S., MOREIRA, R. A., GRANGEIRO, T. B. Establishment of a heterologous system for the expression of *Canavalia brasiliensis* lectin: a model for the study of protein splicing. *Genetics and Molecular Research*, 5 (1), 216-223, 2006.
- BLIXT, O., HEAD, S., MONDALA, T., SCANLAN, C., HUFLEJT, M.E., ALVAREZ, R., BRYAN, M.C., FAZIO, F., CALARESE, D., STEVENS, J., RAZI, N., STEVENS, D. J., SKEHEL, J. J., VAN DIE, I., BURTON, D. R., WILSON, I. A., CUMMINGS, R., BOVIN, N., WONG, C. H., PAULSON, J. C. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci USA*, 101, 17033-17038, 2004.
- BOISTELLE, R., ASTIER, J. P. Crystallization mechanisms in solution. *J. Crystal Growth*, 90, 14-30, 1988.
- BONNEIL, E., YOUNG, N. M., LIS, H., SHARON, N., THIBAUT, P. Probing genetic variation and glycoform distribution in lectins of the *Erythrina* genus by mass spectrometry. *Archives of Biochemistry and Biophysics*, 426, 241-249, 2004.
- BOTOS, I., O'KEEFE, B. R., SHENOY, S. R., CARTNER, L. K., RATNER, D. M., SEEBERGER, P. H., BOYD, M. R. AND WLODAWER, A. Structures of the complexes of a potent anti-HIV protein cyanovirin-N and high mannose oligosaccharides. *J. Biol. Chem.*, 277, 34336-34432, 2002.
- BOUCKAERT, J., DEWALLEF, Y., POORTMANS, F., WYNS, L., LORIS, R. The structural features of concanavalin A governing non-proline peptide isomerization. *Journal of Biological Chemistry*, 275, 19778–19787, 2000.
- BOUCKAERT, J., LORIS, R., POORTMANS, F., WYNS, L. Crystallographic structure of metal-free concanavalin A at 2.5 Å resolution. *Proteins Struct Funct Genet.*, 23, 510-524, 1995.
- BOURNE, Y., ROUSSEL, A., FREY, M., ROUGE, P., FONTECILLA-CAMPS, J. C., & CAMBILLAU, C. Three-dimensional structures of complexes of *Lathyrus ochrus* isolectin I with glucose and mannose: fine specificity of the monosaccharide-binding site. *Proteins: Struct. Funct. Genet.* 8, 365–376, 1990.

- BRANCO, A. T., BERNABÉ, R. B., FERREIRA, B. S., OLIVEIRA, M. V. V., GARCIA, A. B., SOUZA-FILHO, G. A. Expression and purification recombinant lectin from rice (*Oryza sativa* L.). *Protein Expression & Purification*, 33, 34-38, 2004.
- BRINDA, K. V., SUROLIA, A., VISHVESHWARA, S. Review Article: Insights into the quaternary association of proteins through structure graphs: a case study of lectins. *Biochemical Journal*, 391, 1-15, 2005.
- BROW, M. A. D. PCR Protocols: A Guide to Methods and Applications (Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J., eds.), *Academic Press*, San Diego. 189, 1990.
- BUCKLER, A. J., CHANG, D. D., GRAW, S. L., BROOK, D., HABER, D. A., SHARP, P. A., AND HOUSMAN, D. E. Proc. Natl. Acad. Sci. USA, 88, 4005, 1991.
- BUTS, L., GARCIA-PINO, A., IMBERTY, A., AMIOT, N., BOONS, G. J., BEECKMANS, S., VERSÉES, W., WYNS, L., LORIS, R. Structural basis for the recognition of complex-type biantennary oligosaccharides by *Pterocarpus angolensis* lectin. *FEBS J.*, 273, 2407-2420, 2006.
- CALVETE, J. J., SANTOS, C. F., MANN, K., GRANGEIRO, T. B., NIMTZ, M., URBANKE, C., AND SOUSA-CAVADA, B. Amino acid sequence, glycan structure, and proteolytic processing of the lectin of *Vatairea macrocarpa* seeds. *FEBS Letters*, 425, 286-292, 1998.
- CALVETE, J. J., THOLE, H. H., RAIDÁ, M., URBANKE, C., ROMERO, A., GRANGEIRO, T. B., RAMOS, M. V., ROCHA, I. M. A., GUIMARÃES, F. N., CAVADA, B. S. Molecular characterization and crystallization of Diocleinae lectins. *Biochimica et Biophysica Acta*, 1430 (2), 367-375, 1999.
- CARLINI, C. R., GROSSI-DE-SÁ, M. F. Plant toxic proteins with insecticidal properties. A review on their potentialities as bioinsecticides, *Toxicon*. 40, 1515-1539, 2002.
- CARRINGTON, D. M., AUFFRET, A., HANKE, D. E. Polypeptide ligation occurs during post-translational modification of Concanavalin A. *Nature*. 313 (3), 64-67, 1985.
- CARTER, C. W., JR., CARTER, C. W. Protein crystallization using incomplete factorial experiments. *J. Biol. Chem.* 254, 12219, 1979.
- CAVADA, B. S., BARBOSA, T., ARRUDA, S., GRANGEIRO, T. B., BARRALNETTO, M. Revisiting proteus: do minor changes in lectin structure matter in biological activity? lessons from and potential biotechnological uses of the Diocleinae subtribe lectins. *Current Protein and Peptide Science*, 2, 123-135, 2001.
- CAVADA, B. S., MARINHO, E. S., BENEVIDES, R. G., DELATORRE, P., SOUZA, L. A. G., NASCIMENTO, K. S., SAMPAIO, A. H., MORENO, F. B. M. B., AZEVEDO JR, W. F., DEBRAY, H. Purification, partial characterization and preliminary x-ray diffraction analysis of a mannose specific lectin from *Cymbosema roseum* seeds. *Acta Crystallographica. Section F*, 62, 235-237, 2006.

CAVADA, B. S., SANTOS, C. F., GRANGEIRO, T. B., NUNES, E. P., SALES, P. V. P., RAMOS, R. L. Purification and characterization of a lectin from seeds of *Vatairea macrocarpa* Ducke. *Phytochemistry*, 49, 675-680, 1998.

CHANDRA, N. R., PRABU, M. M., SUGUNA, K. & VIJAYAN, M. Structural similarity and functional diversity in proteins containing the legume lectin fold. *Protein Eng.* 14, 857-866, 2001.

CHAO, Q., CASALONGUE, C., QUINN, J. M., ETZLER, M. E. Expression and partial characterization of *Dolichos biflorus* seed lectin in *Escherichia coli*. *Archives of Biochemistry and Biophysics*, 313, 346-350, 1994.

CHEN, V. B., BRYAN, A. W., HEADD, J. J., KEEDY, D. A., IMMORMINO, R. M., KAPRAL, G. J., MURRAY, L. W., RICHARDSON, J. S., RICHARDSON, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica D*, 66, 12-21, 2010.

CHOI, J. H., LEE, S. Y. Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol.* 64, 625-635, 2004.

CIOCCI, G., MITCHELL, E. P., CAUTIER, C., WIMMEROVA, M., SUDAKEVITZ, D., PEREZ, S., GILBOA-GARBER, N. & IMBERTY, A. Structural basis of calcium and galactose recognition by the lectin PA-IL of *Pseudomonas aeruginosa*. *FEBS Lett*, 555, 297-301, 2003.

CLARK, J. M. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucl. Acids Res.* 16, 9677-86, 1988.

COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4. The CCP4 suite: programs for protein crystallography. *Acta Cryst D.*, 50(4), 760-763, 1994.

D'ALESSIO, J. M., GRUBER, C. E., CAIN, C., AND NOON, M. C. Construction of directional cDNA libraries using the SUPERScript Plasmid System; *Focus*®, 12, 47, 1990.

DAM, T. K., AND BREWER, C. F. Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.*, 102, 387-429, 2002.

DAM, T. K., GERKEN, T. A., CAVADA, B. S., NASCIMENTO, K. S., MOURA, T. R., BREWER, C. F. Binding studies of alpha-GalNAc-specific lectins to the alpha-GalNAc (Tn-antigen) form of porcine submaxillary mucin and its smaller fragments. *The Journal Of Biological Chemistry*. 282, 28256-28263, 2007.

DAM, T. K., OSCARSON, S., ROY, R., DAS, S. K., PAGE, D., MACALUSO, F. & BREWER, C. F. Thermodynamic, kinetic, and electron microscopy studies of concanavalin A and *Dioclea grandiflora* lectin cross-linked with synthetic divalent carbohydrates. *J. Biol. Chem.*, 280, 8640-8646, 2005.

DELANO, W. L. The PyMOL Molecular Graphics System. *DeLano Scientific*, San Carlos, CA, USA, 2002.

DELATORRE, P. Lectinas de leguminosas: novos enfoques estruturais para velhas moléculas. *Tese de Doutorado* apresentada ao Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, 2006.

DELATORRE, P., AND AZEVEDO, W. F. J. Simulation of electron density maps. *J. Appl. Cryst.*, 34, 658-660, 2001.

DELATORRE, P., ROCHA, B. A. M., SOUZA, E. P., OLIVEIRA, T. M., BEZERRA, G. A., MORENO, F. B., FREITAS, B. T., SANTI-GADELHA, T., SAMPAIO, A. H., AZEVEDO JR., W. F., CAVADA, B. S. Structure of a lectin from *Canavalia gladiata* seeds: new structural insights for old molecules, *BMC Struct. Biol.* 7, 52, 2007.

DELBAERE, L. T., VANDONSELAAR, M., PRASAD, L., QUAIL, J. W., WILSON, K. S., & DAUTER, Z. Structures of the lectin IV of *Griffonia simplicifolia* and its complex with the Lewis b human blood group determinant at 2.0 Å resolution. *J. Mol. Biol.* 230, 950-965, 1993.

DENNIS, J. W., GRANOVSKY, M., WARREN, C. E. Protein glycosylation in development and disease. *BioEssay.* 21, 412-21, 1999.

DENNIS, J. W., GRANOVSKY, M., WARREN, C. E.. Protein glycosylation in development and disease. *BioEssay.* 21, 412-21, 1999.

DEREWENDA, Z., YARIV, J., HELLIWELL, J. R., KALB, A. J., DODSON, E. J., PAPIZ, M. Z. The structure of the saccharide-binding site of concanavalin A. *EMBO J.* 8, 2189-2193, 1989.

DINCTURK, H. B., DUNN, A. J., JONES, D. H. Recombinant pre-pro-Concanavalin A (jack bean) is stable but of low solubility. *Journal of Biosciences (Bangalore)*, 26, 635-40, 2001.

DOWN, R. E., GATEHOUSE, A. M. R., HAMILTON, W. D. O., GATEHOUSE, J. A. Snowdrop lectin inhibits development and fecundity of the glasshouse potato aphid (*Aulacorthum solani*) when administered *in vitro* and via transgenic plants both in laboratory and glasshouse trials. *J Insect Physiol.*, 42, 1035-1045, 1996.

DUBOIS B., PEUMANS, W. J., VAN DAMME, E. J. M., VAN DAMME, J., OPDENAKKERA, G. Regulation of gelatinase B (MMP-9) in leukocytes by plant lectins. *FEBS Letters*, 427, 275-278, 1998.

ECK, J., LANGER, M., MÖCKEL, B., BAUR, A., ROTHE, M., ZINKE, H., LENTZEN, H. Cloning of the mistletoe lectin gene and characterization of the recombinant Achain. *European Journal of Biochemistry.* 264, 775-784, 1999.

ELGAVISH, S., SHAANAN, B., Structures of the *Erythrina corallodentron* lectin and of its complexes with mono- and disaccharides. *Journal of Molecular Biology*, 277(4), 917-932, 1998.

EMSLEY, P., AND COWTAN, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, 60, 2126-2132, 2004.

- ETZLER, M. E. Plant lectins. Carbohydrates in Chemistry and Biology. Part II Biology of Saccharides. ERNST, B.; HART, G. W.; SINAÏ, P. (Eds). WILEY-VCH: Weinheim. 4, 535-548, 2000
- EVANS, P. R. "Data reduction", Proceedings of CCP4 Study Weekend, *Data Collection & Processing*, 114-122, 1993.
- EWING, B., HILLIER, L., WENDL, M. C., GREEN, P. Base-calling of automated sequences traces using phred. I. Accuracy assessment, *Genome Res.*, 8, 175–185, 1998.
- FAN, E., ZHANG, Z., MINKE, W. E., HOU, Z., VERLINDE C. L. M. J., HOL, W. G. J. High-affinity pentavalent ligands of Escherichia coli heat-labile enterotoxin by modular structure-based design. *J. Am. Chem. Soc.*, 122, 2663-2664, 2000.
- FELDSTED, R., EGORIN, M., LEAVITT, R., & BAUCHUR, N. Recombination of subunits of *Phaseolus vulgaris* isolectins. *Journal of Biological Chemistry*, v. 252, p. 2967–2971. 1977.
- FIGUEIREDO, J. G., BITENCOURT, F., BESERRA, I. G., TEIXEIRA, C. S., LUZ, P. B., BEZERRA, E. H., MOTA, M. R., ASSREUY, A. M., DE QUEIROZ CUNHA, F., CAVADA, B. S., DE ALENCAR, N. M. Antinociceptive activity and toxicology of the lectin from *Canavalia boliviana* seeds in mice. *Naunyn Schmiedebergs Arch Pharmacol.* 380(5), 407-414, 2009.
- FREIRE, M. G. M., DE SOUZA, I. A., SILVA, A. C. M., MACEDO, M. L. R., LIMA, M. S., TAMASHIRO, W. M. S. C., ANTUNES, E., MARANGONI, S. Inflammatory responses induced in mice by lectin from *Talisia esculenta* seeds. *Toxicon.* 42, 275-280, 2003.
- FROHMAN, M. A. PCR Protocols: A Guide to Methods and Applications (Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J., eds.) p. 28, *Academic Press*, San Diego, 1990.
- FROHMAN, M. A., DUSH, M. K. AND MARTIN. G. R. Rapid production of fulllength cDNAs from rare transcripts by amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci.*, 85, 8998–9002, 1988.
- FROHMAN, M.A. MARTIN, G.R. Rapid amplification of cDNA ends using nested primers. *Techniques*, 1, 165–173, 1989.
- GABIUS, H. J., SIEBERT., H. C., ANDRE,S., JIMENEZ-BARBERO, J., RUDIGER, H. Chemical Biology of the Sugar Code. *Chem. Bio. Chem.* 5, 740-764, 2004.
- GABIUS, H.-J. Animal Lectins. *European Journal of Biochemistry*, 243: 543–576, 1997.
- GADELHA, C. A., MORENO, F. B., SANTI-GADELHA, T., CAJAZEIRAS, J. B., ROCHA, B. A., ASSREUY, A. M. Native crystal structure of a nitric oxide-releasing lectin from the seeds of *Canavalia maritima*. *Journal Structural Biology*, 152, 185–194, 2005.

