



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

**CLAUDENER SOUZA TEIXEIRA**

**DETERMINAÇÃO DA ESTRUTURA TRIDIMENSIONAL DE UMA LECTINA  
DA SEMENTE DE *Camptosema pedicellatum* BENTH POR CRISTALOGRAFIA  
DE RAIOS X**

**FORTALEZA**

**2012**

CLAUDENER SOUZA TEIXEIRA

DETERMINAÇÃO DA ESTRUTURA TRIDIMENSIONAL DE UMA LECTINA DA  
SEMENTE DE *Camptosema pedicellatum* POR CRISTALOGRAFIA DE RAIOS X

Dissertação apresentada ao  
Programa de Pós-Graduação em  
Bioquímica do Departameneto de  
Bioquímica e Biologia Molecular  
da Universidade Federal do Ceará  
como parte dos requisitos para  
obtenção do título de Mestre em  
Bioquímica.

Orientador: Prof. Dr. Benildo  
Sousa Cavada

FORTALEZA

2012

CLAUDENER SOUZA TEIXEIRA

DETERMINAÇÃO DA ESTRUTURA TRIDIMENSIONAL DE UMA LECTINA DA  
SEMENTE DE *Camptosema pedicellatum* POR CRISTALOGRAFIA DE RAIOS X

Dissertação apresentada ao Programa de Pós Graduação de Bioquímica do  
Departamento de Bioquímica e Biologia Molecular da Universidade Federal do Ceará  
como parte dos requisitos para obtenção do título de Mestre em Bioquímica.

Aprovada em \_\_\_\_/\_\_\_\_/\_\_\_\_

BANCA EXAMINADORA

Prof. Dr. Benildo Sousa Cavada  
Depto. Bioquímica e Biologia Molecular  
Universidade Federal do Ceará UFC

**Orientador**

Prof. Dr. Bruno Anderson Matias da Rocha  
Depto. Bioquímica e Biologia Molecular  
Universidade Federal do Ceará UFC

**Co-Orientador**

Prof. Dr. Plínio Delatorre  
Depto. Biologia Molecular  
Universidade Federal da Paraíba UFPB

**Examinador**

A meus pais biológicos Maria Lucia e Luiz Braz por todo apoio, dedicação e amor, a meus pais de coração Francisca Lucena e Afonso por toda força e carinho, a meus irmãos Claudenir e Junior por sempre apostarem em mim e a meu amor, Raquel, por estarem sempre ao meu lado. Amo vocês!

## AGRADECIMENTOS

Agradeço primeiro a Deus, criador do nosso rico e biodiverso universo, autor de todas as coisas. Agradeço pela sua grande obra e pelo dom da minha vida.

Ao meu orientador, Prof. Dr. Benildo Sousa Cavada, pelo o apoio, conselhos, amizade, e todos os ensinamentos que levarei por toda minha vida, muito obrigado.

Ao prof. Dr. Bruno Anderson Matias da Rocha, em primeiro lugar pela amizade que construímos durante o início da minha graduação e pós graduação, por sempre acreditar no meu potencial, mesmo distante nunca deixou de investir no meu conhecimento, você sem dúvidas foi quem me despertou para o mundo da ciência e hoje lhe tenho como um espelho profissional, muito obrigado por tudo meu amigo.

Ao Prof. Dr. Celso Shiniti Nagano, por sempre me aconselhar nos experimentos e por conversas que julgo de suma importância para a minha formação profissional, muito obrigado.

A Prof<sup>a</sup>. Dr<sup>a</sup> Kyria Santiago do Nascimento pelos conselhos nos experimentos e por sempre nos proporcionar um ambiente alegre, sempre conduzindo o Biomol-Lab de forma coerente.

Aos amigos, Helton, Rafael Simões, Ito, Mayron, Bruno Lopes, Sâmia, Batista, Ana Claudia, Caio, Lia Claudia, Clareane, Fernando, Alysson, Suzete, Alfa Bari, Kássia, Francisco Júnior, Arthur, Vanir Reis, Mayara, Rafael Queiroz, Camila Nobre Guilherme e Maria Julia, por todos os momentos no laboratório.

Em especial aos amigos Raquel Benevides, Raniére Moura, Rômulo Farias, Vinícius Osterne, Raphael Nóbrega por todos os momentos de aprendizagem e descontração, sempre contei com a amizade verdadeira de vocês desde o primeiro dia de BioMol-lab, muito obrigado por tudo.

Ao meu grande mestre o professor José Carlos Marques Freitas por todos os ensinamentos valiosos sobre a vida, muito obrigado por me guiar neste caminho.

Aos professores Dr Plínio Delatorre e Dr<sup>a</sup> Beatriz Tupinambá pela grande amizade e por fazerem parte da base de minha formação ainda na graduação e até hoje com

ensinamentos e conselhos. Muito obrigado por tudo, vocês sempre acreditaram no meu potencial.

Em especial aos meus amigos de departamento, Vanessa Moraes, Lucas Pinheiro, Washington, Ednésio, Gerardo, Paulo Carvalho, e os demais colegas por todos os momentos de felicidade e tristeza compartilhados.

Aos meus grandes amigos de graduação, Katia, Daniele Kelly (As Fúrias), Bodocó, Laura, Paulinha, Thaianny, Joemília, Ísis, Daniele Fernandes, Tales e Olga por toda amizade que até hoje é conservada.

A todos os Professores e Funcionários do Departamento de Bioquímica e Biologia Molecular da UFC.

Aos meus eternos amigos/irmãos da A8, Niel, Neilson, Renan, Marcos, Anderson, Darlon e Junior por toda força durante todos esses anos, vocês foram minha força motriz, muito obrigado por todos os momentos vividos, das noites de diversão e descontração ao som de Fagner, das farras, da ajuda em momentos difíceis, tenho certeza que essa amizade é verdadeira.

Em especial a um anjo que caiu do céu e que sempre está ao meu lado apesar da distância, meu porto seguro, uma linda rosa que me contempla com o seu cheiro e me tranquiliza com sua beleza, RAQUEL você foi fundamental no meu/nosso caminhar. Te amo minha bela flor!

## **ESTE TRABALHO FOI REALIZADO COM O AUXILIO DAS SEGUINTE INSTITUIÇÕES**

Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), através de bolsa de Mestrado concedida ao candidato, e por auxílios de pesquisa concedidos ao BioMol-Lab (Laboratório de Moléculas Biologicamente Ativas do Departamento de Bioquímica e Biologia Molecular do Centro de Ciências da Universidade Federal do Ceará.

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelos convênios estabelecidos ao Curso de Pos-Graduação em Bioquímica do Centro de Ciências da Universidade Federal do Ceará.

Ministério de Ciência e Tecnologia, através do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), além do financiamento aos Projetos de Cristalografia e Bioinformática.

Laboratório de Moléculas Biologicamente Ativas do Departamento de Bioquímica e Biologia Molecular do Centro de Ciências da Universidade Federal do Ceará, onde se encontra o laboratório de cristalização.

Laboratório de Espectrometria de Massa Aplicado a Proteínas (LEMAP) do Departamento de Engenharia de Pesca da Universidade Federal do Ceará, onde foram realizados os experimentos de espectrometria de massas.

Laboratório Nacional de Luz Síncrotron (LNLS), onde foram realizados os experimentos de difração de Raios X.

“No mundo da ciência, é sempre escabroso fazer da ignorância uma arma de combate, nunca se permita a tal coisa. Não se violente ao consentir ou querer facilidades ao se curvar a criaturas servis.”

(Prof. José Carlos Marques Freitas)



## RESUMO

Lectinas têm sido usadas como modelos em estudos das bases moleculares da interação e especificidade proteína-carboidrato por serem capazes de decifrar o glicódigo de estruturas celulares. O objetivo do presente estudo foi purificar e resolver a estrutura primária completa usando espectrometria de massas em tandem e a estrutura tridimensional da lectina de *Camptosema pedicellatum* (CPL) complexada com 5-bromo-4-cloro-3-indol- $\alpha$ -D-manose (X-Man) por cristalografia de raios X. CPL foi purificada em um único passo por cromatografia de afinidade. Os resultados de espectrometria de massas revelaram que CPL apresenta uma combinação de cadeias com pesos moleculares de  $25.298 \pm 2$  Da (cadeia  $\alpha$ ),  $12.835 \pm 2$  Da (cadeia  $\beta$ ) e  $12.481 \pm 2$  Da (cadeia  $\gamma$ ). A CPL apresenta uma mutação conservada em sua sequência de aminoácidos que corresponde na estrutura tridimensional à região do subsítio hidrofóbico, um componente do domínio de reconhecimento a carboidrato (CDR), indicando a relevância da interação hidrofóbica no estabelecimento de interações com o carboidrato. A substituição do aminoácido e análise da interação da CPL com o X-Man também revelou que o efeito hidrofóbico causado por uma pequena mudança no subsítio hidrofóbico interfere na formação de pontes de hidrogênio devido à reorientação espacial do grupo indol no CDR.

**Palavras-chave:** *Camptosema pedicellatum*. Lectina. Subsítio hidrofóbico.

## ABSTRACT

Lectins have been used as models for studies of the molecular basis of protein-carbohydrate interaction and specificity by deciphering codes present in the glycan structures. The purpose of the present study was to purify and solve the complete primary and crystal structure of the lectin of *Camptosema pedicellatum* (CPL) complexed with 5-bromo-4-chloro-3-indolyl- $\alpha$ -D-mannose (X-Man) using tandem mass spectrometry. CPL was purified by single-step affinity chromatography. Mass spectrometry findings revealed that purified CPL features a combination of chains weighing  $25,298 \pm 2$  Da ( $\alpha$ -chain),  $12,835 \pm 2$  Da ( $\beta$ -chain) and  $12,481 \pm 2$  Da ( $\gamma$ -chain). The CPL amino acid sequence features a conservative mutation which correspond to the hydrophobic subsite, a constituent of the carbohydrate recognition domain (DRC), indicating the relevance of hydrophobic interactions in the establishment of interactions with carbohydrates. The substitution and the analysis of the interactions with X-Man also revealed that the hydrophobic effect caused by a minor change in the hydrophobic subsite interferes in the formation of H-bonds due to the reorientation of the indolyl group in the DRC.

**Keywords:** *Camptosema pedicellatum*.Lectin. Hydrofobic subsite

## LISTA DE FIGURAS

<b>FIGURA 1.</b> Classificação estrutural de lectinas.....	19
<b>FIGURA 2.</b> Classificação das lectinas vegetais.....	21
<b>FIGURA 3.</b> Modificações pós-traducionais durante a biossíntese da Concanavalina A (ConA). .....	23
<b>FIGURA 4.</b> Monômero típico de uma lectina de leguminosa .....	26
<b>FIGURA 5.</b> Representação dos diferentes tipos de oligomerização das galectinas.....	27
<b>FIGURA 6.</b> Representação da interação das lectinas de leguminosas com metais.....	28
<b>FIGURA 7.</b> Representação da área de superfície de um sítio de ligação a compostos hidrofóbicos.....	31
<b>FIGURA 8.</b> <i>Camptosema pedicellatum</i> .....	34
<b>FIGURA 9.</b> Purificação da CPL.....	43
<b>FIGURA 10.</b> Análise por espectrometria de massas da CPL.....	46
<b>FIGURA 11.</b> Sequência de aminoácidos da CPL.....	47
<b>FIGURA 12.</b> Alinhamento múltiplo de sequencias de lectinas ConA-like.....	49
<b>FIGURA 13.</b> Análise MS/MS do pepetídeo T5 e E2.....	50
<b>FIGURA 14.</b> Cristais de CPL.....	52
<b>FIGURA 15.</b> Grafico de Ramachandran das coordenadas da CPL.....	55
<b>FIGURA 16.</b> Estrutura Completa da CPL complexada com X-Man.....	56
<b>FIGURA 17:</b> Mapa da densidade eletrônica da Val99 e do X-Man.....	57
<b>FIGURA 18.</b> Representação da interação a metais da CPL.....	58
<b>FIGURA 19.</b> Representação da área de superfície do domínio de reconhecimento a carboidratos.....	60
<b>FIGURA 20.</b> Superposição do domínio de reconhecimento a carboidrato da CPL e DVL.....	61