- GALEGO DEL SOL, F., NAGANO, C. S., CAVADA, B. S., CALVETE, J. J. The First Crystal Structure of a Mimosoideae Lectin Reveals a Novel Quaternary Arrangement of a Widespread Domain. *J. Mol. Biol.*, 353, 574–583, 2005.
- GARCIA-PINO, A., BUTS, L., WYNS, L., IMBERTY, A., LORIS, R. How a plant lectin recognizes high mannose oligosaccharides. *Plant Physiol.* 144 (4), 1733-41, 2007.
- GARCIA-PINO, A., BUTS, L., WYNS, L., LORIS, R. Interplay between metal binding and cis/trans isomerization in legume lectins: structural and Thermodynamic Study of *P. angolensis* Lectin, *J. Mol. Biol.*, 361, 153–167, 2006.
- GASTEIGER E., HOOGLAND C., GATTIKER A., DUVAUD S., WILKINS M.R., APPEL R.D., BAIROCH A.; Protein Identification and Analysis Tools on the ExPASy Server;(In) John M. Walker (ed): The Proteomics Protocols Handbook, *Humana Press*, 571-607, 2005.
- GERARD, G. F, SCHMIDT, B. J., KOTEWICZ M. L., CAMPBELL, J. H. cDNA synthesis by Moloney murine leukemia virus RNase H-minus reverse transcriptase possessing full DNA polymerase activity. *Focus*, 14, 91–93, 1992.
- GERARD, G. F., D'ALESSIO, J. M., KOTEWICZ, M. L. cDNA synthesis by cloned Moloney murine leukemia virus reverse transcriptase lacking RNase H activity. *Focus*, 11, 66-69, 1989.
- GOLDSTEIN, I. J. AND PORETZ, R. D., in The Lectins: Properties, Functions and Applications in Biology and Medicine, ed. I. E. Liener, N. Sharon, and I. J. Goldstein, *Academic Press*, Orlando, 33, 1986.
- GOMES, J. C.; ROSSI, R. R.; CAVADA, B. S.; MOREIRA, R. A.; OLIVEIRA, J. T. A. Histamine release induced by glucose (mannose)-specific lectins isolated from Brazilian beans. Comparison with concanavalin A. *Agents Actions* 1994, 41, 132.
- GOUET, P., COURCELLE, E., STUART, D. I. AND METOZ, F. ESPript: multiple sequence alignments in PostScript. *Bioinformatics*. 15, 305-308, 1999.
- GRANGEIRO, T. B., FREIRE FILHO, F., CAJAZEIRAS, J. B., RAMOS, M. V., SALES, F. J. M., CAVADA, B. S. Efeito de lectinas vegetais sobre o crescimento e desenvolvimento de insetos. IN: SALES, F. J. M. (Ed.): *Leitura Entomológica*. Itaiçaba-CE. BioTools. 286p, 87-139. 1999b.
- GRANGEIRO, T. B., MONTENEGRO DE SALES, F. J., FREIRE FILHO, F. R., CAJAZEIRAS, J. B., RAMOS, M. V. and CAVADA, B. S. Lectinas de leguminosas da subtribo *Diocleinae*: proteínas com atividades inseticida e probiótica. IN: SALES, F.J.M. (Ed.): *Leitura Entomológica*. Itaiçaba-CE. BioTools., 286p, 207-246, 1999a.
- GUPTA, D., DAM, T. K., OSCARSON, S. AND BREWER, C. F. Thermodynamics of lectin carbohydrate interactions. Binding of the core trimannoside of asparagine linked carbohydrates and deoxy analogs to concanavalin A. *J. Biol. Chem.*, 272, 6388–6392, 1997.

- GUZMAN-PARTIDA, A. M., ROBLES-BURGUENO, M. R., ORTEGA-NIEBLAS, M., VAZQUEZ-MORENO, I. Purification and characterization of complex carbohydrate specific isolectins from wild legume seeds: *Acacia constricta* is (vinorama) highly homologous to *Phaseolus vulgaris* lectins. *Biochimie*. 86, 335–42, 2004.
- HAKOMORI, S. Aberrant glycosylation in cancer cell membranes as focused on glycolipid: overview and perspectives. *Cancer Res.*, 45, 2405–2414, 1985.
- HALLIDAY, J. & NAKAO, P. L. The symbiotic affinities of woody species under consideration as nitrogen-fixing trees. *NifTAL Resource Document*. Hawaii, University of Hawaii, 1982.
- HAMELRYCK, T. W., LORIS, R., BOUCKAERT, J., WYNS, L. Properties and structure of the legume lectin family. *Trends in Glycoscience and Glycobiology*, 10, 349-404, 1998.
- HAMPEL, A., LABANANSKAS, M., CORNNERS, P. G., KIRKEGARD, L., RAJBHANDARY, U. L., SIGLER, P. B., AND BOCK, R. M. Single crystals of transfer RNA from formylmethionine and phenylalanine transfer RNA's. *Science*, 162, 1384, 1968.
- HARAGUCHI, T., NOMURA K., YAGI, F. Cloning and Expression of a Mannose-Binding Jacalin-Related Lectin from Leaves of Japanese Cycad (*Cycas revoluta* Thunb.), *Biosci. Biotechnol. Biochem.*, 70, 2222-2229, 2006.
- HELENIUS, A., AEBI, M. Intracellular functions of N-linked glycans. *Science*. 291, 2364-9, 2001.
- HERNANDEZ, P., DEBRAY, H., JAEKEL, H., GARFIAS, Y., JIMENEZ, M. D. MDEL, C., MARTINEZ-CAIRO, S., ZENTENOO, E. Chemical characterization of the lectin from *Amaranthus leucocarpus* syn. *Hypocondriacus* by 2-D proteome analysis. *Glycoconj J.*, 18, 321–9, 2001.
- HERVÉ, C., SERRES, J., DABOS, P., CANUT, H., BARRE, A., ROUGÉ, P., LESCURE, B. Characterization of the Arabidopsis *lecRK-a* genes: members of a superfamily encoding putative receptors with an extracellular domain homologous to legume lectins. *Plant Mol Biol.*, 39, 671–682, 1999.
- HOFFMAN, L. M., DONALDSON; D. D. Synthesis of mitogenic phytohemagglutinin-L in *Escherichia coli*. *Biotechnology*. 5: 157-160, 1987.
- IMBERTY, A., GAUTIER, C., LESCAR, J., PEREZ, S., WYNS, L. & LORIS, R. An unusual carbohydrate binding site revealed by the structures of two *Maackia amurensis* lectins complexed with sialic acid-containing oligosaccharides. *J. Biol. Chem.* 275, 17541–17548, 2000.
- IMBERTY, A., MITCHELL, E. P., AND WIMMEROVÁ, M. Structural basis for high affinity glycan recognition by bacterial and fungal lectins. *Curr. Opin. Struct. Biol.*, 15, 525-534, 2005.
- ISLAM, B., KHAN, S. N., NAEEM, A., SHARMA, V., KHAN, A. U. Novel effect of plant lectins on the inhibition of *Streptococcus mutans* biofilm formation on saliva- coated surface. *J. Appl. Microbiol.*, 106, 1682-1689, 2009.

- JANAKIRAMAN, M. N., WHITE, C. L., LAVER, W. G., AIR G. M. AND LUO, M. Structure of influenza virus neuraminidase B/Lee/40 complexed with sialic acid and a dehydro analog at 1.8-Å resolution: implications for the catalytic mechanism. *Biochemistry*, 33, 8172-8179, 1994.
- JANAKIRAMAN, M. N., WHITE, C. L., LAVER, W. G., AIR G. M. AND LUO, M. Structure of influenza virus neuraminidase B/Lee/40 complexed with sialic acid and a dehydro analog at 1.8-Å resolution: implications for the catalytic mechanism. *Biochemistry*, 33, 8172-8179, 1994.
- JANCARIK, J, KIM, S.-H. Sparse-matrix sampling - A screening method for crystallization of proteins. *J. Appl. Crystallogr.* 24, 409-411, 1991.
- JORDAN, E. T., GOLDSTEIN, I. J.. The sequence of a second member of the lima bean lectin gene family and the expression and characterization of recombinant lectin in *Escherichia coli*. *J. Biol. Chem.*, 269 (10), 7674-7681, 1994.
- JOUBERT, F. J., SHARON, N., MERRIFIELD, E. H. Purification and properties of a lectin from *Lonchocarpus capassa* (apple-leaf) seed. *Phytochemistry*, 25 (2), 323-327, 1986.
- KANEDA, Y.; WHITTIER, R. F.; YAMANAKA, H.; CARREDANO, E.; GOTOH, M.; SOTA, H.; HASEGAWA, Y.; SHINOHARA, Y. The high specificities of *Phaseolus vulgaris* erythro- and leukoagglutinating lectins for bisecting GlcNAc or β 1-6-linked branch structures, respectively, are attributable to loop B. *Journal of Biological Chemistry*, 277 (19): 16928-16935, 2002.
- KAUR, M., SINGH, K., RUP, P.J., KAMBOJ, S.S., SAXENA, A.K., SHARMA, M., BHAGAT, M., SOOD, S.K., SINGH, J. A tuber lectin from *Arisaema jacquemontii* Blume with anti-insect and anti-proliferative properties. *Journal of Biochemistry and Molecular Biology*, 39 (4), 432-440, 2006.
- KIMBLE, B., RAMIREZ NIETO, G., PEREZ, D. R. Characterization of influenza virus sialic acid receptors in minor poultry species. *Viol. J.*, 7, 365, 2010.
- KITADA, M., KURODA, Y., DEZAWA, M. Lectins as a Tool for Detecting Neural Stem/Progenitor Cells in the Adult Mouse Brain. *Anat. Rec. Hoboken.*, 1, 1-19, 2010.
- KITOV, P. I., SADOWSKA, J. M., MULVEY, G., ARMSTRONG, G. D., LING, H., PANNUS, N. S., READ, R. J. AND BUNDLE, D. R., Shiga-like toxins neutralized by tailored multivalent carbohydrate ligands. *Nature*, 403, 669-672, 2000.
- KOCOUREK, J., AND HOREJSI, V. Defining a lectin. *Nature*, 290, 188, 1981.
- KUSUI, K.; YAMAMOTO, K.; KONAMI, Y.; OSAWA, T. cDNA cloning and expression of *Bauhinia purpurea* lectin. *Journal Biochemistry*. 109, 899-903, 1991.
- LAEMMLI, U. K., Cleavage of structural protein during the assembly of the bacteriophage t4, *Nature*, 227, 680-685, 1970.

LASKOWSKI, R. A., MACARTHUR, M. W., MOSS, D.S., THORNTON, J.M., PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Cryst.*, 26, 283-291, 1993.

LEE, Y. C., LEE, R. T., Carbohydrate-protein interactions: Basis of glycobiology, *Acc. Chem. Res.*, 28, 321, 1995.

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

Liang CH, Wu CY. Glycan array: a powerful tool for glycomics studies. *Expert Rev Proteomics.*, 6(6), 631-45, 2009.

LIS, H., SHARON, N. Lectins: Carbohydrate-specific proteins that mediate cellular recognition. *Chem. Rev.*, 98, 637-674, 1998.

LORENZI, H. Árvores brasileiras: Manual de identificação e cultivo de plantas arbóreas nativas do Brasil. *Nova Odessa: Plantarum*, 368p, 1998.

LORIS, R. Principles of structures of animal and plant lectins. *Biochim. Biophys. Acta*, 1572, 198-208, 2002.

LORIS, R., DE GREVE, H., DAO-THI, M. H., MESSENS, J., IMBERTY, A., WYNS, L. Structural basis of carbohydrate recognition by lectin II from *Ulex europaeus*, a protein with a promiscuous carbohydrate-binding site. *J. Mol. Biol.* 301, 987–1002, 2000.

LORIS, R., HAMELRYCK, T., BOUCKAERT, J., WYNS, L. Legume lectin structure. *Biochimica et Biophysica Acta*, 1383, 9-36, 1998.

LORIS, R., IMBERTY, A., BEECKMANS, S., VAN DRIESSCHE, E., READ, J. S., BOUCKAERT, J. Crystal structure of *Pterocarpus angolensis* lectin in complex with glucose, sucrose and turanose. *J. Biol. Chem.*, 278, 16297–16303, 2003.

LORIS, R.; VAN WALLE, I.; DE GREVE, H.; BEECKMANS, S.; DEBOECK, F.; WYNS, L.; BOUCKAERT, J. Structural basis of oligomannose recognition by the *Pterocarpus angolensis* seed lectin. *Journal of Molecular Biology*, 335, 1227. 2004.

LUO, S., ZHANGSUN, D., TANG, K. Functional GNA expressed in *Escherichia coli* with high efficiency and its effect on *Ceratovacuna lanigera* Zehntner. *Applied Genetics and Molecular Biotechnology*. 69, 184-191, 2005.

LUTOMSKI, D., IMAM-SGHIOUAR, N., BLONDEAU, K., CARON, M., JOUBERTCARON, R. Purification of human galectin-1 produced in high-cell density cultures of recombinant *Escherichia coli*: a comparison with classic shake flask cultivation. *Journal of Chromatography B*, 808, 105-109, 2004.

MACEDO, M. L. R., FREIRE, M. G. M., SILVA, M. B. R., COELHO, L. C. B. B. Insecticidal action of *Bauhinia monandra* leaf lectin (BmoLL) against *Anagasta kuehniella* (Lepitoptera:

Pyralidae), *Zabrotes fubsasciatus* and *Callosobruchus maculatus* (Coleoptera: Bruchidae). *Comp. Biochem. Physiol.*, 146, 486–498, 2007.

MACEDO, M. L. R.; FREIRE, M. G. M. & CASTRO, M. M., Mechanisms of the insecticidal action of TEL (*Talisia esculenta* lectin) against *Callosobruchus maculatus* (Coleoptera: Bruchidae). *Archives of Insect Biochemistry and Physiology*. 56 (2), 84–96, 2004.

MAMMEN, M., CHOI S. K. WHITESIDES, G. M. Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. *Angew. Chem., Int. Ed. Engl.*, 37, 2754–2794, 1998.

MANDAL, D. K., KISHORE, N. & BREWER. C. F. Thermodynamics of lectin-carbohydrate interactions. Titration microcalorimetry: Measurements of the binding of N-linked carbohydrates and oval buniin to concanavalin A, *Biochemistry*, 33, 1149–1156, 1994.

MANN, K., FARIAS, C. M. S. A., GALLEGRO DEL SOL, F., SANTOS, C. F., GRANGEIRO, T. B., NAGANO, C. S., CAVADA, B. S., CALVETE, J. J. The amino-acid sequence of the glucose/ mannose-specific lectin isolated from *Parkia platycephala* seeds reveal three tandemly arranged jacalin related domains. *Eur. J. Biochem.*, 218, 4414–4422, 2001.

MANOJ, N. SUGUNA, K. Signature of quaternary structure in the sequences of legume lectins. *Protein Eng.*, 14, 735–745, 2001.

MARONDEDZE, C., CHIDZWONDO, F., SITHOLE-NIANG, I., READ, J. S. Pterocarpus rotundifolius lectin gene. Submitted to the EMBL/GenBank/DDBJ databases, 2004.

MAROTTE, K., PRÉ VILLE, C., SABIN, C., MOUMÉ-PYMBOCK, M., IMBERTY, A., ROY, R. Synthesis and binding properties of divalent and trivalent clusters of the Lewis a disaccharide moiety to *Pseudomonas aeruginosa* lectin PA-III. *Org. Biomol. Chem.*, 5, 2953–2961, 2007b.

MAROTTE, K., SABIN, C., PRÉ VILLE, C., MOUMÉ-PYMBOCK, M., WIMMEROVA, M., MITCHELL, E.P., IMBERTY, A., AND ROY, R. X-ray structures and thermodynamics of the interaction of PA-III from *Pseudomonas aeruginosa* with disaccharide derivatives. *Chem. Med. Chem.*, 2, 1328–1338, 2007a.

MARTINS, A.M., MONTEIRO, A.M., HAVT, A., BARBOSA, P.S., SOARES, T.F., EVANGELISTA, J.S., DE MENEZES, D.B., FONTELES, M.C., TEIXEIRA, E.H., PINTO, V.P., NASCIMENTO, K.S., ALENCAR, N.M., CAVADA, B.S. Renal effects induced by the lectin from *Vatairea macrocarpa* seeds. *J Pharm Pharmacol.*, 57, 1329–1333, 2005.

MCCOY, A. J., STORONI, L. C. & READ, R. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Cryst. D*, 60, 1220–1228, 2004.

MCREE, D.E. Practical Protein Crystallography. *Academic Press*, inc. USA. 386 pp. 1999.

MEAGHER, J. L., WINTER, H. C., EZELL, P., GOLDSTEIN, I. J., STUCKEY, J. A. Crystal structure of banana lectin reveals a novel second sugar binding site. *Glycobiology*. 15(10),1033–42, 2005.

- MERGULHÃO, F. J. M.; SUMMERSB, D. K.; MONTEIROA, G. A. Recombinant protein secretion in *Escherichia coli*. *Biotechnology Advances*, 23,177–202, 2005.
- MERRITT, E. A., HOL, W. G. J. AB5 toxins. *Curr. Opin. Struct. Biol.*, 5, 165-171, 1995.
- MEZEI, L. M. STORTS, D. R. Purification of PCR products. In: PCR Technology: Current Innovations, Griffin, H.G. and Griffin, A.M., eds., *CRC Press*, Boca Raton, FL, 21, 1994.
- MIN W., JONES, D.H. In vitro splicing of concanavalin A is catalyzed by an aspariginil endopeptidase. *Nat Struct. Biol.*, 1(8), 502-504, 1994.
- MIN, A. P., JONES, D. H. Stability and detection of recombinant pre-pro-concanavalin A after cytoplasmatic expression in *Escherichia coli*. *FEBS Lett.*, 301-318, 1992.
- MIN, W.; DUNN, D. J.; JONES, D. H. Non-glycosilated recombinant pro-concanavalin A is active without polypeptide cleavage. *Embo. J.*, 11(4), 1303-1307.1992.
- MITCHELL, E., HOULES, C., SUDAKEVITZ, D., WIMMEROVA, M., GAUTIER, C., PEREZ, S., WU, A. M., GILBOA-GARBER, N. & IMBERTY, A. Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lung of cystic fibrosis patients. *Nat Struct Biol.*, 9, 918–921, 2002.
- MITCHELL, S. W. Researches upon the venom of the rattlesnake. *Smithson Contrib Knowl.*, XII, 89-90, 1860.
- MOLEMA, G., MEIJER, D. K. F., Targeting of drugs to various blood cell types using (neo)glycoproteins, antibodies and other protein carriers. *Adv. Drug Delivery Rev.*, 14(1), 25-50, 1994.
- MOOTHOO, D. N., CANAN, B., FIELD, R. A. & NAISMITH, J. H. Man(α 1-2)Man α -OMe-concanavalin A complex reveals a balance of forces involved in carbohydrate recognition. *Glycobiology*, 9, 539–545, 1999.
- MOREIRA, R. A., PERRONE, J. C. Purification and partial characterization of a lectin from *Phaseolus vulgaris*. *Plant Physiol.*, 59, 783-787, 1977.
- MORENO, F. B. M. B., DE OLIVEIRA, T. M., MARTIL, D. E., VICOTI, M. M., BEZERRA, G. A., ABREGO, J. R. B., CAVADA, B. S., AZEVEDO, W. F. JR. Identification of a new quaternary association for legume lectins. *J Struct Biol.*, 161,133–143, 2008.
- MORENO, F. B., BEZERRA, G. A., DE OLIVEIRA, T. M., DE SOUZA, E. P., DA ROCHA, B. A., BENEVIDES, R. G., DELATORRE, P., CAVADA, B. S., DE AZEVEDO, W. F. JR. New crystal forms of Diocleinae lectins in the presence of different dimannosides. *Acta Crystallogr Sect F Struct Biol Cryst Commun.*, 62(11), 1100-3, 2006.
- MURDOCK, L. L., HUESING, J. E., NIELSEN, S. S., PRATT, R. C., SHADE, R. E. Biological effects of plant lectins on the cowpea weevil. *Phytochemistry*, 29, 85-89, 1990.

- MURSHUDOV, G. N., VAGIN, A. A, DODSON, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallgr. D Biol. Crystallogr.* 1 (53), 240-55, 1997.
- NAGANO, C. S., CALVETE, J. J., BARETTINO, D., PÉREZ, A., CAVADA, B. S., SANZ, L. Insights in to the structural basis of the pH-dependent dimer–tetramer Equilibrium through crystallographic analysis of recombinant Diocleinae lectins. *Biochem. J.*, 409, 417 – 428, 2008.
- NAKAMURA, S., IKEGAMI, A., MATSUMURA, Y., NAKANISHI, T., NOMURA, K. Molecular Cloning and Expression of the Mannose/Glucose Specific Lectin from *Castanea crenata* Cotyledons1. *J. Biochem.*, 131, 241-246, 2002.
- NAPIMOGA, M.H., CAVADA, B.S., ALENCAR, N.M., MOTA, M.L., BITTENCOURT, F.S., ALVES-FILHO, J.C., GRESPAN, R., GONÇALVES, R.B., CLEMENTE-NAPIMOGA, J.T., DE FREITAS, A., PARADA, C.A., FERREIRA, S.H., CUNHA, F.Q. *Lonchocarpus sericeus* lectin decreases leukocyte migration and mechanical hypernociception by inhibiting cytokine and chemokines production. *Int Immunopharmacol.*, 7 (6), 824–835, 2007.
- NASCIMENTO, K. S.; GRANGEIRO, T. B.; CAVADA, B. S.; SAMPAIO, A. H. Composição mínima, teores proteicos e detecção de atividade hemaglutinante em sementes de *Andira retusa* (Leguminosae, Papilionoidae, Dalbergieae); *Resumos do 53o Congresso Nacional de Botânica; 53o Congresso Nacional de Botânica; Recife; BRASIL, 2002.*
- NELSON, D.L. & COX, M.M. Principles of Biochemistry, Fifth Edition (*Freeman Publishers*), New York, 2009.
- NELSON, D.L. & COX, M.M. Principles of Biochemistry, Fourth Edition (*Freeman Publishers*), New York, 2004.
- NEWTON, C.R., GRAHAM, A. PCR, part 1: Basic principles and methods. In: PCR, *BIOS Scientific Publishers, Ltd.*, Oxford, UK, 13, 1994.
- NIELSEN, H., ENGELBRECHT, J., BRUNAK, S., VON HEIJNE, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.*, 10, 1–6, 1997.
- NISHIGUCHI, M., YOSHIDA, K., SUMIZONO, T., TAZAKI, K. Studies by site-directed mutagenesis of the carbohydrate-binding properties of a bark lectin from *Robinia pseudoacacia*. *FEBS Lett.* 403(3), 294-298, 1997.
- NOGUEIRA, N. A.; GRANGEIRO, M. B.; CUNHA, R. M.; RAMOS, M. V.; ALVES, M. A.; TEIXEIRA, E. H.; BARRAL-NETTO, M.; CALVETE, J. J.; CAVADA B. S.; GRANGEIRO, T. B. Expression and purification of the recombinant ConBr (*Canavalia brasiliensis* lectin) produced in *Escherichia coli*. *Protein Pet Lett.*, 9(1), 59-66, 2002.
- NUNES, B. S., RENSONNET, N. S., DAL-SECCO, D., VIEIRA, S. M., CAVADA, B. S., TEIXEIRA, E. H., MOURA, T. R., TEIXEIRA, C. S., CLEMENTE-NAPIMOGA, J. T., CUNHA, F. Q., NAPIMOGA, M. H. Lectin extracted from *Canavalia grandiflora* seeds presents

potential anti-inflammatory and analgesic effects. *Naunyn Schmiedebergs Arch Pharmacol.*, 379(6), 609-616, 2009.

OGILVIE, M. L., GARTNER, T. K. Identification of snake venoms. *J. Herpetol.*, 19: 285-290, 1984.

Ohara, O., Dorit, R.L., and Gilbert, W. One-sided polymerase chain reaction: the amplification of cDNA. *Proc. Natl. Acad. Sci.*, 86, 5673, 1989.

OHIZUMI, Y., GAIDAMASHVILI, M., OHWADA, S., MATSUDA, K., KOMINAMI, J., NAKAMURA-TSURUTA, S., HIRABAYASHI, J., NAGANUMA, T., OGAWA, T., MURAMOTO, K. Mannose-binding lectin from yam (*Dioscorea batatas*) tubers with insecticidal properties against *Helicoverpa armigera* (Lepidoptera: Noctuidae). *J Agric Food Chem.* 57 (7), 2896-902, 2009.

OLIVEIRA, C., COSTA, S., TEIXEIRA, J. A., DOMINGUES, L. cDNA Cloning and Functional Expression of the α -d-Galactose-Binding Lectin Frutalin in *Escherichia coli*. *Molecular Biotechnology.* 43 (3), 212-220, 2009.

PEREIRA-JUNIOR, F.N. Purificação e Caracterização Parcial de uma Lectina Isolada de Sementes de *Platymiscium floribundum* Vogel. *Monografia de Graduação*, URCA, Crato, 45 p., 2008.

PEUMANS, W. J., VAN DAMME, E. J. M. Prevalence, biological activity and genetic manipulation of lectins in foods. *Trends Food Science Technology*, 7 (4), 132-138, 1996.

PEUMANS, W. J., VAN DAMME, E. J. M., Lectins as plant defense proteins. *Plant Physiol.*, 109, 347-352, 1995.

PEUMANS, W. J., WINTER, H. C., BEMER, V., VAN LEUVEN, F., GOLDSTEIN, I. J., TRUÇA-BACHI, P., VAN DAMME, E. J. M. Isolation of a novel plant lectin with an unusual specificity from *Calystegia sepium*. *Glycoconjugate J.*, 14, 259-265, 1997.

PEUMANS, W. J.; VAN DAMME, E. J. M. Plant Lectins: versatile proteins with important perspectives in biotechnology. *Biotechnology and Genetic Engineering Reviews*, 15, 199-227, 1998.

PEUMANS, W.J., VAN DAMME, E.J., BARRE, A., ROUGÉ, P. Classification of plant lectins in families of structurally and evolutionary related proteins. *Adv. Exp. Med. Biol.*, 491, 27-54, 2001.

PIRES, A. F.; RODRIGUES, N. V. F. C.; SIQUEIRA, R. C. L.; SILVA, M. S. J.; MARINHO, M. M.; CAVADA, B. S.; ALENCAR, N. M. N.; ASSREUY, A. M. S. Atividade Antiinflamatória Da Lectina De Sementes De *Lonchocarpus araripensis*. 3ª Reunião Regional *FESBE*, Fortaleza, CE, 2008.

PRAKASHKUMAR, R., PUSHPANGADAN, P., VIJAYAKUMAR, T. Search for lectins in seeds of tropical trees of Kerala, India. *Biol. Planta*, 40(1), 155-158, 1998.

PRASTHOFER, T., PHILLIPS, S. R., SUDDATH, F. L., ENGLER, J. A. Design, expression, and crystallization of recombinant lectin from the garden pea (*Pisum sativum*). *J. Biol. Chem.*, 264 (12), 6793-6796, 1989.

PUSTZAI, A. *Plant Lectins. Cambridge University Press. Cambridge, 1991.*

RAEMAEEKERS, J. M. R., MURO, L., GATEHOUSE, J. A., FORDHAM-SKELTON, A. P. Functional phytohemagglutinin (PHA) and *Galanthus nivalis* agglutinin (GNA) expressed in *Pichia pastoris*. Correct N-terminal processing and secretion of heterologous proteins expressed using the PHA-E signal peptide. *European Journal of Biochemistry*, 265, 394-403, 1999.

RAIKHEL, N. V., LEE, H.-I., AND BROEKAERT, W. F.. Structure and function of chitin-binding proteins. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 44, 591–615, 1993.

RAMOS, M. V., BOMFIM, L. R., CAVADA, B. S., ALENCAR, N. M. N., GRANGEIRO, T. B., DEBRAY, H. Further characterization of the glycan-binding specificity of the seed lectin from *Vaitarea macrocarpa* and its dependence of pH. *Protein Peptide Letters*, 7, 241–248, 2000.

RANGEL, T. B. A., CARVALHO, A. U., NASCIMENTO, K. S., CAVADA, B. S., FERNANDES, P. M. B. Isolamento e caracterização parcial de lectina proveniente de sementes da espécie vegetal *Andira fraxinifolia* (LEGUMINOSAE-FABACEAE). Anais/Resumos da 61ª Reunião Anual da SBPC-ISSN 2176-1221, 2009.

RATANAPO, S., NGAMJUNYAPORN, W., CHULAVATNATOL. M. Interaction of a mulberry leaf lectin with a phytopathogenic bacterium, *P. syringae pv mori*. *Plant Sci.*, 160, 739–744, 2001.

RINI, J. M., HARDMAN, K. D., EINSPAHR, H., SUDDATH, F. L. & CARVER, J. P. X-ray crystal structure of a pea lectin-trimannoside complex at 2.6 Å resolution. *J. Biol. Chem.*, 268, 10126–10132, 1993.

ROBLES, J., DOERS, M. pGEM®-T Vector Systems troubleshooting guide. *Promega Notes*, 45, 19–20, 1994.

ROZWARSKI, D. A., SWAMI, B. M., BREWER, C. F. & SACCHETTINI, J. C. Crystal structure of the lectin from *Dioclea grandiflora* complexed with core trimannoside of asparagine-linked carbohydrates. *J. Biol. Chem.* 273, 32818–32825, 1998.

RUDIGER, H. On the physiological role of plant lectins. *Bioscience*, 34, 95–99, 1984.

RUDIGER, H. On the physiological role of plant lectins. *Bioscience*, 34, 95–99, 1984.

RUDIGER, H. Structure and function of plant lectin. *Glycosciences: status and perspectiva*. ed. GABIUS, H., J. & GABIUS, S., Published by *Chaoman & Hall GmbH*, Weinheim, Germany, 415-438, 1997.

RUDIGER, H., AND GABIUS, H.J. Plant lectins: Occurrence, biochemistry, functions and application. *Glycoconj. J.*, 18, 589-613, 2001.

RUPP, B. Biomolecular crystallography: Principles, practice, and application to structural biology, *Garland*, 2006.

SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G.T., MULLIS, K. B., ERLICH, H. A. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239, 487-491, 1988.

SAIKI, R. K., SCHARF, S., FALOONA, F., MULLIS, K. B., HORN, G. T., ERLICH, H. A., AND ARNHEIM, N. Enzymatic amplification of β -globin sequences and restriction site analysis for diagnosis of sickle cell anemia, *Science* 230, 1350-1354, 1985.

SAMBROOK, J. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press*, 2222 p., 2002.

SAMPIETRO, A. R., ISLA, M. A. QUIROGA, E. N., VATTUONE, M. A. An N-acetylglucosamine oligomer binding agglutinin (lectin) from ripe *Cyphomandra betacea* Sendt. Fruits. *Plant Science.*, 160, 659-67, 2001.

SANTOS, A. F., CAVADA, B. S., ROCHA, B. A., NASCIMENTO, K. S., SANT'ANA, A. E. Toxicity of some glucose/mannose-binding lectins to *Biomphalaria glabrata* and *Artemia salina*. *Bioresour Technol.* 101 (2), 794-798, 2010.

SANZ-APARICIO, J., HERMOSO, J., GRANGEIRO, T. B., CALVETE, J. J., CAVADA, B. S. The crystal structure of *Canavalia brasiliensis* lectin suggests a correlation between its quaternary conformation and its distinct biological properties from concanavalin A. *FEBS Letters*, 405, 114-118, 1997.

SARKAR, M., WU, A. M., AND KABAT, E. A. Immunochemical studies on the carbohydrate specificity of *Maclura pomifera* lectin. *Arch. Biochem. Biophys.*, 209, 204-218, 1981.

SASTRY, M. V. K., BANERJEE, P., PATANJALI, S. R., SWAMY, M. J., SWARNALATHA, G. V., SUROLIA, A. Analysis of the saccharide binding to *Artocarpus integrifolia* lectin reveals specific recognition of T-antigen (b-D Gal(1,3)DGalNAc). *J. Biol. Chem.*, 261, 11726-11733, 1986.

SCHAFFER, C., GRANINGER, M., MESSNER, P. Prokaryotic glycosylation. *Proteomics.*, 1, 248-61, 2001.

SCHORR, M., SPONHOLZ, N., PORTILHO, I. S. Desenvolvimento sustentável para o Vale do Ribeira. <http://www.planetaorganico.com.br/TrabSchorr.htm>, 09 Ago. 2011.

SHAANAN, B., LIS, H., SHARON, N. Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose. *Science*, 254 (5033), 862-6, 1991.

SHARMA V., SUROLIA A. Cloning by genomic PCR and production of peanut agglutinin in *Escherichia coli*. *Gene*, 148, 299-304, 1994.

- SHARMA, V., SUROLIA, A. Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J Mol Biol*, 267, 433–445, 1997.
- SHARON, N. & LIS, H. Legume Lectins – A large family of homologous proteins. *In FASEB Journal*, 4, 3198-3208, 1990.
- Sharon, N. and Lis, H. Lectins as cell recognition molecules. *Science*, 46(4927), 227- 234, 1989.
- SHARON, N. Lectins: Carbohydrate-specific Reagents and Biological Recognition Molecules. *The Journal of Biological Chemistry*, 282(5), 2753-2764, 2007.
- SHARON, N., AND LIS, H. Lectin's, 2nd Ed., *Kluwer Academic Publishers*, Dordrecht, The Netherlands, 2003.
- SHARON, N., LIS, H. The structural basis for carbohydrate recognition by lectins. *Adv. Exp. Med. Biol.*, 491, 1–16, 2001
- SHARON, N.; LIS, H. Carbohydrates in cell recognition. *Scientific American*, 74-80, 1993.
- SINGH, J., SINGH, J., KAMBOJ, S. S. A novel mitogenic and antiproliferative lectin from a wild cobra lily, *Arisaema flavum*. *Biochemical and Biophysical Research Communications*, 318 (4), 1057-1065, 2004.
- SØRENSEN, H. P.; MORTENSEN, K. K. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial Cell Factories*, 4, 1. 2005(b).
- SØRENSEN, H. P.; MORTENSEN, K. K.. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology*, 115, 113–128, 2005(a).
- SRINIVAS, V. R., REDDY, G. B., AHMAD, N., SWAMINATHAN, C. P., MITRA, N., SUROLIA, A. Legume lectin family, the ‘natural mutants of the quaternary state’, provide insights into the relationship between protein stability and oligomerization. *Biochim. Biophys. Acta*, 1527, 102–111, 2001.
- SRINIVASAN, N., RUFINO, S. D., PEPYS, M. B., WOOD, S. P., BLUNDELL, T. L. A superfamily of proteins with the lectin fold. *Chemrracs-Biochem Mol Biol.*, 6, 149-164, 1996.
- STANCOMBE, P. R.; ALEXANDER, F. C. G.; LING, R.; MATHENSON, M. A.; CHONE, C. C.; CHADDOCK, J. A. Isolation of the gene and large-scale expression and purification of recombinant *Erythrina cristagalli* lectin. *Protein expression and purification*. 30, 283-292. 2003.
- STREICHER, H. R. G.; SHARON, N. Recombinant Plant Lectins and Their Mutants. *Methods in enzymology*. 363, 47-77, 2003.
- STUBBS, M. E., CARVER, J. P., DUNN, R. J. Production of pea lectin in *Escherichia coli*. *J. Biol. Chem.*, 261 (14), 1986.