**LISTA DE TABELAS**

<b>TABELA 1</b> - Sete diferentes tipos de interface formados durante a oligomerização de lectinas de leguminosas.....	30
<b>TABELA - 2.</b> Atividades farmacológicas relacionadas a lectinas vegetais.....	33
<b>TABELA 3</b> - Efeito inibidor de monossacarídeos e dissacarídeos na atividade hemaglutinante da CPL.....	44
<b>TABELA 4</b> - Sequência de peptídeos da CPL e suas respectivas massas moleculares.....	48
<b>TABELA 5</b> - Estatística da coleta de dados, refinamento e qualidade da estrutura da CPL com X-Man.....	53
<b>TABELA 6</b> - Interações de Van der Waals e contatos polares entre a lectina da semente de <i>C. pedicellatum</i> e X-Man.....	59
<b>TABELA 7.</b> Distâncias entre OH de Tyr12 e N1 do X-Man no domínio de reconhecimento a carboidrato de lectinas Diocleinae.....	62

**LISTA DE ABREVIATURAS E SIGLAS**

- ABU** ácido-aminobutírico;
- AIA** ácido indolacético;
- cDNA** DNA complementar;
- CFL** Lectina de *Cratylia floribunda*;
- CG** Complexo de Golgi;
- CGL** Lectina de *Canavalia gladiata*;
- Cgran** Lectina de *Canavalia grandiflora*;
- CID** – Dissociação induzido por colisão;
- CMI** – Concentração Inibitória Mínima;
- ConA** Concanavalina A;
- ConBr** Lectina de *Canavalia brasiliensis*;
- ConM** Lectina de *Canavalia maritima*;
- CRLI** Lectina manose específica de *Cymbosema roseum*;
- CRLII** Lectina lactose específica de *Cymbosema roseum*;
- DGL** Lectina de *Dioclea grandiflora*;
- Dgran** Lectina de *Dioclea grandiflora*;
- Dgui** Lectina de *Dioclea guianensis*;
- DRC** Domínio de reconhecimento a carboidrato;
- DvirL** Lectina de *Dioclea Virgata*;
- DVL** Lectina de *Dioclea violacea*;
- DwL** Lectina de *Dioclea wilsonii*
- EG** Etileno glicol;
- Fuc** Fucose;
- Gal** Galactose;
- GalNAc** N-acetil galactosamina;
- Glc** Glicose;

**GlcNAc** *N*-acetil glicosamina;

**GNA** Aglutinina de *Galanthus nivalis*;

**GS-I** Isoforma I da lectina de *Griffonia simplicifolia*;

**GS-IV** Isoforma IV da lectina de *Griffonia simplicifolia*;

**KeV** Kilo elétrons volts;

**LNLS** – Laboratório Nacional de Luz Síncroton;

**LTA** Aglutinina de *Lotus tetragonolobus*

**Man** Manose;

**NeuAc** Ácido neuramínico;

**NO** Óxido Nítrico;

**PAGE** Eletroforese em gel de poliacrilamida;

**PBS** Tampão fosfato salino;

**PDB** Banco de Dados de Proteínas;

**PEG** Polietileno glicol;

**PHA** Aglutinina de *Phaseolus vulgaris*;

**PNA** Aglutinina de *Peanuts*;

**rDGL** Lectina recombinante de *Dioclea grandiflora*;

**RE** Retículo endoplasmático;

**RIP** Proteína Inativadora de Ribossomos;

**SBA** Aglutinina de *Glicine max*;

**SDS** Dodecil sulfato de sódio;

**Tris** Tris(hidroximetil)aminometano;

**UEA** Aglutinina de *Ulex europaeus*;

**U.H.** – Unidade Hemaglutinante;

**V<sub>m</sub>** Volume de Matthews;

**X-Man** 5-bromo-4-cloro-3-indol- $\alpha$ -D-manose;

## SUMÁRIO

<b>1.</b>	<b>INTRODUÇÃO</b> .....	16
1.1	Considerações iniciais sobre lectinas vegetais e estruturas.....	16
1.2	Breve histórico.....	16
1.3	Definições.....	18
1.4	Aspectos gerais de lectinas vegetais.....	19
1.5	Lectinas de leguminosas.....	22
1.6	Estruturas tridimensionais das lectinas vegetais.....	24
1.7	Aplicações biológicas.....	32
1.8	Definições botânicas.....	33
<b>2.</b>	<b>OBJETIVOS</b> .....	<b>35</b>
2.1	Objetivos gerais.....	35
2.2	Objetivos específicos.....	35
<b>3.</b>	<b>MATERIAL E MÉTODOS</b> .....	<b>36</b>
3.1	Purificação.....	36
3.2	Atividade hemaglutinante.....	36
3.3	Inibição da atividade hemaglutinante.....	37
3.4	Eletroforese em gel de poliacrilamida em presença de SDS.....	37
3.5	Determinação da massa molecular por espectrometria de massas...38	38
3.6	Digestão em gel e sequenciamento dos peptídeos por espectrometria de massas.....	38
3.7	Cristalização.....	39
3.8	Difração e coleta de dados.....	40
3.9	Substituição molecular e refinamento.....	41
3.10	Validação da estrutura por bioinformática.....	42
<b>4.</b>	<b>RESULTADOS E DISCUSSÃO</b> .....	<b>43</b>
4.1	Purificação da CPL.....	43
4.2	Determinação da massa molecular por espectrometria de massas e análise da sequência primária.....	45
4.3	Estrutura completa da CPL complexada com X-Man.....	51
<b>5.</b>	<b>CONCLUSÃO</b> .....	<b>64</b>
<b>6.</b>	<b>REFERÊNCIAS</b> .....	<b>65</b>
<b>7.</b>	<b>ANEXO</b> .....	<b>76</b>

## 1. INTRODUÇÃO

### 1.1 Considerações iniciais sobre lectinas vegetais e estruturas

Lectinas com propriedades de interesse biotecnológico foram isoladas, caracterizadas, e, um grande número dessas proteínas tem sido obtido por expressão heteróloga com ou sem mutação dirigida no domínio de reconhecimento a carboidratos (GOLDSTEIN, 2010; Werz; Seeberger, 2005).

Nestes últimos anos, várias estruturas de lectinas foram determinadas por cristalografia de raios X com alta resolução e depositadas no banco de dados do *Protein Data Bank* ([www.pdb.org](http://www.pdb.org)). Embora alguns trabalhos tenham reportado o envolvimento dessas proteínas na defesa das plantas (MURDOCK; SHADE, 2002), ainda há muitas perguntas que não foram respondidas: Qual a função biológica das lectinas? Por que algumas plantas têm seu desenvolvimento comprometido quando um determinado gene de lectina é silenciado e em outras plantas esse desenvolvimento não é dependente? Será que lectinas representam um sistema imune primitivo?

A resolução estrutural dessas proteínas tem auxiliado nas pistas dos possíveis papéis biológicos que elas desempenham nas plantas, traçando vias que vão desde o reconhecimento do carboidrato em domínios proteicos até mecanismos pós-reconhecimento, elaborando, assim, uma correlação estrutura/função. Como exemplo, a ligação da lectina da semente de *Dolichos biflorus* à adenina (HAMELRYCK *et al.*, 1999) e da lectina de *Canavalia maritima* complexada ao fitohormônio AIA (dados não publicados). Tais correlações entre lectinas e moléculas envolvidas em processos biológicos importantes como desenvolvimento e defesa da planta são potenciais alvos de estudos dessas proteínas.

### 1.2 Breve histórico das lectinas

No final do século XIX, alguns estudos relacionados a proteínas, com capacidade de aglutinar hemácias, foram explorados em trabalhos científicos por grupos que pesquisavam a toxicidade de extratos de plantas e venenos de serpentes. Essas proteínas eram chamadas de hemaglutininas ou fitohemaglutininas, pois eram em sua maioria encontradas em plantas. A primeira observação de uma atividade aglutinante derivada de lectinas foi descrita por Silas Weir Mitchell em 1860, em extrato de veneno de cobra *Crotalus durissus* (MITCHEL *et al.*, 1860, 1886). Já para as lectinas de



plantas, acredita-se que o primeiro relato foi feito por Peter Herrman Stillmark em sua tese de doutoramento na universidade de Dorpat na Estônia com o título “Über Ricin ein giftiges Ferment aus den Samen von *Ricinus communis* L. und einigen anderen Euphorbiaceen” (Sobre a Ricina: Um fermento tóxico de sementes de *Ricinus communis* L. e algumas outras espécies de Euforbiácias). Ele havia isolado uma hemaglutinina altamente tóxica extraída de sementes de mamona (*Ricinus communis*) a qual denominou como ricina (STILLMARK, 1888). Subsequentemente, H. Hellin isolou também uma hemaglutinina tóxica de sementes de Jequiriti (*Abrus precatorius*). Ambas se tornaram comercialmente disponíveis e foram utilizadas por Paul Ehrlich como antígenos modelos para estudos imunológicos, estabelecendo vários princípios fundamentais da imunologia através dessas proteínas (VAN DAMME *et al.*, 1998; SHARON & LIS, 2004).

Apenas em 1919, a primeira hemaglutinina foi purificada de sementes de feijão de porco (*Canavalia ensiformis*) por James B. Sumner na Universidade de Cornell (Estados Unidos). Em 1936, Sumner e Howell demonstraram que a Concanavalina A (ConA) possuía habilidade de aglutinar células como eritrócitos e fungos e precipitar glicoconjugados em solução. Sumner também demonstrou que esta atividade poderia ser inibida por sacarose, demonstrando pela primeira vez a especificidade dessas lectinas por açúcar. (SHARON; LIS, 2004).

A ideia de que a toxicidade é uma propriedade intrínseca das lectinas foi abandonada no início do século XX, depois que Landsteiner e Raubitscheck, em 1907, relataram pela primeira vez a presença de uma lectina não tóxica nas leguminosas *Phaseolus vulgaris* (feijão comum), *Pisum sativum* (ervilha), *Lens culinaris* (lentilha) e *Vicia sativa*. Após o trabalho de Landsteiner e Raubitscheck, muitas outras hemaglutininas vegetais não tóxicas foram descobertas. Tornou-se, a partir de então, evidente que lectinas estão difundidas no reino vegetal e que a toxicidade atribuída a elas é exceção, e não regra (VAN DAMME *et al.*, 1998). O marco seguinte na história das lectinas vegetais foi a descoberta realizada por Renkonen (1948), Boyd e Reguera (1949) de que algumas hemaglutininas exibem uma clara preferência por eritrócitos de um grupo particular de tipo sanguíneo dentro do sistema ABO. Começou-se a observar que algumas proteínas vegetais obtidas de sementes de plantas podiam reconhecer um grupo específico e aglutiná-lo, em que hemácias do sistema ABO respondiam ao contato com essas hemaglutininas de maneira distinta, umas aglutinando e outras não

(VAN DAMME *et al.*, 1998). A habilidade das aglutininas de plantas em distinguir eritrócitos de diferentes tipos sanguíneos levou Boyd e Shapleigh a proporem em 1954 o termo “lectina” (do latim *legere*, que significa selecionar ou escolher) para nomear essas proteínas (SHARON; LIS, 2004).

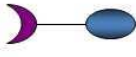
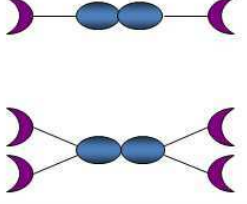
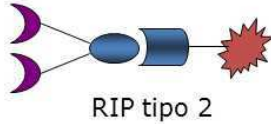

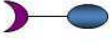
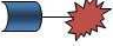
Outro grande marco na história das lectinas ocorreu em 1960, quando Petter C. Nowell demonstrou que a lectina de *Phaseolus vulgaris* possui atividade mitogênica sobre linfócitos. Essa descoberta teve um impacto revolucionário sobre a imunologia, pois até aquele momento acreditava-se que os linfócitos eram células incapazes de se dividir ou de se diferenciar em outros tipos celulares (SHARON; LIS, 2004).

### 1.3 Definições

Lectinas podem ser definidas como proteínas que possuem pelo menos um domínio não catalítico que se liga reversivelmente a um mono/oligossacarídeo específico (PEUMANS; VAN DAMME, 1995).

As lectinas são divididas estruturalmente em quatro grupos, merolectinas, hololectinas, quimerolectinas e superlectinas. As merolectinas consistem de apenas um domínio de ligação a carboidrato sendo, portanto, monovalentes e não têm a capacidade de aglutinar células e/ou precipitar glicoconjugados. As hololectinas são constituídas de domíniosligantes a carboidratos, porém contêm ao menos dois de tais domínios idênticos ou, seja, que se liga a açúcares estruturalmente relacionados, sendo neste grupo que se encontram a maioria das lectinas de planta, que são di ou multivalentes e aglutinam células e/ou precipitam glicoconjugados. As quimerolectinas são proteínas de fusão consistindo de um ou mais domínios de ligação a carboidratos, dispostos de forma sequencial a um domínio não relacionado. Este último domínio pode ter uma atividade enzimática bem definida ou alguma outra atividade biológica que atua independentemente do domínio de ligação a carboidratos. Dependendo do número dos sítios de ligação a carboidratos, as quimerolectinas podem se comportar como merolectinas ou como hololectinas. Finalmente, as superlectinas são aquelas lectinas que se ligam pelo menos a dois tipos de carboidratos estruturalmente diferentes (Figura 1) (PEUMANS; VAN DAMME, 1995). Mais recentemente, lectinas foram definidas como um dos grupos mais importantes de proteínas responsáveis por decifram o glicocódigo (GABIUS, 2000). Elas desempenham um papel central em uma variedade

de eventos biológicos, como reconhecimento de moléculas em células e interações célula-célula (SHARON; LIS, 2004).

Merolectina	Hololectina	Quimerolectina	Superlectina
 Heveína (PDB cod. 1Q9B)	 ConM (PDB cod. 20W4)	 RIP tipo 2	 Lectina da tulipa
 Domínio de ligação a carboidratos  Domínio catalítico			

**Figura 1. Classificação estrutural de lectinas.** Representação esquemática de merolectinas (Heveína;PDB: 1Q9B), hololectinas (ConM; PDB: 20W4), quimerolectinas (Ricina; PDB: 2R2X) e superlectinas, estas últimas sem representantes com estrutura tridimensional elucidada.

#### 1.4 Aspectos gerais de lectinas vegetais

Lectinas de vegetais são as mais bem caracterizadas dentre os grupos de lectinas estudadas. Algumas 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).

A interação de lectinas com carboidratos é bastante específica, tanto quanto a interação entre uma enzima e seu substrato, ou antígenos com anticorpos. As lectinas

ligam-se a açúcares livres ou polissacarídeos, glicoproteínas ou glicolípídeos, que estão livres ou associados a membranas (PUSZTAI, 1991).

Lectinas vegetais apresentam-se distribuídas em muitos grupos botânicos, incluindo mono e dicotiledôneas, musgos e líquens, mas são mais frequentemente estudadas em Leguminosae e Euphorbiaceae (PUSZTAI, 1991). Nos vegetais, as lectinas são encontradas em vários tecidos e apresentam diferentes localizações celulares, bem como diferentes propriedades moleculares. As lectinas são mais abundantes nas sementes, porém podem facilmente ser encontradas em outras partes vegetativas. Em sementes de leguminosas estão prevalentemente nos tecidos cotiledonares (PUSZTAI, 1991).

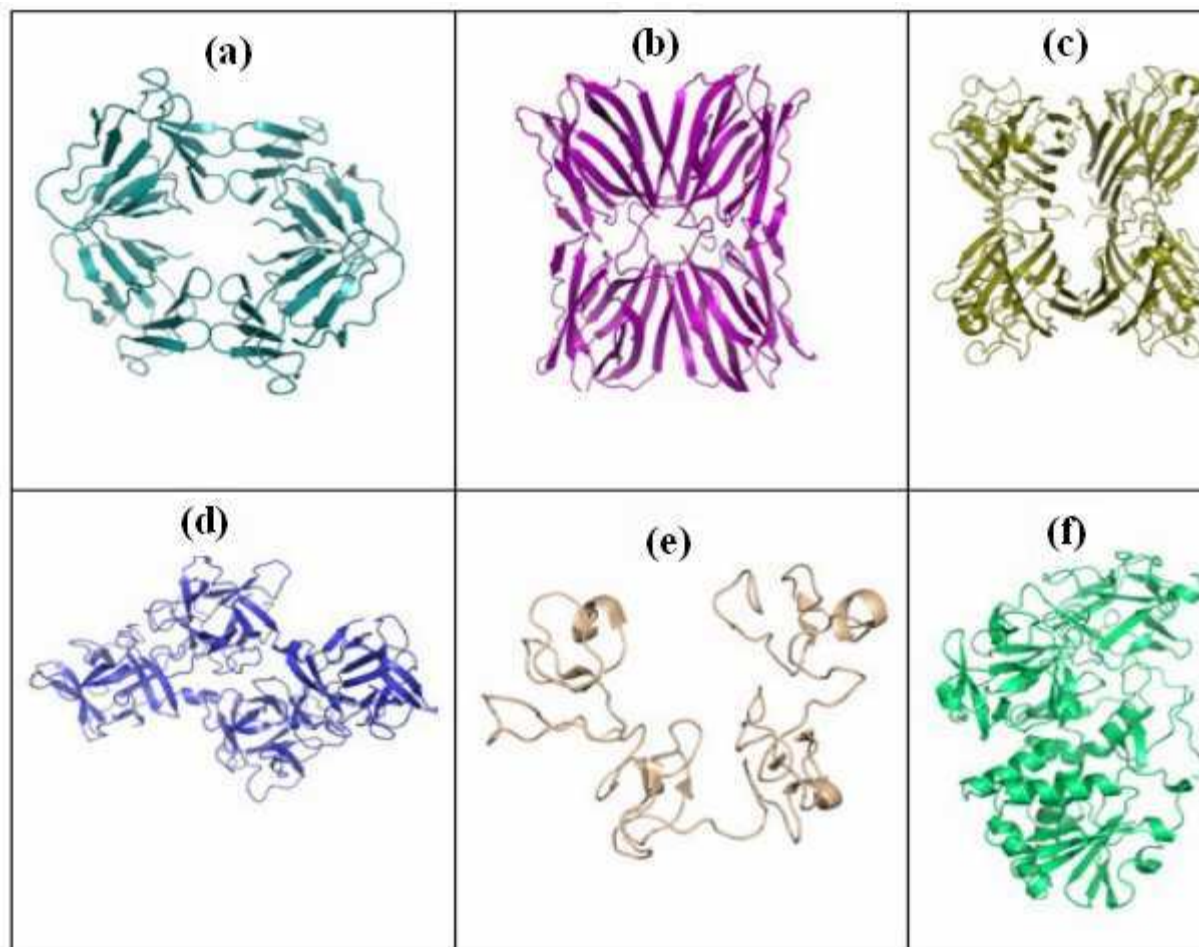
As lectinas de plantas são classificadas em sete grupos de acordo com as semelhanças estruturais e evolutivas. Esses grupos são nomeados de lectinas de leguminosas, lectinas ligantes à quitina, lectinas de monocotiledôneas, proteínas inativadoras de ribossomo tipo 2 (RIP), lectinas de floema de Cucurbitaceae, família da Jacalina e das Amaranthaceae (Figura 2) (VAN DAMME *et al.*, 1998).

As lectinas de leguminosas têm como principal representante a lectina isolada de sementes de *Canavalia ensiformis* (ConA). É uma proteína que requer cátions divalentes em sítio ligante a metal. Cada subunidade possui um íon  $\text{Ca}^{2+}$  e  $\text{Mn}^{2+}$ , que é essencial para a atividade de ligação a carboidrato da lectina. Os aminoácidos envolvidos na ligação desses íons são altamente conservados (VAN DAMME *et al.*, 1998).

As proteínas ligantes à quitina (polímero de *N*-acetil-glicosamina) contêm o chamado domínio heveínico. O termo 'heveína' refere-se a uma pequena proteína de 43 resíduos de aminoácidos encontrada no látex da seringueira (*Hevea brasiliensis*) (WALJUNO *et al.*, 1975). As quitinases classe I, por exemplo, são quimerolectinas compostas de um único domínio heveína ligado a um domínio catalítico com atividade quitinásica. Essas lectinas também são encontradas em Gramineae, Solanaceae, Urticaceae, Papaveraceae e Amaranthaceae (COLLINGE *et al.*, 1993).

A lectina de bulbos de *Galantus nivalis* (GNA) foi a primeira lectina de monocotiledônea ligante de manose a ser isolada e caracterizada (VAN DAMME *et al.*, 1987). Desde então, proteínas similares têm sido encontradas em espécies de

monocotiledôneas, como nas famílias Amalylidaceae, Alliaceae, Araceae, Liliaceae, Orchidaceae e Bromeliaceae (PEUMANS; VAN DAMME *et al.*, 1995).



**Figura 2. Classificação das lectinas vegetais.** (a) Lectina de monocotiledôneas ligantes à manose, aglutinina de *Galanthus nivalis* (GNA PDB: 1MSA) (b) Lectinas relacionadas às Jacalinas, Jacalina (PDB: 1UGW), (c) Lectinas de leguminosas, aglutinina de *Ulex europaeus* (UEA PDB: 1FX5 ), (d) Amarantina, Amarantina (1JLX), (e) Lectina ligante à quitina compostas por domínios heveínicos, Heveína (PDB: 1Q9B) (f) RIP's do tipo II, Abrina (PDB: 2ZR1).

As proteínas inativadoras de ribossomos são quimerolectinas compostas de um polipeptídeo constituído de um domínio adenosina glicosidase (cadeia A) ligado ao domínio ligante a carboidrato (cadeia B) (STIRPE; BATTELLI, 2006). Ambas as cadeias são sintetizadas por uma única molécula precursora a qual é processada após a tradução através da excisão de uma ligação entre as cadeias A e B (VAN DAMME *et al.*, 1998; STIRPE; BATTELLI, 2006). A ricina (*R.communis*) e abrina (*A.*

*preparatorius*) são exemplos de RIP's tipo 2 com potentes atividades biológicas (VAN DAMME *et al.*, 1998).

A Jacalina é uma lectina de sementes da jaca (*Artocarpus integrifolia*) que apresenta especificidade a galactose. Lectinas de Moraceae são consideradas um grupo de lectinas que ocorrem excessivamente nesta particular família de plantas com características da jacalina. Por repetirem-se em lectinas de outras famílias de plantas, muito provavelmente, a família jacalina também é uma superfamília de lectinas que ocorre em diferentes grupos taxonômicos (PEUMANS *et al.*, 2000).

Muitas espécies de Curcubitaceae contêm alta concentração de lectinas que se ligam a oligômeros de GlcNAc (WANG *et al.*, 1994). Todas as lectinas conhecidas do floema de Curcubitaceae mostram um grau de similaridade sequencial entre si, mas não têm qualquer similaridade sequencial com outras lectinas de plantas ou outras proteínas (PEUMANS; VAN DAMME, 1998).

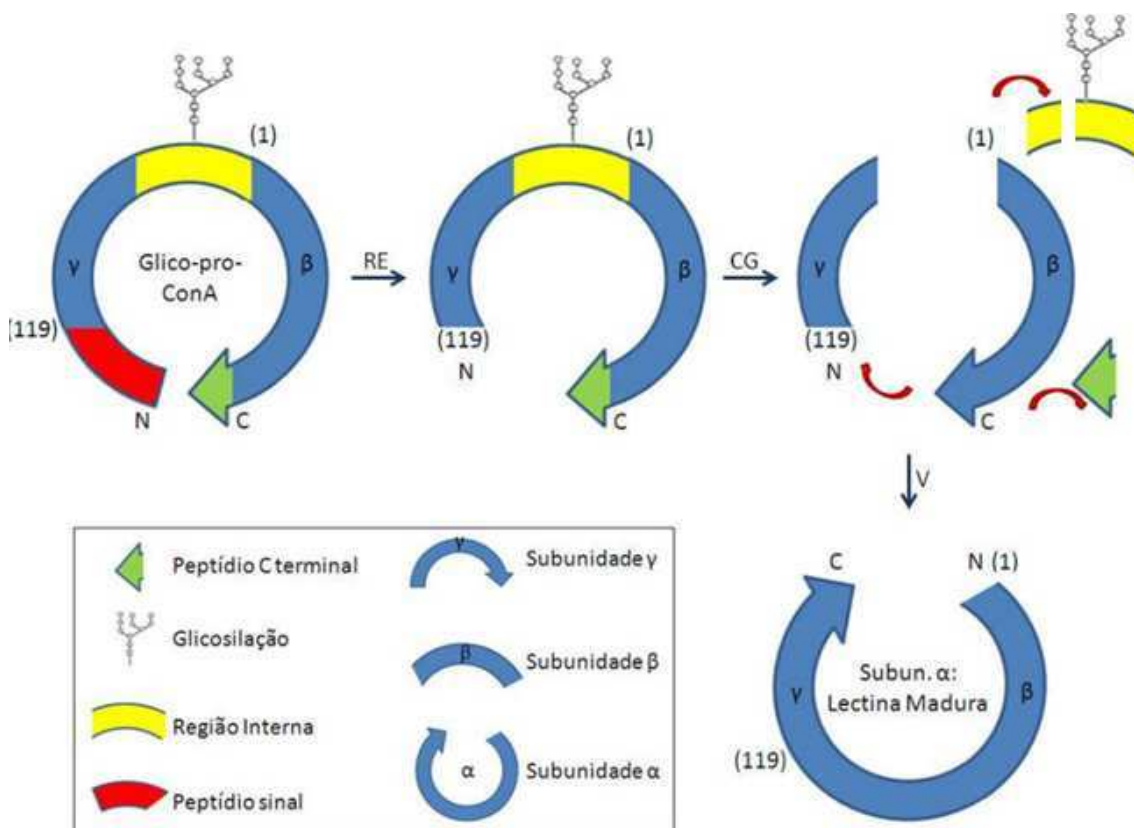
A clonagem molecular e a análise de difração de raios X revelaram que a amarantina, uma lectina de semente de *Amaranthus caudatus*, não se assemelha a qualquer outra lectina de plantas com relação a sua sequência de aminoácidos ou à sua estrutura tridimensional (TRANSUE *et al.*, 1997). Baseado nesses dados, a amarantina é agora considerada como um protótipo da família de lectina amarantina, uma lectina ligante específica de *N*-acetil-galactosamina (VAN DAMME *et al.*, 1998).