SUROLIA, A., SHARON, N., SCHWARZ, F. P. Thermodynamics of monosaccharide and disaccharide binding to Erythrina corallodendron lectin. *J. Biol. Chem.*, 271, 17697–17703, 1996.

SUSEELAN, K. N.; MITRA, R.; PANDEY, R.; SAINIS, K. B.; KRISHNA, T. G. Purification and characterization of a lectin from wild sunflower (*Helianthus tuberosus* L.) tubers. *Archives of Biochemistry and Biophysics*, 407, 241-247, 2002.

TAIPALENSUU, J., ERIKSSON, S., RASK, L. The myrosinase-binding protein from *Brassica napus* seeds possesses lectin activity and has a highly similar vegetatively expressed wound-inducible counterpart. *Eur J Biochem.*, 250, 680–688, 1997.

TAYLOR, M., DRICKAMER, K. Introduction to Glycobiology. *Oxford University Press*, 2003.

TEIXEIRA, E. H.; NAPIMOGA, M. H.; CARNEIRO, V. A.; DE OLIVEIRA, T. M.; CUNHA, R. M.; HAVT, A.; MARTINS, J. L.; PINTO, V. P.; GONÇALVES, R. B.; CAVADA, B. S. In vitro inhibition of Streptococci binding to enamel acquired pellicle by plant lectins. *J Appl Microbiol.*, 101, 111-116, 2006.

THAMOTHARAN S, KARTHIKEYAN T, KULKARNI KA, SHETTY KN, SUROLIA A, VIJAYAN M, SUGUNA K. Modification of the sugar specificity of a plant lectin: structural studies on a point mutant of Erythrina corallodendron lectin. *Acta Crystallogr D Biol Crystallogr.*, 67 (3), 218-27, 2011.

TOONE, E. J. Structure and Energetics of Protein-Carbohydrate Complexes, edited by Bundle, D.R., Fiezi, T., *Current Opinion in Structural Biology*, 4, 719-728, 1994.

TURTON, K., NATESH, R., THIYAGARAJAN, N., CHADDOCK, J. A., ACHARYA, K. R. Crystal structures of Erythrina cristagalli lectin with bound N-linked oligosaccharide and lactose. *Glycobiology*, 14 (10), 923-9, 2004.

UPADHYAY, S. K., SAURABH, S., RAI, P., SINGH R., CHANDRASHEKAR, K., VERMA, P. C., SINGH, P. K., TULI. R. SUMO fusion facilitates expression and purification of garlic leaf lectin but modifies some of its properties. *J Biotechnol.*, 146, 1–8, 2010.

VAN DAMME, E. J. M., PEUMANS, W. J., BARRE, A. & ROUGÉ, P. Plant lectin: A composite of several distinct families of structural and evolutionary related proteins with diverse biological roles. *Crit. Ver. Plan Sciences*, 17(6), 575-692, 1998.

VAN DAMME, E. J., BARRE, A., SMEETS, K., TORREKENS, S., VAN LEUVEN, F., ROUGE, P. The bark of Robinia pseudoaccacia contains a complex mixture of lectins. *Plant Physiol.*, 107, 833–43, 1995.

VAN DAMME, E. J., NAKAMURA-TSURUTA, S., HIRABAYASHI, J., ROUGE, P., PEUMANS, W. J. The Sclerotinia sclerotiorum agglutinin represents a novel family of fungal lectins remotely related to the Clostridium botulinum non-toxin haemagglutinin HA33/A. *Glycoconjugate J.*, 24, 143-156, 2007.

- VAN EIJSDEN, R. R., DIAZ, C. L., DE PATER, B. S., KIJNE, J. W. Sugarbinding activity of pea (*Pisum sativum*) lectin is essential for heterologous infection of transgenic white clover hairy roots by *Rhizobium leguminosarum* biovar viciae. *Plant Mol Biol.*, 29, 431-439, 1995.
- VAN EIJSDEN, R. R., HOEDEMAEKER, F. J., DIAZ, C. L., LUGTENBERG, B. J., DE PATER, B. S. AND KIJNE, J. W. Mutational analysis of pea lectin. Substitution of Asn125 for Asp in the monosaccharide-binding site eliminates mannose/glucose-binding activity. *Plant Mol. Biol.*, 20 (6), 1049-1058, 1992.
- VAN PARIJS, J., BROEKAERT, W. F., GOLDSTEIN, I. J., & PEUMANS, W. J. Hevein: an antifungal protein from rubber-tree (*Hevea brasiliensis*) latex. *Planta*, 183, 258–262, 1991.
- VAREJÃO, N., ALMEIDA, M. D. A. S., DE CICCIO, N. N., ATELLA, G. C., COELHO, L. C., CORREIA, M. A., FOGUEL, D. Heterologous expression and purification of a biologically active legume lectin from *Cratylia mollis* seeds (CRAMOLL 1). *Biochim Biophys Acta.* 1804(9), 1917-1924, 2010.
- VARKI, A. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, 3, 97-130, 1993.
- VARROT, A., BLANCHARD, B. AND IMBERTY, A. Lectin Binding and its Structural Basis, in Carbohydrate Recognition: Biological Problems, Methods, and Applications (eds B. Wang and G.-J. Boons), *John Wiley & Sons, Inc.*, Hoboken, NJ, USA, 2011.
- VASCONCELOS, I.M.; OLIVEIRA, J.T.A. Antinutritional properties of plant lectins. *Toxicon*, 44, (4), 385-403, 2004.
- VOELKER, T. A., STASWICK, P. CHRISPEELS, M. J. Molecular analysis of two phytohemagglutinin genes and their expression in *Phaseolus vulgaris* cv. Pinto, a lectin-deficient cultivar of the bean. *Embo J.*, 5 (12), 3075-3082, 1986.
- VOGELSTEIN, B., GILLESPIE, D. Preparation and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci.*, USA 76, 615, 1979.
- WAH, D. A., ROMERO, A., GALLEGO DEL SOL, F., CAVADA, B. S., RAMOS, M. V., GRANGEIRO, T. B. Crystal structure of native and Cd/Cd-substituted *Dioclea guianensis* seed lectin. A novel manganese-binding site and structural basis of dimer-tetramer association. *J. Mol. Biol.*, 310, 885–894, 2001.
- WALJUNO, K., SCHOLMA, R.A., BIENTEM, J., MARIONO, A., HAHN, A. M. Amino acid sequence of heveine. *Proc. Int. Rubber Conf.*, Kuala Lumpur, 2, 518-531, 1975.
- WALLACE, A. C., LASKOWSKI, R. A., THORNTON, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.*, 8, 127-134, 1995.
- WANG, C. C., HUANG, Y.-L., REN, C.-T., LIN, C.-W., HUNG, J.-T., YU, J.-C., YU, A. L., WU, C.-Y., WONG, C.-H. Glycan microarray of Globo H and related structures for quantitative analysis of breast cancer. *PNAS*, 105, 11661-11666, 2008.

- WANG, L. C.; KANG, L.; HU, T. M.; WANG, J. L. Abrin-a A chain expressed as soluble form in *Escherichia coli* from a PCR-synthesized gene is catalytically and functionally active. *Biochimie*, 86, 327-333, 2004.
- WEIS, W. I., DRICKAMER, K. Structural basis of lectin-carbohydrate recognition. *Annu. Rev. Biochem.*, 65, 441-473, 1996.
- WIERZBA-ARABSKA, E., MORAWIECKA, B. Purification and properties of lectin from potato tubers and leaves; interaction with acid phosphatase from potato tuber. *Acta Biochim. Pol.* 34, 407-420, 1987.
- WISEMAN, T., WILLISTON, S., BRANDTS, J. F., LIN, L. N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.*, 179, 131-137, 1989.
- YAMAUCHI, D., MINAMIKAWA, T. Structure of the gene encoding concanavalin A from *Canavalia gladiata* and its expression in *Escherichia coli* cells. *FEBS Letters*, 260(1), 127-130, 1990.
- YIM, M., ONO, T., IRIMURA, T. Mutated plant lectin library useful to identify different cells. *Proc. Natl. Acad. Sci.*, 98, 2222-2225, 2001.
- YOUNG, N. M., WATSON, D. C., THIBAUT, P. Mass spectrometric analysis of genetic and post-translational heterogeneity in the lectins jacalin and *Maclura pomifera* agglutinin. *Glycoconjugate Journal*, 12, 135-141. 1995.
- ZHU, K., HUESING, J. B., SHADE, R. E., BRESSAN, R. A., HASEGAWA, P. M., MURDOCK, L. L. An insecticidal N-acetylglucosamine-specific lectin gene from *Griffonia simplicifolia* (Leguminosae). *Plant Physiol.*, 110, 195-202, 1996.

ANNEXE 2:

**Sequence, structure and specificity of
a lectin from *Platypodium elegans*
seeds with unusual affinity for
asymmetric complex N-glycans**

(Paper to Journal of Biological Chemistry)

**Sequence, structure and specificity of a lectin from *Platypodium elegans*
with unusual affinity for asymmetric complex N-glycans**

Raquel Guimarães Benevides^{1,2}, Geraldine Ganne¹, Rafael da Conceição Simões², Carlo Unverzagt³, Valérie Chazalet¹, Christelle Breton¹, Annabelle Varrot¹, Benildo Sousa Cavada^{2*} & Anne Imberty^{1*}

¹CERMAV-CNRS (affiliated with University Joseph Fourier and to ICMG), BP53, F-38041 Grenoble cedex 09, France.

²Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, CE, Brazil.

³Bioorganische Chemie, Gebäude NW1, Universität Bayreuth, 95440 Bayreuth, Germany.

* Correspondance to Anne Imberty or Benildo Cavada

E-mail: imberty@cermav.cnrs.fr or bscavada@ufc.br

Short title: Lectin from *Platypodium elegans*

Abstract

Lectin activity with specificity for mannose and glucose has been detected in the seed of *Platypodium elegans*, a legume from the *Dalbergiae* tribe. The gene of the lectin PELa has been cloned and the resulting 261 amino acid protein belongs to the legume lectin family with similarity with *Pterocarpus angolensis* agglutinin (PAL) from the same tribe. The recombinant lectin has been expressed in *E. coli* and refolded from inclusion bodies. Analysis of specificity by Glycan Array evidenced a very unusual preference for complex type N-glycans with asymmetrical branches. A short branch consisting of one mannose residue is preferred on the 6-arm of the N-glycan, while extension by GlcNAc, Gal and NeuAc are favorable on the 3-arm. Affinities have been obtained by microcalorimetry using symmetrical and asymmetrical Asn-linked heptasaccharide prepared by semi-enzymatic method. Strong affinity of 5 μ M was obtained for both ligands. Crystal structures of PELa complexed with branched trimannose and symmetrical complex type Asn-linked heptasaccharide have been solved at 2.1 and 1.65 Å resolution respectively. The lectin adopts the canonical dimeric organization of legume lectins. The trimannose bridges the binding sites of two neighbouring dimers, resulting in the formation of infinite chains in the crystal. The Asn-linked heptasaccharide binds with the 6-arm in the primary binding site and extensive additional contacts on both arms. The GlcNAc on the 3-arm is bound in a constrained conformation that may rationalize the higher affinity that is observed on chips for oligosaccharide with shorter 3-arm that do not present this monosaccharide.

Keywords : Legume lectin, *Dalbergiae*, N-linked glycan, crystallography, microcalorimetry

Abbreviations :

AAP : abridged anchor primer; AP : anchor primer; ConA: concanavaline A; DGL : *Dioclea grandiflora* lectin; Gal: galactose; GalNAc: N-acetyl-galactosamine; Glc : glucose; GlcNAc: N-acetyl-glucosamine; ITC : isothermal titration calorimetry; K_d : dissociation constant; LacNAc: N-acetyl-lactosamine; LCL : *Lens culinares* lectin; Man: Mannose; PAL: *Pterocarpus angolensis* lectin; PELa: *Platypodium elegans* lectin A; PSL : *Pisum sativum* lectin; TriMan: trimannose α Man1-3(α Man1-6)Man

Introduction

Legume lectins comprise a large family of homologous carbohydrate-binding proteins that are found primarily in the seeds of legume plants. Despite similarities in their physicochemical properties and their relatively conserved primary sequences, Leguminosae lectins display considerable diversity in their carbohydrate binding specificities, not only in terms of the recognition of monosaccharides, but also in terms of differential binding to complex carbohydrates.¹ Legume lectins are central to the study of the molecular basis and specificity of protein–carbohydrate interactions² and they also have technological implications for the understanding of cell–cell recognition, adhesion, tumor spread, bacterial and viral infection, and inflammation.³ However, the function of plant lectins remains elusive,⁴ although they are generally proposed to play a role in defense mechanisms against pathogens and invertebrate or vertebrate predators⁵ or in establishment of symbiosis.⁶ The mechanism of action still remains unclear even though induced response mechanisms are proposed for many pathogen-mediated injuries in plants. The direct interference with viruses and microorganisms is rather exceptional, but the deleterious effects of plant lectins on both predatory invertebrates and animals are well documented.⁵ The existence of hydrophobic sites within the structure of leguminous lectins which bind phytohormones suggests a possible role in certain aspects of hormonally regulated plant growth and development.⁷

Leguminous plant lectins, although differing in carbohydrate-binding specificity have similar amino acid sequences¹ and similar fold at the monomer level.⁸ The architecture of the so-called ‘legume lectin fold’ is structurally related to a jelly-roll motif characterized by a β -sandwich of 25–30 kDa that contains a carbohydrate-combining site, a tightly bound Ca^{2+} and a transition metal ion, usually Mn^{2+} . Variability occurs both at the level of the quaternary fold, with a variety of dimeric and tetrameric arrangement⁹ and at the level of the binding site. The floor of the binding site consists of few conserved key amino acids but its perimeter depends on five loops with very variable size and sequence that control the carbohydrate specificity.^{8; 10; 11}

Lectins isolated from legume seeds have been regarded as paradigm for investigating the structural basis and thermodynamics of selective sugar recognition.^{12; 13} The interaction between legume lectin and carbohydrate interactions is generally characterized by a low affinity for monosaccharides (mM range). Longer oligosaccharides can yield to micromolar

affinity, generally by generating a multivalent effect by cross-linked neighboring lectins in solution.^{14; 15}

The large majority of leguminous lectins that have been isolated and characterized belong to the *Phaseoleae* and *Vicieae* tribes of the *Papilionoideae* subfamily of *Leguminosae*.¹⁶ In this context, it would be of interest to characterize the structure and fine specificity of other lectins in different tribes in order to extend our knowledge on protein-carbohydrate interaction. The *Dalbergiae* tribe is of particular interest since only one lectin has been structurally characterized, i.e. the agglutinin from *Pterocarpus angolensis* seed (PAL). Crystallography and thermodynamic studies demonstrated high affinity for both oligomannose N-glycans and for biantennary complex type, correlated with the presence of unusually extended binding region close to the primary binding site.^{17; 18; 19; 20}

In this study, we characterize the lectin from *Platypodium elegans* Vog, another member of the *Dalbergiae* tribe. The genus *Platypodium* comprises two species of leguminous trees. *Platypodium elegans* Vog. is a natural specie founded in the Brazil cerrado and in transition areas between cerrado and seasonal forest.²¹ We demonstrate here the presence of a mannose/glucose lectin (PELa) in *P. elegans* seeds. The cloning of the gene has been performed, resulting in the amino acid sequence determination of PELa. From the recombinant lectin, a complete study of the structure-specificity relationship has been performed and we describe here the molecular basis of the unusual affinity of PELa for asymmetric complex N-glycans.