### **1.5 Lectinas de leguminosas**

A maioria das lectinas de dicotiledôneas estudadas foi isolada de plantas de tribos da família das leguminosas (SHARON; LIS, 1990). As lectinas de leguminosas constituem uma ampla família de proteínas estreitamente relacionadas, encontradas exclusivamente em espécies representativas da família Leguminosae. Elas estão presentes em sementes, folhas, caules e raízes (SHARON; LIS, 1990; LORIS *et al.*, 1998). Várias lectinas desta família já foram isoladas, sendo a maior parte pertencente a diferentes subordens e espécies, caracterizadas, principalmente, de sementes, em que o percentual de lectina varia de 1 a 10% do total de proteínas solúveis, embora algumas espécies possam apresentar concentrações acima de 50% (SHARON; LIS, 1990). Várias leguminosas possuem duas ou mais diferentes lectinas nas sementes. Apesar de

diferirem em suas especificidades para carboidratos são muito semelhantes em suas propriedades físico-químicas (SHARON; LIS, 1990).

Sabe-se que algumas lectinas de leguminosas, especialmente as da subtribo Diocleinae, são sintetizadas no retículo endoplasmático na forma de pré-pró-lectinas ou pró-lectina glicosilada (Figura 3). A região correspondente ao cDNA contém 29 resíduos de uma sequência sinal, que é removida posteriormente durante o transporte da pré-pró-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 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 um 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 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, cujo alinhamento dos resíduos 1-118 e 119-237 é o inverso de como se apresenta no precursor (SHARON; LIS, 1990; SHARON, 2007).



**Figura 3. 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 terminalda 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  $\gamma$  e  $\beta$  (Adaptado de BENEVIDES, 2011). RE: retículo endoplasmático, CG: complexo de golgi, V: vacúolo.

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 por monômero. Suas subunidades são quase sempre compostas de uma cadeia polipeptídica. Entretanto, algumas lectinas possuem subunidades formadas por duas cadeias polipeptídicas, entre elas as lectinas da tribo Viciae, gêneros *Pisum*, *Vicia*, *Lathyrus*, e *Lens* (ROUGÉ *et al.*, 1987). Essas lectinas são dímeros formados por duas subunidades iguais e cada subunidade é constituída de uma cadeia  $\alpha$  (5 a 7 kDa) e uma  $\beta$  (15 a 19 kDa) mantidas por ligações não-covalentes (SHARON; LIS, 1989). 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).

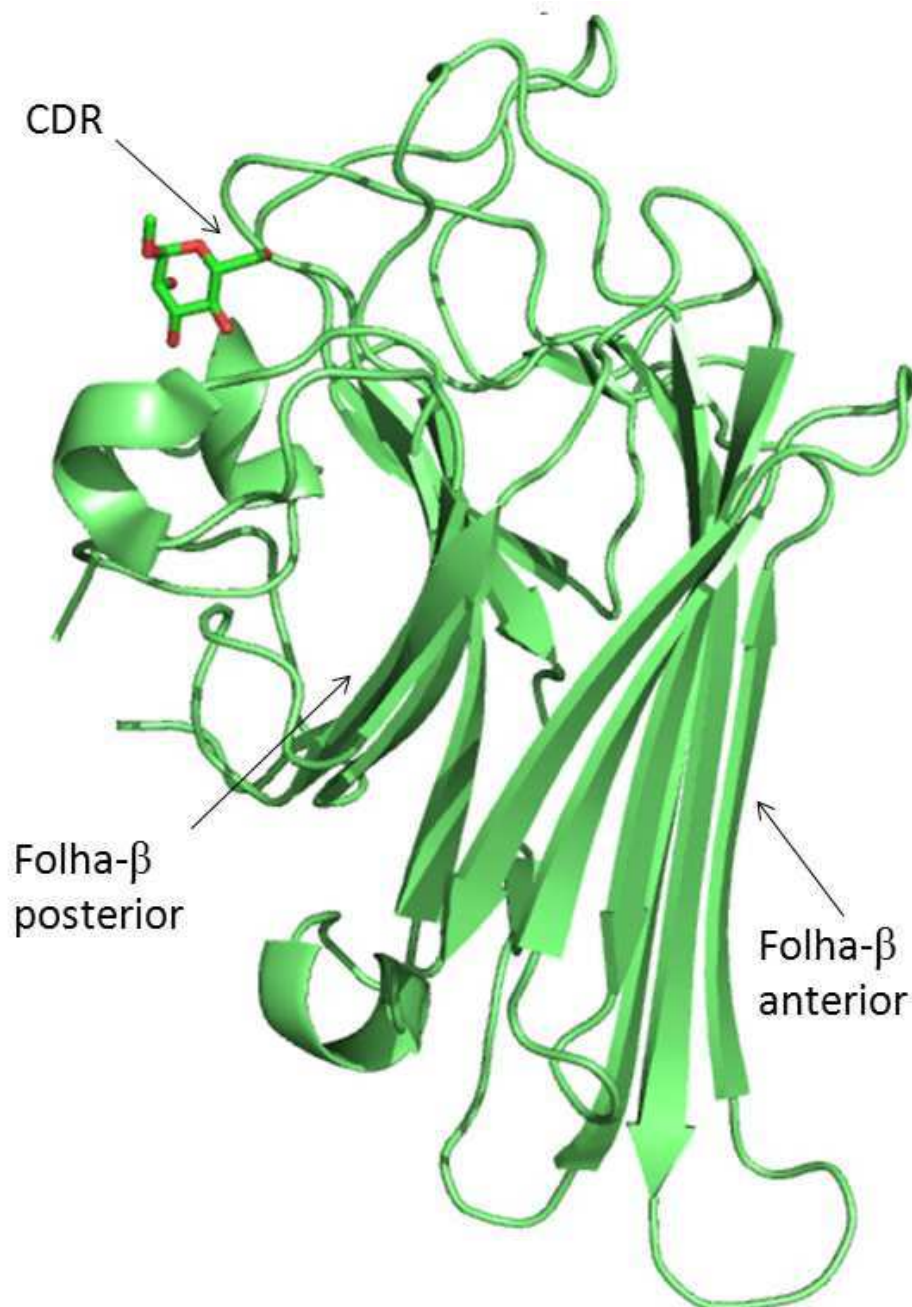
### 1.6 Estruturas tridimensionais das lectinas vegetais

A primeira lectina a ser cristalizada foi a ConA por James B. Sumner que utilizou a técnica para purificar esta proteína em 1919 (SUMNER, 1919). Apesar disso, não foi dada muita importância à sua purificação. Em 1936, Sumner e However reportaram propriedades relevantes dessa proteína que era capaz de aglutinar eritrócitos.



Passaram-se trinta e seis anos até que sua estrutura tridimensional fosse resolvida por cristalografia de raios X em 1972 (EDELMAN *et al.*, 1972; HARDMAN *et al.*, 1972). Subsequentemente, outras lectinas de leguminosas foram cristalizadas e tiveram suas estruturas tridimensionais resolvidas. Hoje, existem vários grupos no mundo trabalhando com purificação e caracterização estrutural de lectinas vegetais.

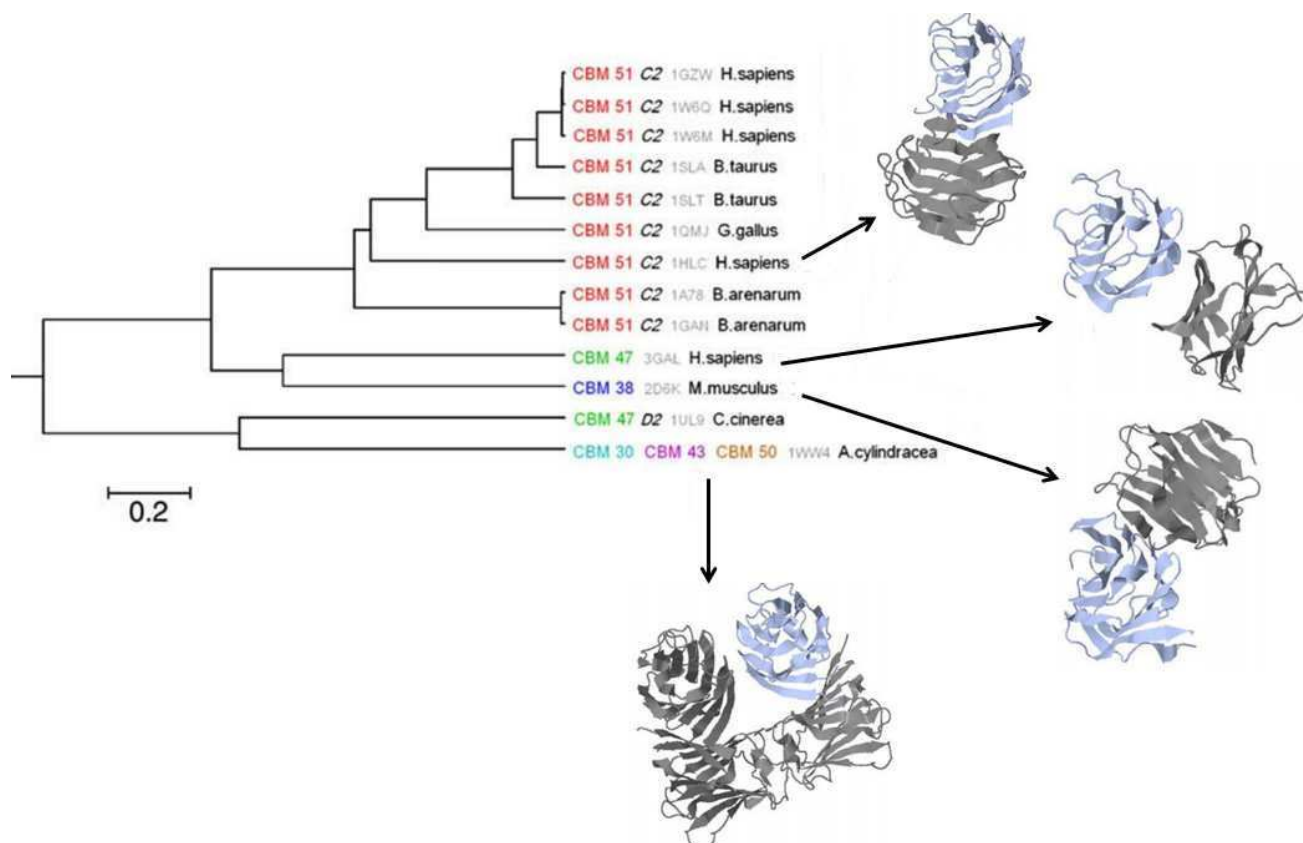
Em geral, nas lectinas de leguminosas, a estrutura do monômero é bem conservada. Ela consiste principalmente de folhas- $\beta$ , que são conectadas por *loops*, em que se encontram cerca de 50% dos aminoácidos. O monômero é formado por três folhas- $\beta$  de fitas antiparalelas: uma folha- $\beta$  composta por seis fitas antiparalelas, a folha- $\beta$  anterior, e outra composta por sete fitas, a folha- $\beta$  frontal (BANERJEE *et al.*, 1996). A principal região hidrofóbica está localizada entre as folhas frontal e anterior. Esse enovelamento terciário é conhecido como motivo “jelly-roll” (Figura 4) (WILLIAMS; WESTHEAD, 2002).



**Figura 4. Monômero típico de uma lectina de leguminosa.** Representação do motivo *jelly-roll* da lectina de *Canavalia brasiliensis* destacando o domínio de reconhecimento a carboidrato (CDR) e folhas- $\beta$  posteriores e anteriores.

O enovelamento acima descrito a partir de folhas betas é bastante comum e parece ter evoluído independentemente várias vezes (LORIS, 2002). Além das lectinas de leguminosas, podem ser citadas as galectinas e pentraxinas, ambas apresentando tal enovelamento, embora, sejam lectinas de animais. Apesar de todas as galectinas serem formadas por folhas- $\beta$ , o enovelamento e o estado de oligomerização em diferentes

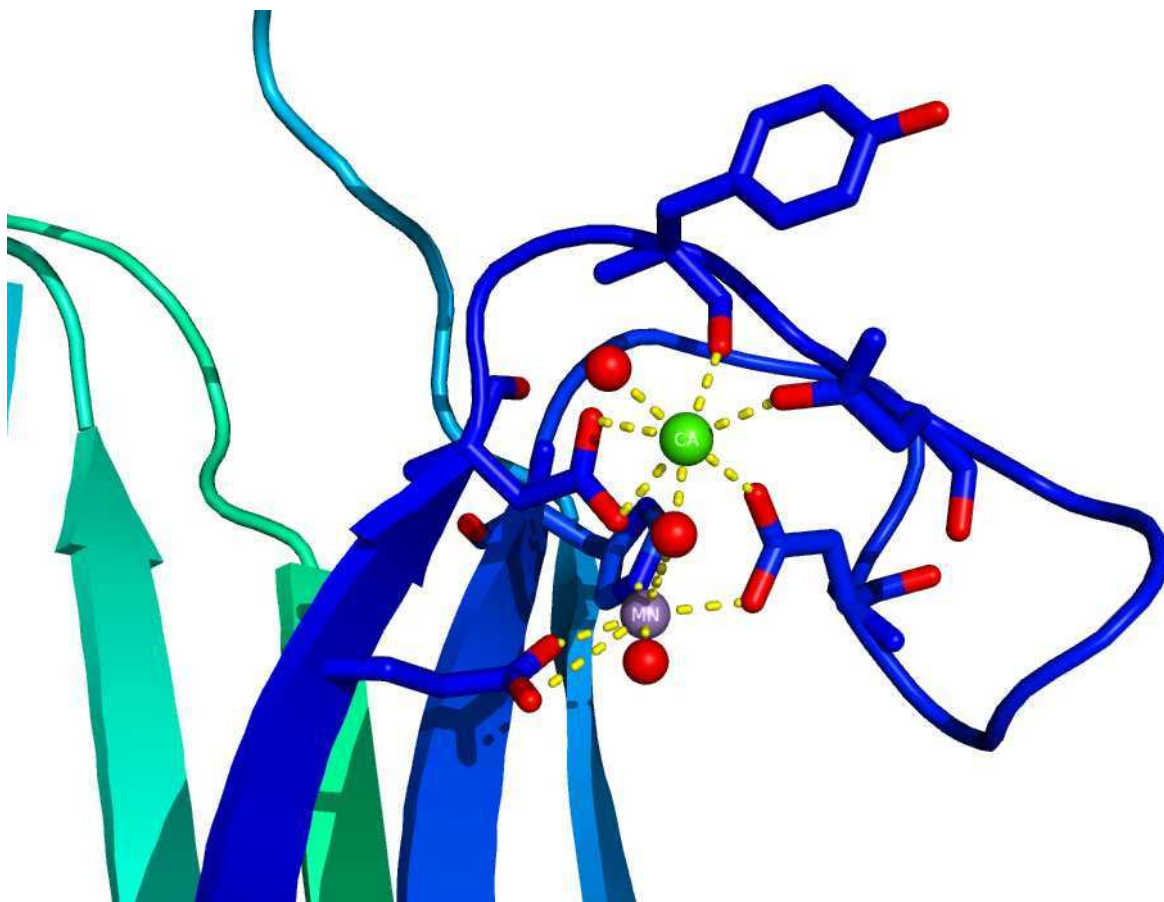
grupos de animais são de extrema importância na variabilidade do reconhecimento de diferentes receptores (Figura 5), e na atividade biológica das galectinas que estão envolvidas na adesão celular (DAYHOFF *et al.*, 2009). Além de lectinas, algumas outras proteínas que não se ligam a carboidratos também apresentam o enovelamento semelhante à lectinas de leguminosas, como exemplos, podem ser citados os inibidores de  $\alpha$ -amilase e uma série de arcelinas (HAMELRYCK, *et al.*, 1996).



**Figura 5. Representação dos diferentes tipos de oligomerização das galectinas.** Árvore filogenética dos diferentes grupos e os estados de oligomerização que as galectinas assumem em cada grupo (Adaptado de DAYHOFF *et al.*, 2009).

Na estrutura das lectinas de leguminosas, o domínio de ligação a carboidrato é estabilizado e ativado por íons de manganês e cálcio. Eles estão a uma distância de aproximadamente 4,5 Å um do outro e estão ligados por dois aspartatos. O evento chave na ativação do DRC pelos íons é a isomerização da ligação peptídica entre alanina 207 e aspartato 208 de trans para cis, cujo aspartato 208 é movido para uma posição de ligação a carboidrato (Figura 6). Na estrutura da ConA desmetalizada, o aspartato 208 se apresenta afastado do DRC (LORIS *et al.*, 1998). A ligação a carboidratos se dá, por interações apolares entre as porções apolares dos açúcares e resíduos de aminoácidos

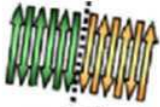
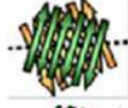
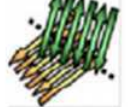
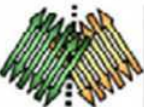


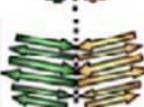
aromáticos (tirosina e fenilalanina), e por pontes de hidrogênio formadas entre as hidroxilas dos carboidratos e com cadeias laterais. Em lectinas de leguminosas o DRC é formado pelos resíduos de aminoácidos: Asn14, Leu99, Tyr100, Asp208 e Arg228. Dentre esses resíduos, há uma região de “loop” (97-102) onde se acredita ser a região que determina a especificidade por monossacarídeos das lectinas de leguminosas, podendo ser específicas a glicose/manose ou a galactose (Loris, *et al.*, 1998). As regiões dos DRCs são conservadas nas lectinas de *Diocleinae*, mas na estrutura da lectina de *Canavalia maritima* (ConM) há uma mutação no resíduo 202, a qual uma serina se encontra no lugar de uma prolina. Esse resíduo, apesar de não estar relacionado diretamente ao DRC, propicia um deslocamento da Tyr12 para dentro do DRC, o que lhe confere um aumento significativo na capacidade da ConM se ligar à dissacarídeos (GADELHA *et al.*, 2005).



**Figura 6. Representação da interação a metais das lectinas de leguminosas.** Sítio de ligação a metais ( $\text{Ca}^{2+}$  e  $\text{Mn}^{2+}$ ) e interações que estabilizam o sítio de ligação ao carboidrato na lectina de *Canavalia ensiformis* (ConA) PDB 1CVN.

A estrutura quaternária das lectinas de leguminosas é composta de duas ou quatro subunidades, cada uma com um único sítio de ligação a carboidratos de mesma especificidade. As subunidades interagem por forças não covalentes. Em relação aos arranjos quaternários, essa família tem sido classificada em nove tipos, compostos por sete interfaces diméricas diferentes (BRINDA *et al.*, 2004; SINHA *et al.*, 2007). Esses sete modos de interação estão representados na Tabela 1. Moreno *et al.*, (2008) relataram um novo modo de interação observado na lectina de *Lotus tetragonolobus* (LTA), que forma um homotetrâmero composto de dois dímeros interagindo de modo X4 que interagem de forma invertida quando comparado com o tetrâmero da lectina de *Griffonia simplicifolia* (GS IV). Dentre os modos de interação, duas formas de associação são encontradas na estrutura quaternária das lectinas da subtribo *Diocleinae*. Os dímeros canônicos ou associação do Tipo II são estabilizados, principalmente, por pontes de hidrogênio e interações hidrofóbicas. A interação entre esses dímeros forma o tetrâmero, sendo essa associação chamada de interface X2. O dímero canônico é caracterizado pela associação lado a lado de duas folhas- $\beta$  anteriores, uma de cada monômero, formando, assim, uma folha- $\beta$  contínua de 12 fitas. A interface do dímero se forma de tal modo que as folhas- $\beta$  ficam face a face, formando estruturas rígidas.

**Tabela 1.** Sete diferentes tipos de interface formados durante a oligomerização de lectinas de leguminosas (Adaptado de Sinha *et al.*, 2007).

Tipos de interface	Exemplo e características	Ilustração
Interface canônica ou tipo II	Presente em lectinas diméricas como PSL e lectinas tetraméricas como ConBr.	
Interface X1	Observado na interface não canônica do tetrâmero de SBA e PHA.	
Interface X2	Observado apenas nos tetrâmeros com a interface não canônica de ConA e ConBr.	
Interface X3 ou “aperto de mão”	Observado nos dímeros de Ecol e WBA.	
Interface X4 ou “costa para costa”	Observado na PNA, GS4 e GS1.	
Interface X5	Observado na GS1, é a forma da interface X4 com um ângulo diferente e orientação mútua.	
Interface não usual	Observado na PNA entre as subunidades 3 e 4.	

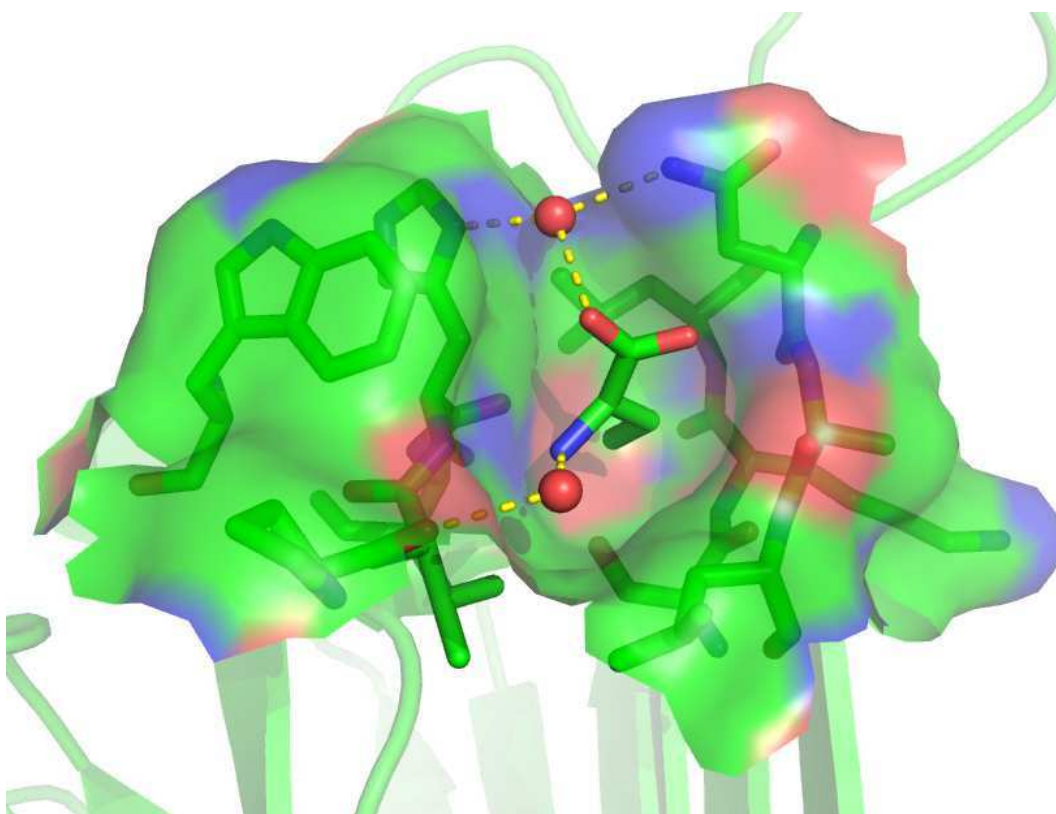
Os DRCs de cada monômero se encontram nos extremos dos dímeros. A estrutura quaternária pode apresentar rotações e/ou translações, se for comparada entre as diversas lectinas da subtribo *Diocleinae*. Por exemplo, no caso da ConBr, os dímeros apresentam uma pequena rotação quando comparado com a ConA e da própria ConA que ao ser complexada com um trimanosídeo apresenta uma pequena rotação nos dois monômeros do dímero comparada com a sua estrutura sem ligante. Isso demonstra que o tetrâmero é uma estrutura dinâmica (LORIS *et al.*, 1998).