Results

Identification of lectin activity in Platypodium elegans seeds extract

Crude extract from *P. elegans* seeds agglutinated rabbit erythrocytes with a titer of 2098 hemagglutination units per milliliter. Human erythrocytes with blood group A, B and O were also tested but no hemagglutination was observed. These results indicate the expression of at least one lectin in seeds with specificity for the oligosaccharides present in rabbit erythrocytes. Several monosaccharides and disaccharides were tested for their ability to inhibit PELA-induced rabbit erythrocytes agglutination as shown in [Table 1](#). Glucose and its derivatives α -methyl-glucoside and GlcNAc demonstrated some inhibition, albeit at rather large concentration. The most efficient compound is mannose that could inhibit the agglutination at concentration of 3 mM. This result indicates the presence of a mannose/glucose-binding lectin in the crude extract from *P. elegans* seeds.

RNA isolation and cDNA Cloning:

Total RNA was extracted from the seeds of *P. elegans* and amplification of cDNA ends (3' RACE and 5' RACE) was performed using degenerated and specific primers. The specific degenerated primers were designed based on sequence fragments involved in binding to calcium and manganese in legume lectins (Figure 1S and Table 1S in Suppl. Mat.). 3' RACE was carried out using primer PLAT2 and an adapter (AP), but the band obtained was poor. On the reamplification, with the band of 500 bp as a template, the fragment was amplified. Sequence analysis of the 500 bp 3' RACE terminal sequence showed a stop codon, a 165 bp sequence of 3'-noncoding region and a poli-(A) tail. 5' RACE was performed using cDNA fragments as a template. The amplified product by the primers PLAT3 and an adapter (AAP) gave a 800 bp sequence of PELA mRNA, with a start codon and a 15 bp sequence of 5'-noncoding region (Figure 2S in Suppl. Mat.).

The amino acid sequence was deduced from 5' RACE product clone. Analysis of the genomic sequence indicates that the protein has 261 amino acids. Prediction of post-translational N- glycosylation from the amino acid sequence evidenced two consensus at Asn7 (NST) and Asn127 (NES). PELA sequence has similarities with other Glc/Man specific legume lectins, such as those from *Pterocarpus angolensis* (72 %), *Pterocarpus rotundifolius* (82 %) and *Arachis hypogaea* (70 %) (ID numbers: CAD19804, AAT57665 and AAA74572, respectively). Alignment with a set of legume lectins with known crystal structures is

displayed in Figure 1. Sequence identities vary from 72 % for the Man/Glc specific lectin from *P. angolensis* to 39 % for the Gal/GalNAc lectin from *E. corallodendron*. The five loops that have been described to build the carbohydrate binding sites^{8; 10} are highlighted. Sequences in these loops present similarities with *P. angolensis* lectin, but not much with other lectins, in agreement with the large panel of monosaccharides that can be bound by legume lectins.

Production of recombinant PELA

Production of recombinant protein in *E. coli* BL21 resulted in the accumulation of the protein in inclusion bodies. A renaturation protocol including solubilisation by 8 M urea was used to obtain the protein in the soluble form (Figure 3S in Suppl. Mat.). Active PELA was obtained by affinity purification on a manose-agarose column with a final yield of 5 mg of PELA per liter of cell culture. The soluble lectin displays a molecular mass of 29 kDa, in agreement with the one calculated from the translated gene (theoretical MW: 29,287 kDa). The activity of the lectin was checked at each step of production and purification by hemagglutination assay.

Specificity of PELA with Glycan Microarray Resource

The lectin specificity was analyzed by Glycan Array (Consortium for Functional Glycomics) that allows the screening of a broad spectrum of natural and synthetic glycans. The Printed Array Version 4.1 containing 465 natural and synthetic glycans was used and PELA was observed to bind to a limited number of oligosaccharides (Figure 2A). Short oligosaccharides are not bound with the exception of linear or branched oligomannose fragments (compounds A and B). The lectin is specific for N-glycans. Oligomannose type N-glycan with terminal non-reducing α -mannose are high affinity ligands (compounds C and D). However, the lectin displays a clear preference for asymmetrical complex type N-glycans with a short 6-arm, containing only the α Man residue, and long 3-arm with N-acetyl-lactosamine (LacNAc) (compound F) or sialylated LacNAc (compound E) on the α Man residue. Symmetrical complex N-glycans (compounds H, K and L) are not efficiently bound.

Glycan array results allowed a semi-quantitative analysis, by reporting the intensities of fluorescence at different protein concentrations. As displayed in Figure 2B, the preference for asymmetrical complex N-glycans is conserved for the three PELA concentrations tested (50, 10 and 1 $\mu\text{g}\cdot\text{ml}^{-1}$) of lectin. The influence of the terminal sialic acid on the 6-arm seems

to be moderate since oligosaccharides E and F have rather equivalent response at low protein concentrations. Comparing the concentration-dependant fluorescence response between oligosaccharides E and M, and also F and N, clearly confirms that the addition of one GlcNAc on the 3-arm significantly decreases the lectin affinity.

Glycan array data for other legume lectins are available in <http://www.functionalglycomics.org/>. The classical Glc/Man legume lectins such as ConA and PSA from *Pisum sativum* binds preferentially to oligomannose-type N-glycans, more weakly to complex N-glycan, and not much to asymmetrical N-glycans. These data therefore demonstrate that the unusual specificity of PELA for asymmetrical glycans is clearly different from the classical Glc/Man legume lectins.

Affinity measurements by titration microcalorimetry

The affinity between PELA and the two Asn-linked oligosaccharides (compounds E and H) was estimated by titration microcalorimetry by injecting aliquots of ligands in the cell containing the protein. The thermograms obtained for both oligosaccharides are displayed in Figure 4. In both cases, strong exothermic peaks are observed indicating a favorable enthalpy of binding as classically observed in protein/carbohydrate interactions.¹³ Using a simple one site model, the two oligosaccharides tested bind with similar affinities, with K_d of 5 μ M. Such affinity is unusually strong for legume lectins and has been obtained only for trimannoside in multivalent binding to ConA or *Dioclea grandiflora* lectin (DGL).^{14; 15} In the case of the symmetrical glycan H, a better fit is obtained using a model with two sites indicating that each branch of the heptasaccharide can bind to the lectin, but with different affinities. However, no quantitative data can be safely extracted from the two sites fit. On the opposite, the unsymmetrical glycan E fits very well with the one site model.

As displayed in Figure 34, the shapes of the thermograms and of the integrated heat curve are very different for the two oligosaccharides. While both compounds have approximately the same affinity for PELA, and therefore similar free energy of binding (ΔG of about -30 kJ/mol) (Table 2), the enthalpy and entropy contributions are very different. The symmetrical compound H has a medium size enthalpy of binding (ΔH =-23 kJ/mol) and a favorable entropy that contributes to approx 20% of the binding energy ($-T\Delta S$ = -7.6 kJ/mol). On the opposite the unsymmetrical compound E has a much higher enthalpy of binding (ΔH = -60 kJ/mol) and a strong entropy barrier with an unfavorable contribution of $-T\Delta S$ =30

kcal/mol. The presence of an additional GlcNAc on the 6-arm seems therefore correlated with loss of enthalpy contribution.

Crystal structure of PELa in complex with oligosaccharides

Overall description. Crystals for different complexes were obtained by the hanging drop vapor diffusion method. The protein at 10 mg ml⁻¹ was incubated with 1 mM oligosaccharide prior to testing crystal growth. Co-crystallization with a the branched trimannose (α Man1-3 (α Man1-6)Man: TriMan), yield to crystals diffracting to 2.1 Å in space group P2₁2₁2₁ (Table 3). The structure was solved by molecular replacement using the coordinates of PAL (PDB code 1N3O). A dimer of lectin and one molecule of ligand are observed in the asymmetric unit. Crystals were also obtained with the symmetrical glycan I, diffracting to 1.6 Å in the same space group. The structure was solved by molecular replacement using the coordinates of the first structure above. A lectin dimer and two oligosaccharides are observed in the asymmetric unit.

In both structures, each monomer of PELa adopts the classical β -sandwich fold observed in all legume lectins, with the presence of one calcium and one manganese ion⁹ (Figure 5). Of the 261 residues resulting from the cloning, only 235 to 239 could be identified in the electron density. The longer observed peptide is chain B of PELa/heptasaccharide complex that spans from Ser8 to Tyr248. The two monomers in the asymmetric unit associate by parallel arrangement of their back β -sheet, resulting in the “canonical legume lectin dimer” that is classically observed in both dimeric and tetrameric legume lectins⁹. The two potential glycosylation sites Asn7 and Asn127 are located in the disordered N-term region and on the small helix at the surface of the protein, respectively and are therefore likely to be glycosylated in plants.

PELa/trimannose complex. The complex of the lectin with TriMan contains very clear density for the trisaccharide bridging two symmetry related PELa dimers (Figure 6A). The two α -linked mannose residues are bound in facing monomers, the 1-3linked mannose in monomer A and the 1-6 linked one in monomer B (Figure 6B). The binding modes of the two mannose residues are very similar, each establishing five direct hydrogen bonds with sides chains of Asp95, Asn 147 and main chain nitrogen of Gly115 and Glu230 (Figure 6C). Hydrophobic contacts involve Ala94 and Phe141. This interaction corresponds to the

classical mannose binding mode observed in PAL²⁰ and ConA.²² A conserved water bridge is also observed in both sites between O2 and the carbonyl oxygen of Ala113.

The central core-mannose residue establishes bonds with both monomers, involving Asn145, Ser146 and Gln231 of chain A and Glu230 of chain B. The bridging TriMan results in rather close contact of the two facing protein monomers. Protein-protein contacts are created with the participation of Arg123B side chain hydrogen bonded to Thr148A and Asp150A main chain. Also Asn145 residue of each monomer interferes with the binding site of the other monomer and participate to hydrogen bond to mannose O2 for Asn145B or to hydrophobic contact to C2 for Asn145A (Figure 6B and C).

As a result of its interaction with facing monomers, the TriMan shape is very extended and differs from what has been previously observed in other complexes with legume lectins such as PAL,²⁰ ConA²³ and DGL.²⁴ Comparison between PELa and PAL confirms the strong similarity between the binding sites, with same conformation of the five loops, labeled A to E (Figure 7A). PAL binds the trimannoside with the α Man residue of the 6-arm in the primary binding site, as in monomer B of PELa. Nevertheless, the observed oligosaccharide conformation is very different since TriMan folds around the phenyl residue in the binding site of PAL and the 3-linked mannose is hydrogen bonded to an asparagine residue in Loop A. The difference in conformation is due to a 90° shift in the Ψ dihedral angle of the α Man1-6Man linkage. Comparison with lectins from the *Dioclae* tribe is displayed in Figure 7B. Binding sites loops are well conserved between *Dioclae* (ConA and DGL) and *Dalbergiae* (PAL and PELa), with the exception of loop C that is shorter in ConA/DGL than in PAL/PELa (14 amino acids instead of 16) resulting in a more open binding site, as previously described.²⁰ As a result, the TriMan conformation in ConA and DGL is different: it folds back on the protein surface and the 1-3 linked α -Man makes additional contact with amino acids of Loop C. The observed conformation of TriMan in ConA could not be possible in PELa due to the different conformation of loop C that would result in a steric conflict.

PELa/Asn-linked heptasaccharide complex. Clear density for the whole biantennary heptose was observed in both monomers of PELa, and the linked asparagine residue could be located in monomer B (Figure 8A) because of its stabilization by salt bridge with Arg221 of a symmetry related protein. Apart for this slight difference, the oligosaccharides in chain A and B are very similar in term of conformation and hydrogen bonds (details available in Table 2S

and 3S in Suppl. Mater.) and all description below will refer to monomer B, unless otherwise indicated. The mannose of the 6-arm is located in the primary binding site (subsite M) with the same orientation and the same hydrogen bond network as described above for the TriMan ligand. With the exception of the asparagine and the first GlcNAc attached to it, all residues are interacting with the protein surface, resulting in an extended surface of contact (Figure 8B). The second GlcNAc of the chitobiose core sits against Loop C and establishes a water-bridged hydrogen bond with Asn145 main chain oxygen. The numerous other contacts involve both arms of the biantennary glycans. The GlcNAc on the 6-arm establishes three hydrogen bonds with amino acids of site -1 (Figure 8C and Table 3S in Suppl. Mater.). The core mannose in site +2 and the 3-arm in sites +2 and +3 also display a large number of contacts. Altogether, more than 15 direct hydrogen bonds and 6 water-bridged ones are observed, rationalizing the high affinity measured by titration microcalorimetry.

The conformation of the oligosaccharides have been analyzed and compared to previously determined energy maps²⁵ (available from <http://glyco3d.cermav.cnrs.fr>). The conformation β GlcNAc1-2Man and α Man1-3Man of the linkages in the 3-arm are close to the global minima (Figure 9A). For the other antenna, the α Man1-6Man linkage is in a low energy conformation, with a GG conformation of the ω angle, but the β GlcNAc1-2Man linkage is clearly distorted with a Ψ value of about 115° , in a region of the energy map that is 4 kcal/mol above the global minimum at $\Psi = 150^\circ$.

The conformation of the biantennary glycan and the network of contact are very similar to what have been described previously for the PAL/pentasaccharide of the complex type,¹⁷ albeit with small differences at the extremities of branches (Figure 8C). The same pentasaccharide complexed with ConA yield to the same position of 6- arm in subsite M and -1, but to a different orientation of the 3-arm, again because of the important difference in Loop C conformation.²⁶ The same distorted orientation of the β GlcNAc1-2Man linkage in subsite -1 was described for the ConA complex. It should be noted that in other complexes such as PAL/heptasaccharide,¹⁷ this linkage adopts a completely different orientation with a large distortion in the Φ angle.

Conclusion

The manuscript describes a very complete sequence/structure/specificity characterization of a legume lectin o from the *Dalbergiae* tribe with unusual affinity for asymmetric complex N-

glycans. The sequence and structure present similarity with the lectin from *P. angolensis*, PAL, the only *Dalbergiae* lectin that was structurally characterized previously.^{17; 18; 19; 20} Some of the structural features of binding of PELa to complex N-glycans are also common to both lectins, such as the occurrence of an extended binding site that results in very large surface of contact between the oligosaccharide and the protein. Nevertheless, very important new data are revealed by the present work. This is the first characterisation of a lectin from this sub-family by Glycan Array. Interestingly, PELa binds to a subset of complex N-glycans, and this binding profile is very different to those already determined for other lectins with Glc/Man specificity such as ConA, DGL, PSL from *Pisum sativum* and LCL from *Lens culinaris* that all bind preferably to oligomannose type N-glycans (data available at <http://www.functionalglycomics.org/>). We could also determine the very high affinity of PELa for Asn-linked complex type oligosaccharide.

The crystallisation of PELa in complex with trimannose results in infinite lectin chains in the crystal with one carbohydrate ligands bridging two lectins monomers through its terminal non-reducing mannose residues. Such trimannose-bridging has already been observed in crystal of snowdrop agglutinin²⁷ and bacterial lectin,²⁸ but not in legume lectins. The resulting conformation of the trimannoside is very extended and drastically different to the folded shape previously observed in other legume lectin complexes.^{20; 23; 24} Interestingly, the conformational change is not due to a rotation around the ω angle, that keeps a *gauche-gauche* conformation in the different legume lectins, but in the Ψ angle that is rotated by 90° as displayed in Figure 9.

The preference for asymmetrical N-glycans with a short 6-arm has been confirmed by a semi-quantitative analysis of the Glycan Array data performed at several lectin conformation. No crystals could be obtained for the complex between PELa and the asymmetric compound E, but the complex with the symmetric compound H (heptasaccharide-Asn) gives some clues about this preference. The GlcNAc on the 3-arm establishes contact in the extended sites+3 and +3 while maintaining a low energy conformation of the linkages involved. On the opposite, the binding of the GlcNAc of the 6-arm in site -1 induces a strong conformational distortion with energy cost. Oligosaccharides with long 3-arm and short 6-arm, i.e. without GlcNAc, would therefore be energetically favoured for binding. There is therefore an excellent agreement between Glycan Array and crystallographic data. However, solution data from ITC are not concordant since both the asymmetric and symmetric glycan binds with same strong affinity to PELa. Nevertheless, the enthalpy profile are different, and the

asymmetric glycan has a much stronger enthalpy of binding, which is in agreement with the hypothesis that an extra GlcNAc on the 6-arm is not favourable. The difference in entropy contribution between the two oligosaccharides is more difficult to explain. However, it is clear that the entropy contribution would be much different for an oligosaccharide presented by its Asn anchor on a chip (or a glycoprotein) than for the same oligosaccharide in solution. This work therefore exemplifies one of the cases where the physicochemical state of the ligand, i.e. solution state or attached to a surface, can influence its affinity for a receptor.