Em grande parte das lectinas isoladas e purificadas de sementes da subtribo *Diocleinae*, o equilíbrio dímero-tetrâmero é dependente de pH. Elas são modelos excelentes para o estudo da interação entre monômeros de proteínas diméricas e tetraméricas (SRINIVAS *et al.*, 2001) e vêm sendo estudadas para um melhor entendimento das interações proteína-proteína. Resultados recentes indicam que *loops* da cavidade central são pontos chave na estabilização da associação tetramérica dependente de pH das lectinas de *Diocleinae* (GALLEGO *et al.*, 2004; GALLEGO *et al.*, 2007). Nagano e colaboradores sugeriram que a força iônica nos contatos



estabelecidos por His51, Arg60 e Asp78 podem ser modulados pelo pH, o que levaria a uma dependência na força iônica para a dissociação do tetrâmero em proteínas que sejam dependentes do pH para tetramerização (NAGANO *et al.*, 2008).

Além dos DRC, há também sítios hidrofóbicos que podem interagir com moléculas não polares, como acontece, por exemplo, entre a ConA e o ácido indolacético, um hormônio vegetal (EDELMAN; WANG, 1978). Também podemos encontrar outro sítio hidrofóbico na lectina de sementes de *Canavalia gladiata* (Figura 7) (DELATORRE *et al.*, 2007) e em outras lectinas de leguminosas, por exemplo, na lectina da sementes de *Dioclea rostrata*, *Canavalia brasiliensis*, etc. (OLIVEIRA *et al.*, 2008; BEZERRA *et al.*, 2011), ambas complexadas com o ácido aminobutírico, um aminoácido não proteico.



**Figura 7. Representação da área de superfície de um sítio de ligação a compostos hidrofóbicos.** Sítio de ligação ao aminoácido não-protéico ácido  $\alpha$ -aminobutírico na lectina de *Canavalia gladiata* (PDB:1WUV).

Como descrito até o momento, estudos cristalográficos de lectinas podem ajudar a desvendar o comportamento da molécula em soluções a diferentes valores de pH, a importância de íons ligantes na estrutura da molécula, a região onde se dá a interação a carboidratos (DRC), e também permite mostrar a região onde é possível

moléculas hidrofóbicas se ligarem; além disso, é um modelo chave para o entendimento do mecanismo desempenhado em uma ação biológica. (KOMATH *et al.*, 2006; ROCHA *et al.*, 2011).

### 1.7 Aplicações biológicas

Lectinas possuem diversas aplicabilidades biológicas em várias áreas que vão desde a aplicação básica como simples agentes determinantes da tipagem sanguínea em humanos, a terapias e diagnósticos patológicos. Grande parte das atividades está ligada à capacidade que essas moléculas possuem em reconhecer glicanos na superfície das membranas biológicas e assim modular diferentes respostas celulares. As primeiras aplicações biológicas de lectinas em larga escala foram realizadas na obtenção da tipagem sanguínea e estimulação mitogênica de linfócitos (SHARON; LIS, 2004).

Várias lectinas marcadas radiotivamente e lectinas conjugadas com outras moléculas (biotina, enzimas, etc.) servem como reagentes específicos e sensíveis para a detecção de glicoproteínas separadas em gel de eletroforese e em diagnósticos de doenças como o câncer (NASI; PICARIELLO; FERRANTI, 2009).

O uso de lectinas no diagnóstico e terapia é tido como o maior interesse biológico para essas proteínas. O enorme desenvolvimento que possibilita inibir, ativar ou explorar interações proteína-carboidrato pode trazer vários benefícios para a química medicinal; por exemplo, o padrão de desenvolvimento de glicosilação de proteínas celulares é influenciado por diversas alterações fisiológicas, decorrentes de patologias. Assim, as alterações nos glicanos são usadas como pistas para diagnósticos clínicos. Alterações em glicanos são bem caracterizadas em alguns tipos de câncer como de fígado e pulmão, havendo a expressão de glicanos com fucose com ligação específica do tipo ( $\alpha$ 1-6) (OKYAMA *et al.* 2009; DENNIS *et al.*, 2002; AMBROSI *et al.*, 2005) .

Um dos estudos mais promissores atualmente é o uso de lectinas de diversas fontes como ferramentas para futuras terapias antivirais. Esses estudos baseiam-se na interação entre essas macromoléculas e glicoproteínas do capsídeo viral (BALZARINI, 2007). Uma aplicação direta desses estudos é a patente que utiliza a lectina GNA acoplada a filtros de hemodiálise que são utilizados para diminuir a carga viral de pacientes com Hepatite C (TULLIS *et al.*, 2009).



Atividades farmacológicas são, também, motivos de interesse por pesquisadores. Recentemente, tem crescido o número de artigos que mostram atividades farmacológicas que essas proteínas apresentam, das quais podemos citar: proliferação de linfócitos e indução da produção de interferon  $\gamma$  (BARRAL *et al.*, 1992), estimulação de macrófagos na peritonite e reação inflamatória (RODRIGUEZ *et al.*, 1992), indução do edema de pata e migração celular na peritonite (BENTO *et al.*, 1993) e liberação de histamina por mastócitos na peritonite (GOMES *et al.*, 1994). A tabela 2 descreve algumas das atividades farmacológicas desempenhadas por lectinas vegetais.

**Tabela 2.** Atividades farmacológicas relacionadas a lectinas vegetais.

<b>Atividade</b>	<b>Lectina</b>	<b>Referência</b>
Produção de interferon $\gamma$	ConBr, DGL, ConA, DVL, DVir, Dgui	BARRAL <i>et al.</i> , 1992
Efeito anti inflamatório	DVL, Dgui, Dvir, Cgran	NUNES <i>et al.</i> , 2009
Efeito pró inflamatório	ConBr, CGL, DwL, DVL, CRL I ConA	RANGEL <i>et al.</i> , 2011
Ativação de linfócitos	ConBr, DVL, DGL	BARBOSA <i>et al.</i> , 2001
Produção <i>in vitro</i> de Óxido Nítrico	ConBr, DVL, DGL	ANDRADE <i>et al.</i> , 1999
Efeito vasodilatador	ConM, ConBr	GADELHA <i>et al.</i> , 2005

### 1.8 Descrição Botânica

Leguminosae é a terceira maior família entre as Angiospermas, compreendendo 727 gêneros e cerca de 19.000 espécies (LEWIS *et al.*, 2005). Essa família apresenta três subfamílias monofiléticas: Caesalpinioideae, Mimosoideae e Papilionoideae (CHAPILL, 1995; KAJITA *et al.*, 2001; LEWIS; SCHRIRE, 2003;

WOJCIECHOWSKI, 2003). A subfamília Papilionoideae compreende 28 tribos e cerca de 13.800 espécies, o que a torna a mais numerosa dentre as subfamílias de Leguminosae. A espécie *Camptosema pedicellatum* pertence à tribo Phaseoleae. A espécie *Camptosema pedicellatum* é uma trepadeira herbácea ou semi-lenhosa, arbusto ou liana com folhas uni ou trifolioladas, inflorescência pseudoracemosas, flores papilionóides com pétalas vermelhas, androceu pseudomonodelfo, ovário estipitado e frutos deiscentes (Figura 8).



**Figura 8.** *Camptosema pedicellatum*. (a) folhas e vagens (b) flor.

## 2. OBJETIVOS

### 2.1 Objetivos gerais

Resolver e analisar a estrutura primária por espectrometria de massas e a estrutura tridimensional da lectina de sementes de *Camptosema pedicellatum* por cristalografia de raios X.

### 2.2 Objetivos específicos

- Purificar a lectina de *Camptosema pedicellatum*.
- Determinar a estabilidade estrutural por atividade hemaglutinante e através da inibição da atividade por carboidratos.
- Obter a sequência primária por espectrometria de massas.
- Cristalizar a lectina de CPL.
- Difratar e obter conjuntos de dados de difração da CPL.
- Resolver a estrutura tridimensional
- Analisar a estrutura e discutir as principais diferenças da estrutura da CPL em relação a outras lectinas *Diocleinae*

### 3. MATERIAIS E MÉTODOS

#### 3.1 Purificação da CPL

As sementes de *Camptosema pedicellatum* foram coletadas na chapada do Araripe na cidade do Crato – CE (Brasil). As sementes foram descascadas e submetidas a maceração até se obter um pó fino na forma de farinha, apropriada para o processo de extração protéica. Foi utilizada a cromatografia de afinidade em matriz de Sephadex G-50 onde a concentração de proteínas nas frações coletadas foi monitorada por absorvância em comprimento de onda de 280 nm (MOURA, 2011).

Previamente, a farinha foi submetida à extração protéica na proporção de 1:10 (m/v) em solução de NaCl 0,15 M por 3 horas. O extrato foi aplicado à coluna de Sephadex e deixado em contato com a matriz por 8 horas previamente equilibrada com solução de NaCl 0,15 M contendo CaCl<sub>2</sub> 5 mM e MnCl<sub>2</sub> 5 mM. Após o contato, a fração não retida foi eluída com a mesma solução de equilíbrio e a fração retida foi eluída com NaCl 0,15 M e glicose 0,1 M. A fração retida, referente à lectina, foi dialisada exaustivamente contra água destilada (ou MilliQ) e liofilizada, sendo o material liofilizado utilizado para os passos seguintes.

#### 3.2 Atividade Hemaglutinante

Amostras de sangue de coelhos foram lavadas adicionando-se solução de NaCl 0,9% e, em seguida, centrifugando-se as misturas a 3.000 rpm por 5 minutos, descartando o sobrenadante. Esse procedimento foi repetido por sete vezes, ou até que o sobrenadante remanescente da centrifugação se tornasse totalmente límpido.

A atividade hemaglutinante foi testada nos extratos totais segundo o método descrito por Moreira & Perrone (1977), utilizando-se solução a 2% de hemácias.

Os testes foram realizados em placas de microtitulação com poços com fundo em U, contendo 8 fileiras de 12 poços cada, que foram preenchidos com 50 µL de solução de NaCl 0,15 M e em seguida acrescidos 50 µL do extrato no primeiro poço de cada fileira. A amostra então foi diluída serialmente, com homogeneização e transferência de 50 µL para o poço seguinte até o penúltimo poço da fileira, sendo descartados os 50 µL referentes ao último poço. Após as diluições, foram adicionados 50 µL da suspensão de eritrócitos a 2%. A placa foi mantida em repouso por 1 hora em estufa a 37 °C.

Os títulos de hemaglutinação foram medidos em termos de U.H. (Unidade Hemaglutinante) como a concentração de proteínas remanescentes no último poço de uma diluição seriada ainda capaz de causar hemaglutinação visível, através da seguinte equação exponencial:  $U.H. = 2^n$ ; onde  $n$  corresponde ao número do último tubo da diluição onde foi possível detectar hemaglutinação.

### **3.3 Inibição da atividade hemaglutinante**

O ensaio de inibição da atividade hemaglutinante foi realizado utilizando-se soluções a 0,1 M dos carboidratos: D-glicose, D-galactose, D-manose, D-arabinose, D-frutose, D-xilose, D-fucose, *N*-acetil glicosamina, metil  $\alpha$ -galactopiranosídeo, e lactose. Realizou-se uma diluição seriada das soluções de carboidratos, a qual foi adicionado 50  $\mu$ L do extrato total das sementes, incubando-se a 37°C por 30 minutos.

Passado o tempo de incubação, foram adicionados 100  $\mu$ L de solução a 2% de eritrócitos humanos ou de coelho. Incubou-se a 37°C por 1 hora para que as proteínas pudessem interagir com os açúcares, e decorrido esse período realizou-se uma leitura da atividade hemaglutinante para detectar se houve ou não a inibição da atividade.

Os títulos de inibição foram medidos em termos de concentração inibitória mínima (CIM) como a concentração de carboidratos remanescentes no último poço de uma diluição seriada ainda capaz de causar inibição da atividade hemaglutinante.

### **3.4 Eletroforese em Gel de Poliacrilamida na Presença de SDS**

A avaliação da homogeneidade, o grau de pureza e a massa molecular aparente da amostra obtida, referente à lectina, foi realizada seguindo metodologia descrita por Laemmli (1970). A amostra não reduzida foi diluída em tampão Tris-HCl 0,0625 M, pH 8,3, a uma concentração protéica final de 4 mg/mL. Foi acrescentado ao tampão utilizado, SDS a uma concentração final de 1% e quantidade suficiente de glicerol de modo a deixar a solução densa, além de azul de bromofenol 0,02%, para monitorar a frente da corrida eletroforética. Para a amostra reduzida, foi acrescentado também  $\beta$ -mercaptoetanol e a suspensão foi exposta a calor úmido de 100°C por 10 minutos.

O gel de empilhamento utilizado continha a concentração de 3,5% de poliacrilamida na presença de tampão Tris-HCl 0,5 M, pH 6,8, com SDS 1%. O gel de

separação continha a concentração de 12,5% de poliacrilamida em Tris-HCl 3 M, pH 8,8 com SDS a 1%.

A corrida eletroforética foi realizada a uma variação de voltagem de até 200 V, de potência de até 5 W, com uma amperagem constante de 20 mA. O tampão de corrida utilizado foi Tris 0,025 M-Glicina 0,192 M SDS 0,1%, pH 8,8.

Após a corrida eletroforética, o gel de separação foi submetido a uma fixação em uma solução aquosa contendo 25% metanol e 10% ácido acético, por um período mínimo de 1 hora. O gel, então fixado, foi corado em solução de Coomassie R-250 a 0,05% diluído em metanol, ácido acético e água a uma proporção 1:3,5:8 (v/v/v). O excesso do corante foi removido com lavagens em água destilada quente.

### **3.5 Determinação da massa molecular por espectrometria de massas**

A massa molecular da proteína em estudo foi determinada por ionização do tipo electrospray (ESI-MS) acoplada a um espectrômetro de massas híbrido (Synapt HDMS System – Waters Corp). Foi preparada uma solução da proteína a ser analisada na concentração de 10  $\mu\text{mol}/\mu\text{L}$  em 50% de acetonitrila contendo 0,2% de ácido fórmico. Esta solução foi infundida no sistema em um fluxo de 10  $\mu\text{L}/\text{min}$ . A voltagem do capilar e do cone foram ajustadas para 3 kV e 40 V, respectivamente. A temperatura da fonte foi mantida a 100°C e o fluxo de nitrogênio ajustado para 150 L/h. A aquisição de dados foi realizada pelo software MassLynx v4.1 (Waters Corp, EUA) e o espectro multicarregado foi deconvoluído utilizando técnicas de maximização de entropia (FERRIGE *et al*, 1992).

### **3.6 Digestão protéica em gel e sequenciamento dos peptídeos por espectrometria de massas**

Após eletroforese em gel de 12% de poliacrilamida (SDS-PAGE), a banda referente à lectina da semente de *Camptosema pedicellatum* foi retirada do gel e recortada com auxílio de uma ponteira plástica. O gel contendo a proteína de interesse foi descorado em uma solução de 50% de acetonitrila contendo 25 mM de bicarbonato de amônio, desidratado em 100% de acetonitrila e seco em Speedvac (LabConco). O gel foi então reidratado em uma solução de 50 mM de bicarbonato de amônio contendo as enzimas, tripsina (Promega), quimiotripsina (Sigma) e Glu-C (Sigma) nas proporções de 1:50 (peso/peso; enzima:substrato). Para digestão da amostra com a enzima

termolisina, o gel foi reidratado numa solução de 50mM de Tris –HCl pH 7,5 com NaCl 5mM e CaCl<sub>2</sub> 5mM nas mesmas proporções que as demais enzimas. A reação de digestão permaneceu *overnight* a 37°C, sendo interrompida com a adição de ácido fórmico a 2%.

Os peptídeos oriundos da digestão foram extraídos do gel utilizando-se uma solução de 5% de ácido fórmico em 50% de acetonitrila sob agitação durante 15 minutos. Este procedimento foi repetido 3 vezes, o sobrenadante contendo os peptídeos extraídos foram unidos e concentrados em Speedvac e ressuspensos com 25µL de ácido fórmico 0,1%. Estes peptídeos foram injetados em um sistema nanoAcquity (Waters Corp, EUA) conectado a uma fonte de nano electrospray de um espectrômetro de massas (SYNAPT HDMS – Waters Corp, EUA). A amostra foi aplicada em uma coluna de fase reversa C18 (75 µm x 100 mm) e eluída com um gradiente partindo de 10% a 85% de acetonitrila contendo 0,1% de ácido fórmico. O espectrômetro de massas operou em modo positivo, com a temperatura da fonte de 90°C e a voltagem do capilar de 3,0 kV. Os experimentos de LC-MS/MS foram realizados de acordo com a função DDA (Data Dependent Acquisition – Aquisição Dependente de Dados), onde os íons precursores com carga entre +2 e +3 foram selecionados para análise de MS/MS sendo fragmentados através de CID (Collision Induced Dissociation – Dissociação induzida por colisão). Os dados foram coletados, processados e analisados utilizando-se o programa MassLynx v4.1 (Waters Corp) e ProteinLynx v2.4 (Waters Corp). Os peptídeos com sequência de aminoácidos comuns a outras proteínas foram identificados por buscas em banco de dados utilizando ferramenta de pesquisa por padrão de fragmentação dos peptídeos.

Foram utilizados os programas ProteinLynx 2.4 (Waters Corp) e MASCOT (Matrix Science). Os demais peptídeos tiveram suas sequências determinadas através da interpretação manual dos espectros de fragmentação (sequenciamento *De novo*).

### 3.7 Cristalização

A proteína liofilizada, previamente purificada, foi suspensa em água Milli-Q na concentração final de 10 mg.mL<sup>-1</sup>. A suspensão da proteína foi então centrifugada à 9.000 rpm por 10 min e o sobrenadante foi utilizado para os passos futuros. Para experimentos com o ligante, a proteína em solução foi encubada com X-Man na concentração final de 5 mM por 5 horas antes dos experimentos de

cristalização. A proteína foi então submetida ao “screen” de cristalização, utilizando o método da matriz esparsa inicialmente descrito por Jancarik e Kim (1991). O kit utilizado foi o “crystal screen 1” (Hampton Research) onde as variáveis iniciais foram pH, força iônica e precipitantes. O método utilizado foi o de difusão de vapor em gota suspensa utilizando placas de cristalização Linbro<sup>®</sup> de 24 poços. Foram colocados em cada poço da placa de cristalização 300 µL da solução do kit (condição de cristalização) e a gota foi composta por 2 µL da solução de proteína e 2 µL da condição de cristalização. O poço foi então coberto lamínula siliconizadas utilizando graxa de silicone para vedação, as placas forma deixadas em repouso a temperatura de 20 °C. Após a obtenção de cristais, foi feita a otimização dessa condição de cristalização variando a concentração de precipitante e o pH (7,0 a 8,5) 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 (monocristal, tamanho, espalhamento) para que ele pudesse ser difratado quando submetido ao feixe de raios X.

### 3.8 Difração e coleta de dados

Os dados de difração de raios X foram coletados a temperatura de 100 K. Para evitar formação de gelo os cristais foram mergulhados em uma solução crioprotetora composta de 70% da solução de cristalização e 30% de polietileno glicol 400. O cristal laçado com o loop foi alinhado manualmente utilizando-se um goniômetro e então submetido à coleta de dados na estação experimental MX1 do Laboratório Nacional de Luz Sincrotron (Campinas-SP), usando comprimentos de onda de 1,42 Å e detector utilizado foi CCD 165 mm (Mar Research). Para determinar a estratégia de coleta, foram coletadas inicialmente 4 imagens em diferentes ângulos *Phi* ( $\phi$ ) para definir os parâmetros de cela e a mosaicidade com o intuito de calcular o melhor ângulo de oscilação através da equação 1:

$$(1) \quad \theta = \left( \text{Arctg} \frac{\text{Resolução}}{\text{Maior parâmetro de cela}} \right) - \text{Mosaicidade}$$

Um conjunto de 180 imagens foi coletado com oscilação de 1°, totalizando 180°. Os dados foram processados, indexados e integrados utilizando-se o programa MOSFLM (LESLIE, 1992) e as intensidades reduzidas utilizando-se o programa SCALA (EVANS, 1997). Todos os programas utilizados fazem parte do pacote CCP4



(Collaborative Computational Project, Number 4, 1994). No processamento são observadas as intensidades de cada ponto de difração e os índices de Miller, que são números inteiros utilizados para identificar os pontos de difração no espaço recíproco.

### 3.9 Substituição molecular e refinamento

Para solucionar o problema da fase foi adotado como método a substituição molecular usando o programa MolRep (VARGIN; TEPLYAKOV, 1997). As coordenadas atômicas usadas como modelo foram obtidas da lectina recombinante de *Dioclea grandiflora* complexada com 5-bromo-4-cloro-3-indol- $\alpha$ -D-manose (PDB code 2JEC) (NAGANO *et al.*, 2008), porém tanto o ligante quanto as moléculas de água foram removidos do arquivo de coordenadas. Após a utilização de operações de simetria de rotação e translação obteve-se a melhor solução de acordo com os parâmetros *Rfactor* e o coeficiente de correlação, admitindo como modelo com qualidade a ser refinado os resultados que apresentaram valores de *Rfactor* e coeficiente de correlação, (equação 2 e 3 respectivamente) acima de 40% e 60%, respectivamente.

$$(2) \quad R_{factor} = \frac{|F_{obs}| - |F_{calc}|}{|F_{obs}|}$$

$$(3) \quad CC = \frac{\Sigma ab - (\Sigma a \Sigma b)/N}{\{ [\Sigma a^2 - (\Sigma a)^2/N ]^{1/2} \} \{ [\Sigma b^2 - (\Sigma b)^2/N ]^{1/2} \}}$$

Onde  $a = |F_{obs}|$  e  $b = |F_{calc}|$

Inicialmente a estrutura foi submetida a 10 ciclos de refinamento de corpo rígido usando a opção *twin*, para cristais geminados, após os resultados obtidos os ciclos foram reduzidos para 6 devido este apresentar a melhor solução para os parâmetros avaliados, *Rfactor*, *Rfree* e FOM. Posteriormente, os dados foram submetidos a 15 ciclos de refinamento de restrições, sendo estes reduzido apenas a 5 ciclos por apresentar melhores soluções para os mesmos parametros, usando o programa REFMAC5

(Collaborative Computational Project, Number 4, 1994). Os dados das coordenadas da estrutura gerada no refinamento foram analisados no programa Coot, onde foi feito refinamento de posição com mudança do posicionamento das cadeias laterais orientadas para dentro das densidades eletrônicas, visualizadas no mapa (2Fo-Fc) a  $1\sigma$ . Também foram realizadas deleção de resíduos de aminoácidos a partir do modelo utilizado na construção da estrutura. O mapa de densidade eletrônica não explicada (Fo-Fc) foi visualizado a  $2,5\sigma$ , e uma varredura foi realizada no intuito de encontrar densidades não explicadas. Após detecção da densidade que correspondia ao carboidrato ligante, as coordenadas do ligante foram adicionadas nas coordenadas da estrutura e posteriormente o ligante foi posicionado dentro da densidade eletrônica. Por fim, moléculas de água foram adicionadas ao mapa não explicado (Fo-Fc) visualizado a  $2,0\sigma$ , seguido da análise individual e posicionamento de cada molécula de água dentro da nuvem eletrônica, finalizando com mais 3 ciclos de refinamento de restrição na estrutura.