Materials and Methods

Materials

Platypodium elegans seeds were purchased from Reserva de Reflorestamento Flora Tietê (São Paulo, Brazil) and were stored at room temperature or at -70 °C. Monosaccharides, disaccharides, sulfated polysaccharides (Fucoidan, Carageenan) and glycoprotein (Pork Stomach Mucine) were purchased from Sigma-Aldrich (USA). The oligosaccharides were gift from Carlos Unverzagt (Germany)

Haemagglutination tests on seeds extract

The seeds were ground and later deflated in the presence of n-hexane. A protein extract was prepared from 1 g of the ground seeds with 10 mL of 0.15 M NaCl at room temperature for three hours. The homogenate was centrifuged at 16000 g for 20 min at 4 °C and the precipitate was discarded. The hemagglutinating activity of the lectin was determined by a 2-fold serial dilution procedure using native rabbit or human erythrocytes as described previously.²⁹ The hemagglutination titer was defined as the reciprocal of the highest dilution exhibiting observable hemagglutination. Inhibition of agglutination by sugar was assayed by serially diluting the solution of sugar in the microtiter wells, followed by the addition of four agglutinin units of the lectin, followed by the addition of erythrocyte suspension after 30 min. The lowest concentration of sugar that visibly decreased the extent of agglutination was defined as the minimum inhibitory concentration (MIC).

Cloning of PELA and sequencing

For RNA isolation, 1 g of seeds was ground to a powder with a pestle under liquid nitrogen. Total cellular RNA was isolated with Concert Plant RNA reagent (Invitrogen), according to the manufacturer's protocol.

3'-RACE and 5'-RACE. Partial cDNA was synthesized from total mRNAs using rapid amplification of cDNA ends (3'-RACE kit, Invitrogen) with the combination of the following primers: PLAT1 (5'-TNGCNGTNGARTTYGAYAC-3') or PLAT2 (5'-CAYATHGGNATHGAYGTNAA-3') how forward primers and adapter primer (Invitrogen) how reverse primer. The degenerated forward primers (PLAT1 or PLAT2) for the gene of the seed lectin of *P. elegans* were designed for minimal degeneration, from the primary sequence of *Pterocarpus angolensis* lectin (accession number: CAD19811). We used the corresponding sequences from the conserved metal-binding sites of legume lectins. All the PCR's were performed using GoTaq polimerase (Promega). After amplification and purification, the parcial cDNA amplified was cloned into pGEM®T-easy (Promega). Sequence analysis was performed to confirm the according sequence and was used to design a specific reverse primer to synthesize the total cDNA from total mRNA using 5'-RACE kit (Invitrogen): PLAT3 (5'-TGATATAGAGATCCCTCTGGA-3'); how forward primer was used the abridged ancor primer (Invitrogen). The PCR fragment was purified, cloned into the pGEM®T-easy (Invitrogen) and submitted to sequence analysis. The amplified DNA was further cloned into pET-29a: the oligonucleotides primers were as follows: 5'-GGAATTCCATATGCTACTGAATAGGGCAAAC-3' (PLAT4) and 5'-GGCAAAGCTTTCACATATCACGTGCAAGATAC-3' (PLAT5), containing NdeI and HindIII restriction sites, respectively (underlined sequences). PCR was performed using the 5'-RACE amplified product as a template. After digestion with NdeI and HindIII, the amplified fragment was introduced into the multiple cloning site of pET-29a (Invitrogen), resulting in the plasmid pET-29a_PELA.

Production and purification of recombinant PELa

Escherichia coli BL21 cells containing the plasmid pET-29a_PELa were cultured in LB broth medium at 37 °C. When the culture reached an OD600 of 0.5–0.6, IPTG (isopropyl b-D-thiogalactoside) was added to a final concentration of 0.5 mM. Cells were harvested after 4 h incubation at 30 °C, washed and suspended in equilibrating buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM CaCl₂, 5 mM MnCl₂). The cells were broken by cell disruption (Constant Cell Disruption System) at 1700 bars. After centrifugation at 20000 g for 30 min at 4 °C and filtration, the pellet was 2x washed with 50 mM Tris-Cl pH 8.0, 4 M Urea, 0.5 M NaCl, 1 mM EDTA, 0.5 % Triton X-100. The final precipitate was denatured and solubilized in 50 mM Tris-Cl pH 8.0, 8 M Urea, 10 mM DTT and, after centrifugation at 20000 g for 30 min at 4 °C, had its protein concentration adjusted to 1 mg/ml. 50 mL of the denatured protein solution at 1mg/mL was slowly diluted in 950 mL of 50 mM Tris-HCl pH 7.5, 240 mM NaCl, 2 mM MnCl₂, 2 mM CaCl₂, 10 mM KCl, 1mM DTT. After 1 hour at 10 °C and filtration, the fresh refolded protein was submitted to the purification step.

The purification of PELa was performed on a D-mannose-Sepharose column (Sigma-Aldrich). PELa refolded was allowed to bind to the immobilized mannose in equilibrating buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM CaCl₂, 5 mM MnCl₂). After washing with equilibration buffer, PELa was eluted using 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM CaCl₂, 5 mM MnCl₂, 200 mM D-Mannose. The purified protein was washed and concentrated as desired using a Vivaspın ultrafiltration device (cutoff 10 kDa; GE Healthcare) and stored at 4 °C.

The fractions and the protein purity were analyzed by 12% SDS-PAGE and the protein was stained with Coomassie brilliant blue (Sigma-Aldrich).

Isothermal Titration Microcalorimetry analysis

ITC experiments were performed with a VP-ITC isothermal titration microcalorimeter (Microcal; GE Healthcare). Experiments were carried out at 25 °C ± 0.1 °C. Protein was prepared in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM CaCl₂, 5 mM MnCl₂ and the same buffer was used to prepare the oligosaccharide solutions. Protein concentrations used for ITC was 0.06 mM. Injections of 10 µL of each oligosaccharide solution at various concentrations (from 0.4 to 0.8 mM) were added automatically to the protein solution present in the calorimeter cell. Control experiments performed by injection of buffer into the protein solution yielded insignificant heats of dilution. Integrated heat effects were analyzed by non-

linear regression using a single-site binding model (Origin 7.0). The experimental data fitted to a theoretical titration curve gave the association constant (K_a) and the enthalpy of binding (ΔH). The other thermodynamic parameters such as changes in free energy (ΔG) and entropy (ΔS) were calculated from the equation $\Delta G = \Delta H - T\Delta S = -RT\ln K_a$, where T is the absolute temperature and R is the molar gas constant ($8.314 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$). All experiments were performed with c values of $10 < c < 100$.

The glycan microarray analysis

Purified PELA sample was labeled with Alexa Fluor 488 (Invitrogen, CA) according to manufacturer's instructions and re-purified on a D-Salt polyacrylamide desalting column (Pierce, Rockford IL). Alexa-labeled protein was used for glycan-array screening with standard procedure of the Protein-glycan Interaction Core (H) of the Consortium for Functional Glycomics (CFG) (<http://www.functionalglycomics.org>).

Crystallization and structure determination

PELa was concentrated in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5mM CaCl_2 , 5 mM MnCl_2 to 10 mg/mL with addition of Triman or compound H at the final concentration of 1 mM. Crystallization conditions for PELA were screened using the hanging-drop vapor diffusion method with a commercially available crystallization solution (Hampton Research Screens I and II, Hampton Research, Riverside, CA, USA) at 20 °C. Drops were prepared by mixing 1.0 μL sample solution and 1.0 μL reservoir solution, followed by equilibration against 500 μL reservoir solution.

Crystals of PELA/TriMan complex were obtained using crystallization condition 42 from the Crystal Screen II kit (1.5 M Ammonium sulfate, 0.1 M Tris pH 8.5, 12 % v/v Glycerol). Crystals were frozen in liquid nitrogen (100K) after soaking them for as short a time as possible in 25 % (v/v) glycerol in precipitant solution. Data sets were collected at ESRF (European Synchrotron Radiation Facility), Grenoble, France, on stations ID14–2 using an ADSC Q4R CCD (charge-coupled device) detector (Quantum Corp.). A diffraction data set was collected to 2.2 Å resolution in the orthorhombic space group $p2_12_12_1$. Crystals for PELA-heptasaccharide complex were obtained from solution 48 of the Hampton screen II containing 100 mM bicine pH 9 10 % PEG 20K and 2 % dioxane. The concentration of glycerol was increased to 30 % for cryoprotection prior mounting and freezing in liquid nitrogen. Data sets were collected at ESRF on station BM30 using an ADSC Q4R CCD

(charge-coupled device) detector (Quantum Corp.) to 1.6 Å in the orthorhombic space group $p2_12_12_1$. The statistics of the both data collections are given in Table 3.

Structures determination

Phasing information was obtained using a monomer of *Pterocarpus angolensis* lectin (PDB code 1N3O, with removal of water molecules) as a probe for molecular replacement for the orthorhombic crystal form of PELa/TriMan complex using the program PHASER.³⁰ For the PELa/heptasaccharide complex, the crystal structure of the PELa/TriMan complex was used as template. Crystallographic refinement was carried out with the program REFMAC³¹ and manual model building was achieved using Coot.³² Refinement statistics of both structures are given in Table 3. Co-ordinates have been deposited at the Protein Data Bank under code 3ZVX and 3ZYR, respectively. Molecular drawings were prepared using PyMOL Molecular Graphics System (DeLano Scientific, Palo Alto, CA) and LIGPLOT³³ (<http://www.ebi.ac.uk/thornton-srv/software/LIGPLOT/>).

Protein Data Bank accession number

Protein sequence : AEK69351

Crystal structures : 3ZVX, 3ZYR

Acknowledgements

The glycan microarray analysis was provided by the Consortium for Functional Glycomics funded by the National Institute of General Medical Sciences grant GM62116. We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities. This work was supported by CNRS and French Ministry of Research. RGB was the recipient of CAPES and CNPq grants from Brazilian government in different intervals.

References

1. Sharon, N. & Lis, H. (1990). Legume lectins--a large family of homologous proteins. *Faseb J* **4**, 3198-208.
2. Sharon, N. & Lis, H. (1995). Lectins--proteins with a sweet tooth: functions in cell recognition. *Essays Biochem* **30**, 59-75.
3. Rudiger, H., Siebert, H. C., Solis, D., Jimenez-Barbero, J., Romero, A., von der Lieth, C. W., Diaz-Marino, T. & Gabius, H. J. (2000). Medicinal chemistry based on the sugar code: fundamentals of lectinology and experimental strategies with lectins as targets. *Curr Med Chem* **7**, 389-416.
4. Rudiger, H. (1998). Plant lectins - more than just tools for glycoscientists: occurrence, structure, and possible functions of plant lectins. *Acta Anat (Basel)* **161**, 130-52.
5. Peumans, W. J. & Van Damme, E. J. (1995). Lectins as plant defense proteins. *Plant Physiol* **109**, 347-52.
6. Hirsch, A. M. (1999). Role of lectins (and rhizobial exopolysaccharides) in legume nodulation. *Curr Opin Plant Biol* **2**, 320-6.
7. Delatorre, P., Rocha, B. A., Souza, E. P., Oliveira, T. M., Bezerra, G. A., Moreno, F. B., Freitas, B. T., Santi-Gadelha, T., Sampaio, A. H., Azevedo, W. F., Jr. & Cavada, B. S. (2007). Structure of a lectin from *Canavalia gladiata* seeds: new structural insights for old molecules. *BMC Struct Biol* **7**, 52.
8. Loris, R., Hamelryck, T., Bouckaert, J. & Wyns, L. (1998). Legume lectin structure. *Biochim. Biophys. Acta* **1383**, 9-36.
9. Srinivas, V. R., Reddy, G. B., Ahmad, N., Swaminathan, C. P., Mitra, N. & Surolia, A. (2001). Legume lectin family, the 'natural mutants of the quaternary state', provide insights into the relationship between protein stability and oligomerization. *Biochim Biophys Acta* **1527**, 102-11.
10. Sharma, V. & Surolia, A. (1997). Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J. Mol. Biol.* **267**, 433-445.
11. Young, N. M. & Oomen, R. P. (1992). Analysis of sequence variation among legume lectins. A ring of hypervariable residues forms the perimeter of the carbohydrate binding site. *J. Mol. Biol.* **228**, 924-934.
12. Ambrosi, M., Cameron, N. R. & Davis, B. G. (2005). Lectins: tools for the molecular understanding of the glycode. *Org Biomol Chem* **3**, 1593-608.
13. Dam, T. K. & Brewer, C. F. (2002). Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. *Chem. Rev.* **102**, 387-429.
14. Dam, T. K., Oscarson, S. & Brewer, C. F. (1998). Thermodynamics of binding of the core trimannoside of asparagine-linked carbohydrates and deoxy analogs to *Dioclea grandiflora* lectin. *J Biol Chem* **273**, 32812-7.
15. Williams, B. A., Chervenack, M. C. & Toone, E. J. (1992). Energetics of lectin-carbohydrate binding. A microcalorimetric investigation of concanavalin A-oligomannoside complexation. *J. Biol. Chem.* **267**, 22907-22911.

16. Cavada, B. S., Moreira, R. A., Oliveira, J. T. A. & Grangeiro, T. B. (1993). Primary structures and functions of plant lectins. *R. Bras. Fisol. Veg.* **5**, 193-201.
17. Buts, L., Garcia-Pino, A., Imberty, A., Amiot, N., Boons, G. J., Beeckmans, S., Versees, W., Wyns, L. & Loris, R. (2006). Structural basis for the recognition of complex-type biantennary oligosaccharides by *Pterocarpus angolensis* lectin. *Febs J* **273**, 2407-20.
18. Garcia-Pino, A., Buts, L., Wyns, L., Imberty, A. & Loris, R. (2007). How a plant lectin recognizes high mannose oligosaccharides. *Plant Physiol* **144**, 1733-41.
19. Loris, R., Imberty, A., Beeckmans, S., Van Driessche, E., Read, J. S., Bouckaert, J., De Greve, H., Buts, L. & Wyns, L. (2003). Crystal structure of *Pterocarpus angolensis* lectin in complex with glucose, sucrose, and turanose. *J Biol Chem* **278**, 16297-303.
20. Loris, R., Van Walle, I., De Greve, H., Beeckmans, S., Deboeck, F., Wyns, L. & Bouckaert, J. (2004). Structural basis of oligomannose recognition by the *Pterocarpus angolensis* seed lectin. *J Mol Biol* **335**, 1227-40.
21. Lorenzi, H. (1998). *Árvores brasileiras: Manual de identificação e cultivo de plantas arbóreas nativas do Brasil*, Plantarum, Nova Odessa.
22. Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb, A. J., Dodson, E. J., Papiz, M. Z., Wan, T. & Campbell, J. (1989). The structure of the saccharide-binding site of concanavalin A. *Embo J* **8**, 2189-93.
23. Loris, R., Maes, D., Poortmans, F., Wyns, L. & Bouckaert, J. (1996). A structure of the complex between concanavalin A and Methyl-3,6-di-O-(α -D-mannopyranosyl)- α -D-mannopyranoside reveals two binding modes. *J. Biol. Chem.* **271**, 30614-30618.
24. Rozwarski, D. A., Swami, B. M., Brewer, C. F. & Sacchettini, J. C. (1998). Crystal structure of the lectin from *Dioclea grandiflora* complexed with core trimannoside of asparagine-linked carbohydrates. *J Biol Chem* **273**, 32818-25.
25. Imberty, A., Delage, M. M., Bourne, Y., Cambillau, C. & Pérez, S. (1991). Data bank of three-dimensional structures of disaccharides : Part II, N-acetylglucosaminic type N-glycans. Comparison with the crystal structure of a biantennary octasaccharide. *Glyconconj. J.* **8**, 456-483.
26. Moothoo, D. N. & Naismith, J. H. (1998). Concanavalin A distorts the β -GlcNac-(1-2)-Man linkage of β -GlcNac-(1-2)- α -Man-(1-3)-[β -GlcNac-(1-2)- α -Man-(1-6)] upon binding. *Glycobiology* **8**, 173-181.
27. Wright, C. S. & Hester, G. (1996). The 2.0 Å structure of a cross-linked complex between snowdrop lectin and a branched mannopentaose: evidence for two unique binding model. *Structure* **4**, 1339-1352.
28. Lameignere, E., Shiao, T. C., Roy, R., Wimmerová, M., Dubreuil, F., Varrot, A. & Imberty, A. (2010). Structural basis of the affinity for oligomannosides and analogs displayed by BC2L-A, a *Burkholderia cenocepacia* soluble lectin. *Glycobiology* **20**, 87-98.
29. Moreira, A. & Perrone, J. C. (1977). Purification and Partial Characterization of a Lectin from *Phaseolus vulgaris*. *Plant Physiol* **59**, 783-7.

30. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674.
31. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55.
32. Emsley, P., Lohkamp, B., Scott, W. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica. Section D: Biological Crystallography* **66**, 486-501.
33. Wallace, A. C., Laskowski, R. A. & Thornton, J. M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127-34.
34. Gouet, P., Courcelle, E., Stuart, D. I. & Metz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-8.
35. Bourne, Y., Rougé, P. & Cambillau, C. (1992). X-ray structure of a biantennary octasaccharide-lectin complex refined at 2.3-Å resolution. *J. Biol. Chem.* **267**, 197-203.

Figure legends

Figure 1. Aligned amino acid sequences of legume lectins. Multiple alignment and secondaries structures prediction were achieved using the webESPrpt program.³⁴ The sequence data were obtained from the Entrez protein sequence database. ID numbers: AEK69351 (*Platypodium elegans* lectin), CAD19804 (*Pterocarpus angolensis* lectin), 2PEL_A (*Arachis hypogaea* lectin), CAA25787 (*Canavalia ensiformes* lectin), BAA36415 (*Robinia pseudocacia* lectin) and CAA36986 (*Erythrina collarodendrom* lectin). Symbols: alpha helices (α), 3_{10} helices (η), beta strands (β), strict alpha turns (TTT), beta turns (TT), sugar binding residues (green square), metal binding residues (purple square).

Figure 2: Glycan array data. Upper panel : binding of labeled PELa (10 μ g/ml) to all carbohydrates of the glycan array v 4.1. Lower panel: comparison for binding at three lectin concentrations (green 50 μ g/ml, orange 10 μ g/ml, violet 1 μ g/ml).

Figure 3: Schematic representation of the oligosaccharides-Asn.

Figure 4: Microcalorimetry data. ITC plot (measured by VP-ITC, Microcal) obtained from the titration of PELa (60 μ M) by Asn-linked glycans H (panel A) and E(panel B). The plots in the lower panel show the total heat released as a function of total ligand concentration for the titration shown in the upper panels. The solid line represents the best least-square fit to experimental data using one site model.

Figure 5: Overall structure of a dimer of PELa complexed with trimannoside. Peptide chains are represented as ribbon colored in red (chain A) and blue (chain B). Oligosaccharides are represented by sticks and calcium and manganese ions by green and violet spheres, respectively.

Figure 6. Crystal structure of PELa/trimannoside complex. A. Chains of PELa lectins bridged by trimannosides with same color code as in Figure 5. B. Close view of the interface with the trimannoside represented together with its $2mF_o-DFc$ electron density maps contoured at 1σ . Chain A is represented by its accessible surface and chain B as cartoon. C. Ligplot representation of the hydrogen bonds and van der Waals contact around the trimannoside.

Figure 7. Comparison of PELa/oligosaccharide complexes with related ones from other legume lectins together with labelling of loops from the binding site. A. Superimposition

of chain A (blue) and chain B (cyan) from PELA/trimannoside complex together with PAL/trimannoside structure (magenta: PDB code 1Q8V).²⁰ B. Orthogonal view for chains A and B of PELA together with trimannoside complexed with ConA (dark green: 1ONA)²³ and with *Dioclea grandiflora* lectin (light green: 1DGL).²⁴ C. Superimposition of chain B (cyan) from PELA/Asn-linked heptose complex with complex oligosaccharides in the crystal of PAL (magenta: 2ARX, salmon: 2A6X)¹⁷, *Lathyrus ochrus* lectin I (yellow: 1LOF)³⁵ and ConA (green: 1TEI).²⁶

Figure 8. Crystal structure of PELA/Gn2Man3Gn2-Asn complex. A. 2mFo-DFc electron density maps contoured at 1σ of the heptasaccharide-Asn in monomer B. B Representation of the oligosaccharide on the protein surface. C. Ligplot representation of the hydrogen bonds for the ligand in monomer A. The chitobiose and asparagine moiety do not establish hydrogen bond with the protein, and for the sake of clarity, they are not displayed in the figure.

Figure 9. Analysis of the conformation of oligosaccharides bound to PELA. A. Energy maps are from 3D-cermav (<http://glyco3d.cermav.cnrs.fr>). Conformation at each linkage of TriMan and heptose-Ans bound to PELA are represented by circles and triangles respectively. Open and filled triangles represent conformations of the 6-arm and 3-arm, respectively. B. Comparison of the three binding modes observed in crystal structures: TriMan bound through the α Man1-3 (cyan) and by the α Man1-6 (green) residues, and heptose-Asn (yellow).

Table 1. Inhibition of hemagglutination Activity of *P. elegans* seeds crude extract by various sugars

<i>Sugars</i>	<i>Minimum inhibitory concentration (MIC)</i>
D-Galactose	-*
Methyl- α -D-Galactopyranoside	-
D-Galactosamine	-
<i>N</i> -acetyl- β -D-Galactosamine	-
L-Fucose	-
D-Glucose	50 mM
Methyl- α -D-Glucopyranoside	12.5 mM
<i>N</i> -acetyl-D-Glucosamine	25 mM
D-Mannose	3.125 mM
<i>N</i> -acetyl-D-Mannosamine	-
D-Arabinose	-
Lactose	-
Lactulose	-
Fuoidan	-
Carageenan	-
Pork Stomach Mucine	-

* No inhibition observed at concentration of 100 mM

Table 2. Microcalorimetry titration data for the binding of the oligosaccharide E and H to PELa. All measured values are averaged over at least two experiments. Standard deviations are <0.03 for n values and < 5% for the other ones.

Ligand	K_a (10^6 M^{-1})	K_d (μM)	n	ΔG (kJ/mol)	$-\Delta H$ (kJ/mol)	$-T\Delta S$ (kJ/mol)
Compound H	0.22	4.4 (0.3)	1	-30.6	23.0 (0.3)	-7.6 (0.4)
Compound E	0.16	6.0 (0.2)	0.75	-29.9	59.3 (1.8)	29.4 (1.8)

Table 3. Data collection and refinement statistics

	<i>Triman</i>	<i>Hepta-Asn</i>
Beamline (wavelength, Å)	Id14EH1/0.933	BM30A/0.9797
Spacegroup	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions <i>a,b,c</i> (Å)	42.62 80.68 148.76	51.40 76.98 125.59
α, β, γ (deg.)	90.0 90.0 90.0	90.0 90.0 90.0
Resolution (outer shell), Å	42.60-2.10 (2.21-2.10)	47.57-1.65 (1.74-1.65)
Measured/ Unique reflections	99535/28894	186416/55022
Average multiplicity	3.4 (2.4)	3.4 (2.6)
R _{merge}	0.082 (0.330)	0.054 (0.250)
Completeness (%)	93.7 (70.1)	91.4 (65.8)
Mean I / σI	9.3 (2.0)	13.6 (3.9)
Wilson B	24.7	14.00
R _{cryst} /R _{free}	19.5/25.2	15.5/18.8
R _{msd} bonds, Å	0.015	0.015
R _{msd} angles, °	1.54	1.59
Rmsd chiral, Å ³	0.106	0.098
Protein atoms	1935/1926	2004/1986
Bfac, Å ²	23.5/23.7	12.4/14.0
Water molecules	177/149	295/223
Bfac, Å ²	29.0/28.3	25.7/26.6
Ligand atoms	34	89/92
Bfac, Å ²	22.6	14.2/16.6
Hetero atoms	18/8	30/12
Bfac, Å ²	32.4/33.5	24.7/25.7
PDBcode	3ZVX	3ZYR

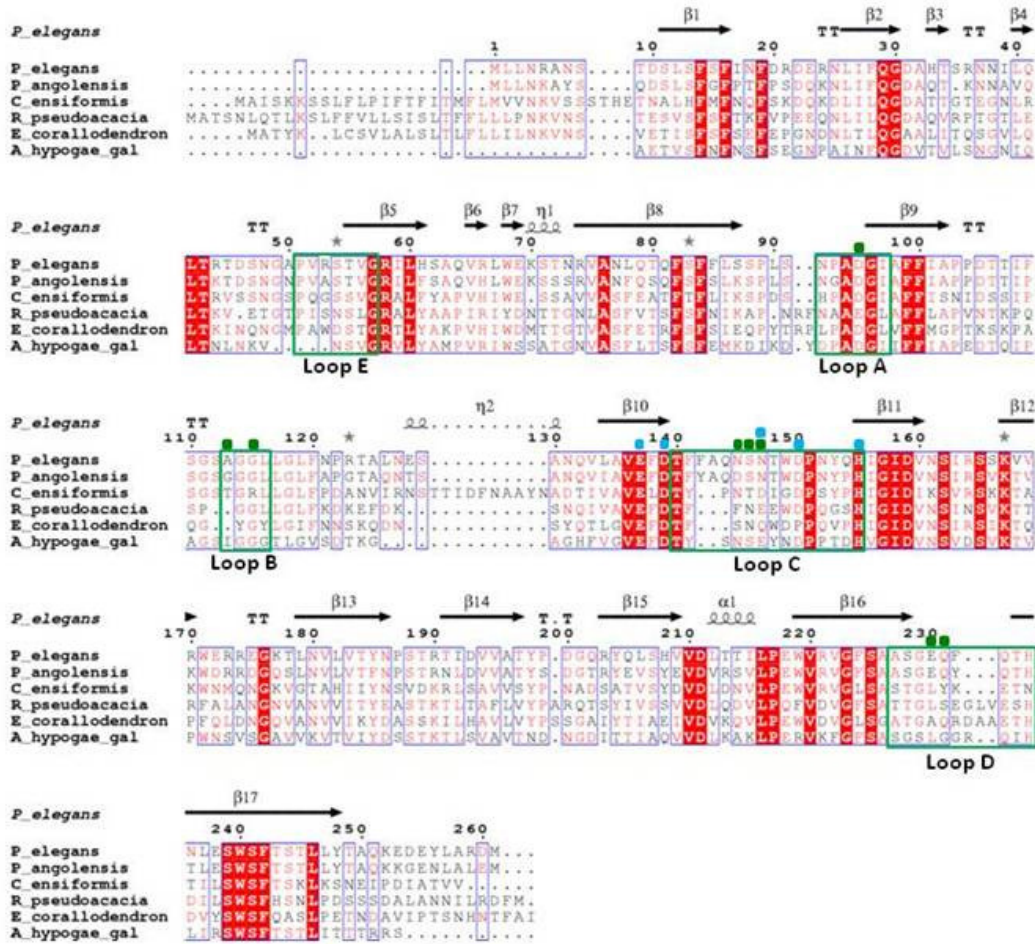


Figure 1

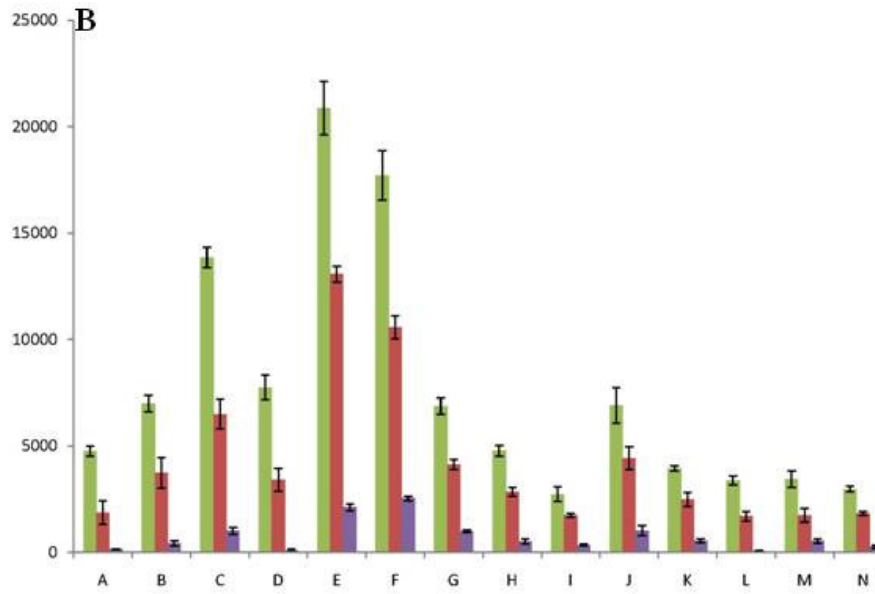
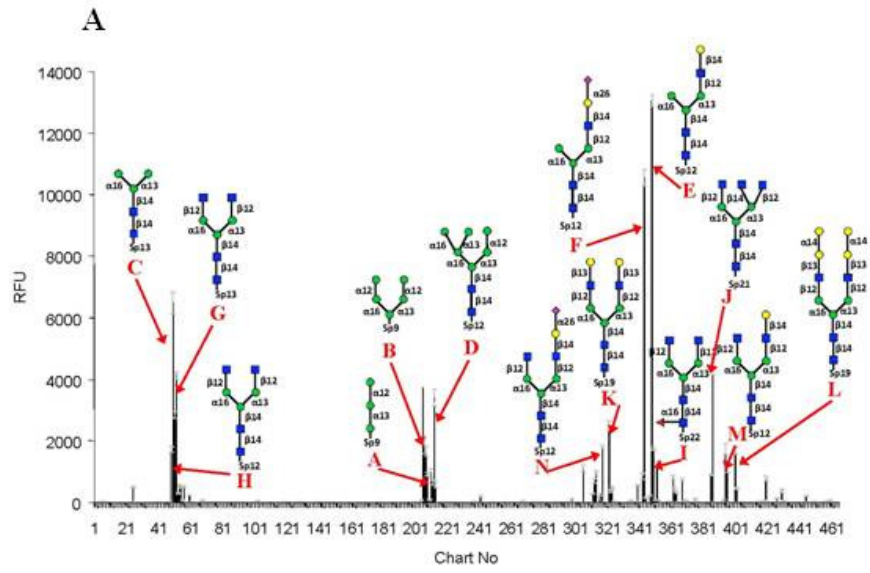


Figure 2

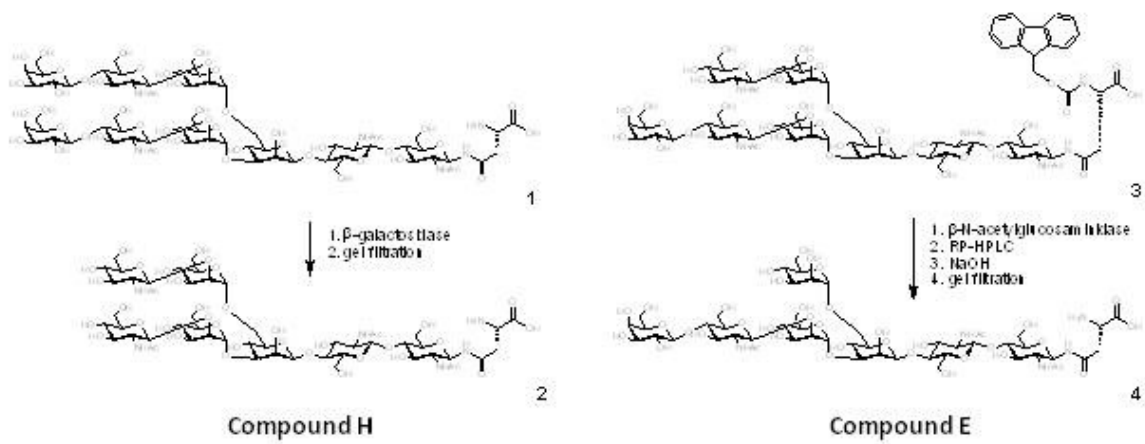


Figure 3

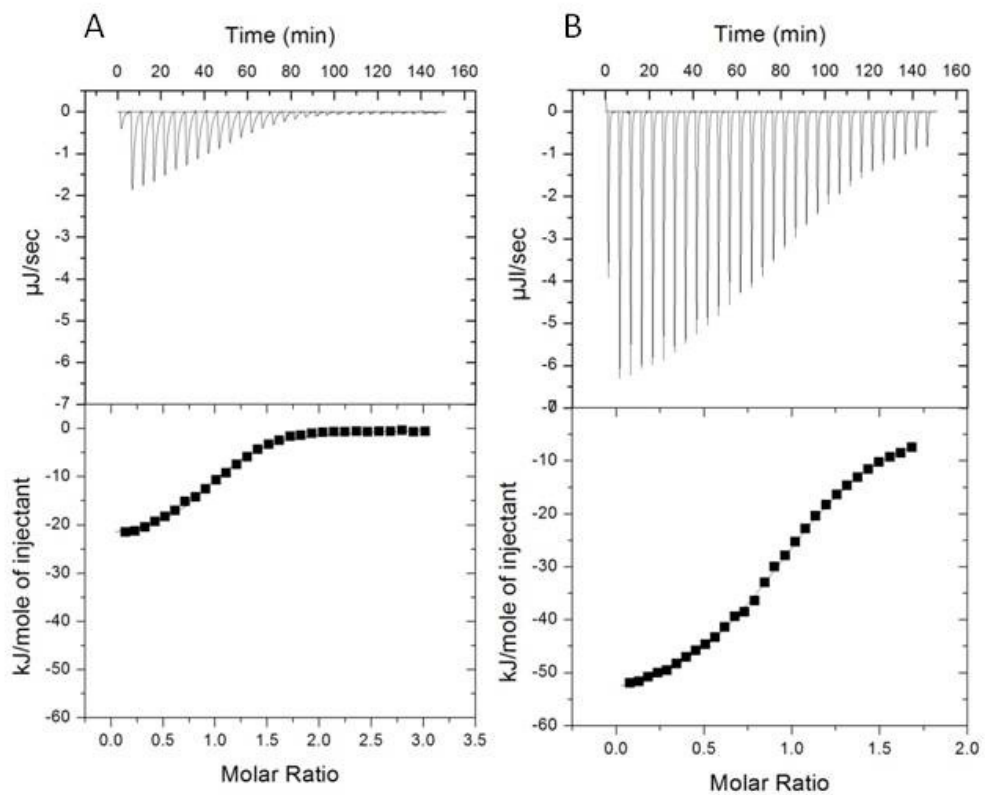


Figure 4

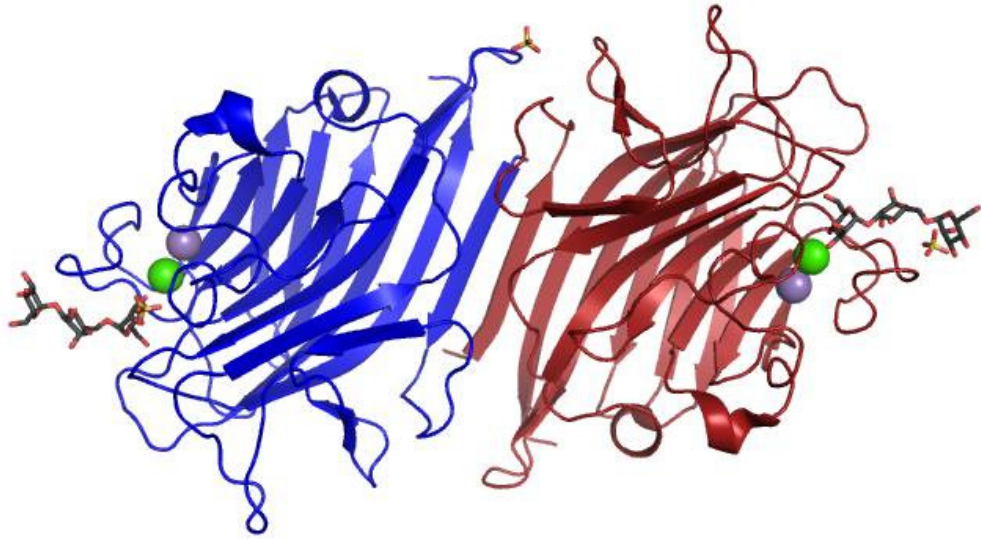


Figure 5

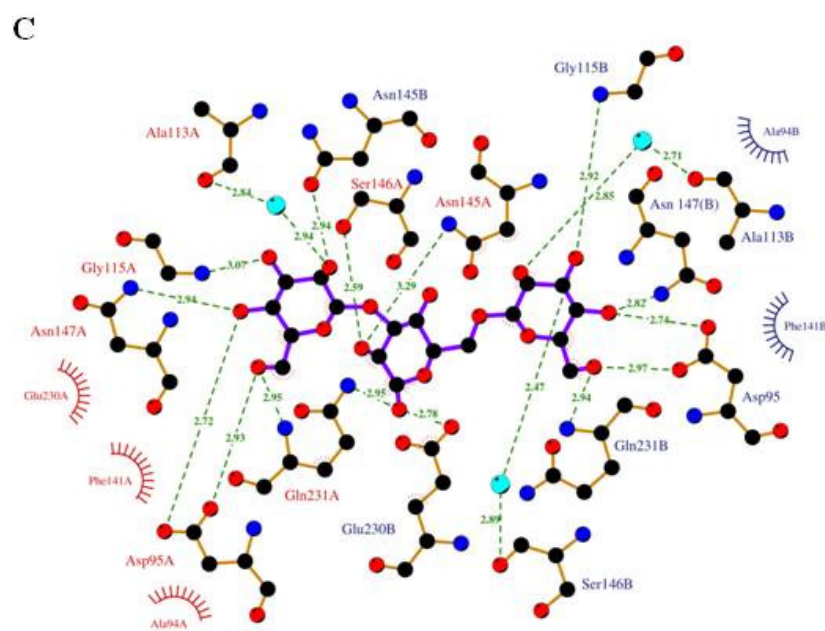
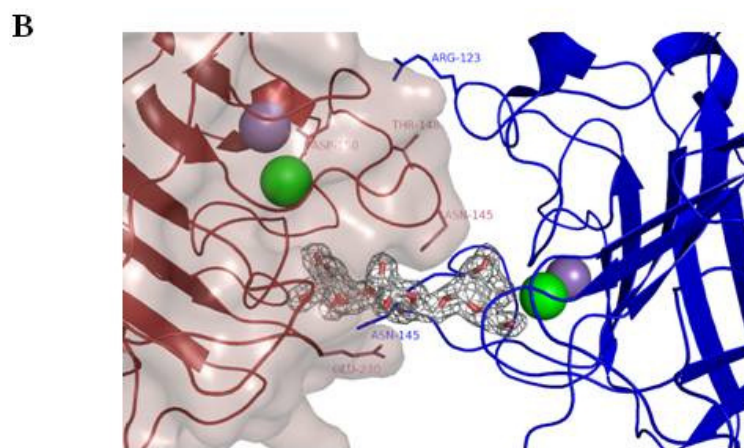
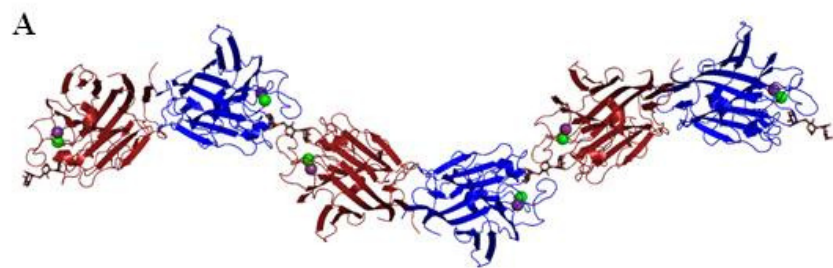


Figure 6

Table 1S. List of primers used for the cloning of PELa

Primer	Type	5'- Sequence – 3'*	Corresponding Aminoacid Sequence	Step
PLAT1	forward	TN GCN GTN GAR TTY GAY AC	(V/L/I)AVEFDT	3'-RACE
PLAT2	forward	CAY ATH GGN ATH GAY GTN AA	HIGIDVN	3'-RACE
PLAT3	reverse	TGATATAGAGATCCCTCTGGA	SRGISIS	5'-RACE
AP	reverse	GGCCACGCGTCGACTAGTACTTTTTTTTTTTTTTTTTT		3'-RACE
AAP	forward	GGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG		5'-RACE
PLAT4	forward	GGAATTCCATATGCTACTGAATAGGGCAAAC	(NdeI)MLLNRA	pET-29a cloning
PLAT5	reverse	GGCAAGCTTTCACATATCACGTGCAAGATAC	LARDM Stop (HindIII)	pET-29a cloning

*R, Y, and N indicate (A or G), (C or T), and (A, C, G or T), respectively. I, inosidic acid.

The underlined sequences indicate the positions of the NdeI and HindIII, respectively.

Table 2S. Torsion angles (Φ, Ψ) or (Φ, Ψ, ω)^a at each glycosidic linkages in the crystal structures of PelA/oligosaccharide complexes. All values in °.

	TriMan	Hept_chain A	Hept_Chain B
αMan1 -6Man	70.2, 104.8, -52.6	90.3, -153.2, -57.1	85.6, -151.4, -60.4
βGlcNAc1-2Man		-78.9, 115.3	-84.3, 117.9
αMan1 -3Man	65.1, 122.6	70.3, 126.5	68.9, 128.0
βGlcNAc1-2Man		-84.6, 152.1	-83.0, 152.3
βMan1-4GlcNAc		-68.2, -127.6	-70.1, -130.3
βGlcNAc1-4GlcNAc		-66.8, -113.2	-57.1, -120.7

^a $\Phi = \theta(O5-C1-O1-Cx')$, $\Psi = \theta(C1-O1-Cx'-Cx_{+1}')$, $\omega = \theta(O1-C6'-C5'-O5')$,

Table 3S. List of hydrogen bonds between the Asn-linked heptose and PELa

	Heptose	Protein
αMan1 -6	O2	
	O3	G115.N Wat1...Ser146.O Wat1...O6 GN(6)
	O4	Ap95.OD1 Asn147.ND2
	O5	Glu230.N
	O6	ASP95.OD2 Gly229.N Gln231.N
βGlcNAc1-2 (6 arm)	O3	Wat2...Ala113.N
	O4	Gly111.O
	O6	Wat1...Ser146.O
	O7	Ser54.OG
αMan1 -3	O6	Asn145.ND2
βGlcNAc1-2 (6 arm)	O6	Asn92.N Wat3...Ser91.A
	O7	Wat4...Asn92.OD1
βMan1-4	O2	Asn145.OD1 Ser146.OG
	O4	Gln231.OE1 Wat5...Gln231.NE2
βGlcNAc1-4	N2	Wat6...Asn145.O

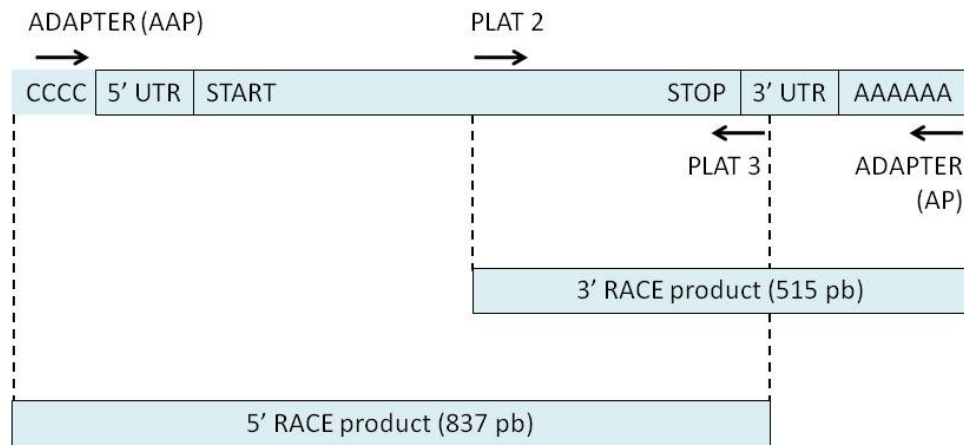


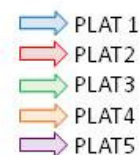
Figure 1S. The complete structure of PELa coding mRNA obtained from 3' and 5' RACE sequence products (coding region: 786 nucleotides), UTR: uncoded terminal region.






```

ATG CTA CTG AAT AGG GCA AAC TCA ACA GAT TCC CTT TCC TTC AGC TTC ATT AAC TTT GAT 60
CGA GAC GAA AGA AAC CTG ATC TTT CAA GGT GAT GCA CAT ACT TCA CGA AAC AAT ATA CTC 120
CAG CTT ACC AGG ACA GAC AGT AAT GGT GCG CCA GTG AGA AGC ACT GTT GGG CGA ATC TTG 180
CAC TCA GCA CAA GTG CGC CTT TGG GAG AAA AGC ACA AAC AGA GTC GCA AAC CTT CAA ACC 240
CAG TTT AGC TTT TTC CTT TCA TCA CCA CTT TCC AAT CCG GCC GAC GGC ATC GCC TTC TTC 300
ATT GCA CCG CCA GAT ACT ACC ATA CCA TCC GGT TCG GCT GGA GGT CTG TTG GGG CTC TTT 360
AAC CCT AGA ACT GCT CTT AAT GAG TCT GCA AAC CAA GTG CTT GCT GTT GAG TTT GAC ACC 420
TTC TTT GCT CAA AAC TCT AAC ACT TGG GAT CCG AAT TAC CAG CAC ATT GGA ATC GAT GTC 480
AAC TCC ATT AGG TCC TCG AAG GTT GTG AGA TGG GAG AGG AGA GAA GGT AAA ACC CTT AAT 540
GTG CTT GTA ACA TAT AAC CCT AGC ACT CGA ACC ATA GAT GTC GTT GCT ACT TAC CCT GAT 600
GGT CAA AGA TAC CAG CTA TCT CAT GTG GTT GAC CTG ACG ACT ATA CTT CCA GAA TGG GTT 660
AGA GTT GGA TTC TCA GCT GCC TCA GGA GAG CAA TTT CAA ACG CAT AAT CTT GAA TCA TGG 720
TCT TTC ACC TCA ACC TTG TTG TAC ACC GCC CAG AAG GAG GAC GAG TAT CTT GCA CGT GAT 780
ATG TGA TCC AGA GGG ATC TCT ATA TCA TAT CTG TCA GGT ATG ATA TGC TAA TAA AGG ACA 840
AGT GTT GCC CAT GGT CTA AGT ACG TTA TGG TTT TTG TTG CTT ATG TTG AGT CAC GTG TGT 900
AAA ATA AAC TTC GGA TGA AGG CGT ATG CTA TCT TGC TAT TAT TGT TTG TCC AAA AAA AAA 960
AAA AAA AAA AAA AAA AAA 978

```

Figure 2S. PELa mRNA sequence. PLAT 1, PLAT2, PLAT 3, PLAT4 and PLAT5 sequence primers are displayed in blue, red, green, orange and violet boxes, respectively. PLAT1 and PLAT2 are forwards and degenerated primers; PLAT3 is specific and reverse; PLAT4 (forward) and PLAT5 (reverse) are specific, with restriction sites for endonucleases (NdeI and HindIII, respectively)



 PLAT 1
 PLAT 2
 PLAT 3
 PLAT 4
 PLAT 5

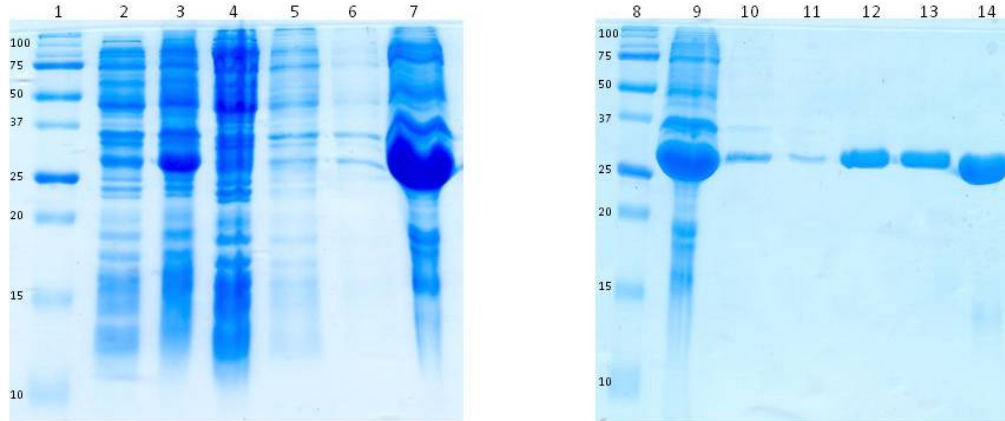


Figure 3S. **Refolding and purification of PELA.** 1,8. Molecular weight marker (kDa) 2. Crude homogenate from *E. coli* cells before induction 3. Crude homogenate from *E. coli* cells after 4h of induction with 0.5 mM IPTG 4. Soluble proteins from disrupted cells 5,6. Inclusion bodies pellet washed with 4 M urea 7,9. Inclusion bodies suspended in 8 M urea 10. Refolded PELA before purification 11. Fraction not retained on Mannose-Sepharose 12,13. Fraction retained on Mannose-Sepharose eluted with 0,2 M D-mannose(PELA) 14.PELA at 3 mg/mL