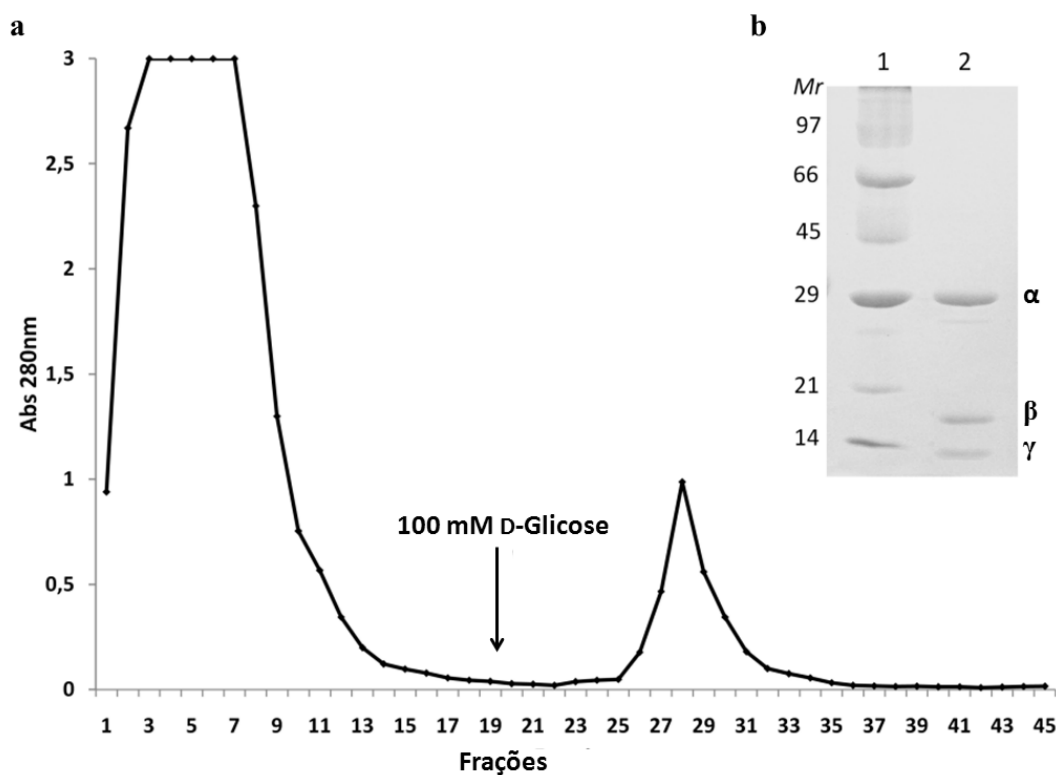
### 3.10 Validação e análise da estrutura

A qualidade estereoquímica da estrutura da CPL avaliadas pelo gráfico de de Ramachandran a partir do programa PROCHECK do pacote CCP4 e o programa online MOLPROBITY (DAVIS *et al.* 2007) que busca os valores das rotações phi ( $\varphi$ ) e psi ( $\psi$ ) do carbono- $\alpha$  dos resíduos de aminoácidos. Os fatores de vibração térmica também foram utilizados para avaliar a correção no posicionamento dos aminoácidos, dos ligantes e das moléculas de água. As interações de Van der Waals, contatos polares e hidrofóbicos foram analisados no programa CONTACT do pacote CCP4 com corte de distâncias de  $3,5\text{ \AA}$ ,  $3,5\text{ \AA}$  e  $5,0\text{ \AA}$ , respectivamente. Todas as figuras e sobreposições de estruturas foram realizadas pelo programa PyMOL (DELANO, 2002).

## 4. RESULTADOS E DISCUSSÃO

### 4.1 Purificação

A CPL foi purificada por cromatografia de afinidade em coluna de Sephadex G-50 a qual foi retida pelas ligações cruzadas do gel de dextrose e eluída por competição com com 0,1 M de glicose como descrito por MOURA (2007). Após a eluição o cromatograma apresentou dois picos, sendo o primeiro correspondente à fração não retida e o segundo à fração retida (Figura 9a). A fração retida corresponde à lectina purificada da semente de *Camptosema pedicellatum*.



**Figura 9. Purificação da CPL.** (a) Perfil cromatográfico da CPL em coluna Sephadex G-50 equilibrada com NaCl 0,15 M contendo 5 mM de CaCl<sub>2</sub> e MnCl<sub>2</sub>, eluída com 100 mM de D-Glicose contendo NaCl 0,15 M (b) SDS-PAGE. Linha 1: marcadores de massa molecular (fosforilase b, 97 kDa; albumina sérica bovina, 66 kDa; ovalbumina, 45 kDa; anidrase carbônica, 29 kDa; inibidor de tripsina, 20,1kDa and  $\alpha$ -lactoalbumina, 14,4 kDa) linha 2: CPL.

Várias outras lectinas da subtribo *Diocleinae* já foram isoladas por cromatografia de afinidade em Sephadex G-50. Essa matriz é caracterizada por separar

moléculas de acordo com o tamanho da proteína, mas, por ser composta basicamente de moléculas de glicose foi utilizada como matriz de afinidade para isolamento dessas lectinas específicas por glicose/manose. Lectinas da subtribo *Diocleinae*, como é o caso da lectina da espécie de *Camptosema pedicellatum*, apresentam perfil cromatográfico bastante semelhantes (MOREIRA; CAVADA, 1984; VASCONCELOS *et al.*, 1991; OLIVEIRA *et al.*, 1991; MOREIRA *et al.*, 1996; CAVADA *et al.*, 1996a; CAVADA *et al.*, 1996b; MOREIRA *et al.*, 1997; CALVETE *et al.*, 1999; CECCATTO *et al.*, 2002). O primeiro experimento de isolamento de uma lectina por cromatografia de afinidade foi realizado pelo pesquisador Irwin J. Goldstein que purificou a lectina de sementes de *Canavalia ensiformis*, a ConA, por cromatografia de afinidade em gel Sephadex a partir do protocolo de extração descrito por James B. Sumner (GOLDSTEIN, 2010).

As atividades hemaglutinantes da CPL contra eritrócitos de coelho nativo e tratado com enzimas proteolíticas foram inibidas por D-manose, D-glicose e  $\alpha$ -metil-D-manopiranosídeo, porém, não foram inibidas por D-galactose, D-arabinose, D-frutose, D-fucose, D-xilose, *N*-acetil-D-glicosamina,  $\alpha$ -metil-D-galactopiranosídeo e lactose (Tabela 3). Similar especificidade de ligação a carboidrato tem sido reportada para outras lectinas de *Diocleinae* (CAVADA *et al.*, 2001; RAMOS *et al.*, 1996)

**Tabela 3.** Efeito inibidor de monossacarídeos e dissacarídeos na atividade hemaglutinante da CPL.

Carboidrato	CIM (mM)
D-manose	12,5
D-glicose	50
D-galactose	NI**
D-arabinose	NI
D-frutose	NI
D-fucose	NI
$\alpha$ -metil-D-manopiranosídeo	3,16
D-xilose	NI
<i>N</i> -acetil-D-glicosamina	NI
$\alpha$ -metil-D-galactopiranosídeos	NI
Lactose	NI

\*CIM Concentração inibitória mínima. \*\* Carboidrato não inibiu a uma concentração máxima de 100 mM.

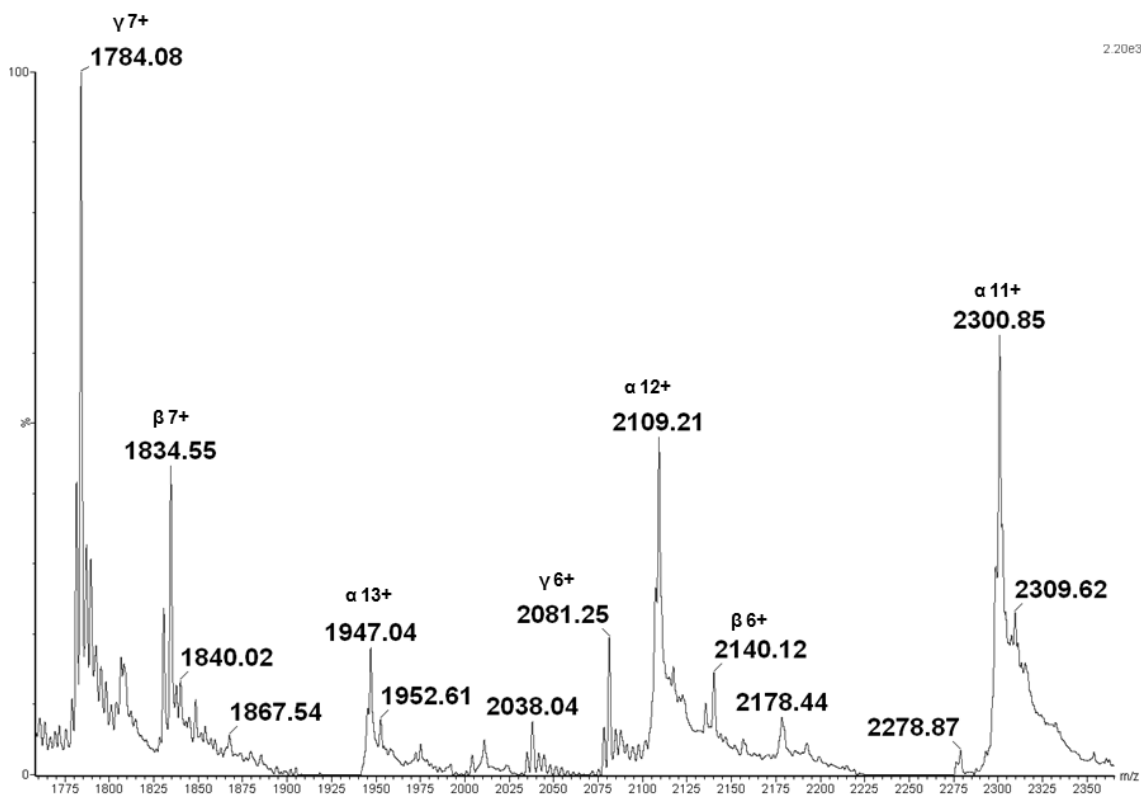
Contudo, a CPL mostrou maior afinidade a  $\alpha$ -metil-D-manopiranosídeo (3,16 mM) que D-manose (12,5 mM), indicando que uma substituição hidrofóbica em C1 aumenta fortemente a ligação do carboidrato a lectinas, possivelmente por permitir um aumento nas interações hidrofóbicas no domínio de reconhecimento a carboidrato (CAVADA *et al.*, 2001).

A eletroforese em gel de poliacrilamida na presença de SDS confirmou a homogeneidade da amostra, revelando uma maior banda de aproximadamente 30 kDa (cadeia  $\alpha$ ) e duas bandas menores correspondendo às cadeias ( $\beta$  e  $\gamma$ ) de aproximadamente 16 kDa e 13 kDa respectivamente, essas bandas são característico das lectinas da subtribo *Diocleinae* (Calvete *et al.*, 1999). A presença de tais fragmentos sugere que a CPL passa pelo mesmo processamento pós-traducional chamado de permutação circular descrito primeiramente por Carrington *et al.*, (1985). Esse processo, citado na introdução, envolve uma clivagem proteolítica em uma região interna de um precursor gerando os fragmentos  $\beta$  e  $\gamma$ . Esses fragmentos são então religados em uma ordem inversa àquela no qual o precursor foi sintetizado, produzindo a cadeia  $\alpha$ . Como a religação não é 100% eficiente, podemos observar, através da eletroforese, que mesmo na fração purificada os fragmentos estão presentes (Figura 9b).

A cadeia  $\alpha$  se mostra mais evidente, mas é possível a visualização clara de  $\beta$  e  $\gamma$ . Perfis de proteínas com alto peso molecular podem ser observados, podendo tais perfis representar oligomerizações das cadeias  $\alpha$  ou o precursor que não foi processado (pre-pró-proteína). O perfil eletroforético da CPL é bastante semelhante a proteínas da subtribo *Diocleinae*, podendo ser caracterizada como uma lectina do tipo ConA.

#### **4.2 Determinação de massa por espectrometria de massas e análise da sequência primária.**

A massa da CPL foi confirmada por análise de espectrometria de massas com ionização do tipo eletrospray, cadeia  $\alpha$  com massa de  $25,298 \pm 2$  Da, cadeia  $\beta$  com  $12,835 \pm 2$  Da e  $\gamma$  com  $12,481 \pm 2$  Da (Figura 10).



**Figura 10. Análise por espectrometria de massas da CPL.** Íons com  $m/z$  2300,85 e estado de carga 11+ corresponde a cadeia  $\alpha$  (25.297 Da), íons com  $m/z$  1834,55 e estado de carga 7+ corresponde a cadeia  $\beta$  (12.835 Da) e os íons com  $m/z$  2081,25 e estado de carga 6+ corresponde a cadeia  $\gamma$  (12.480 Da).

A sequência primária da CPL foi determinada por *overlap* de 27 peptídeos digeridos e sequenciados por espectrometria de massas em *tandem* resultando em 236 resíduos de aminoácidos (cadeia  $\beta$  1-118, cadeia  $\gamma$  119-236) (Figura 11).

```

1  ADTIVAVELD TYPNTDIGDP NYQHIGINIK SIRSKATTRW NVQDGKVGTA HISYNSVAKR
|-----β-chain-----|
|-----T1-----|          |--T2--|-----T3-----|
|-----Q1-----|          |-----Q2-----|
|--E1--|

61  LSAIVSYPPG SSATVSYDVD LNNILPEWVR VGLSASTGVY KETNTILSWS FTSKLTNST
|-----β-chain-----||-
|-----T4-----||-----T5-----||-----T6-----| |---
|-----Q3-----||-----Q4-----|          |--Q5---| |---Q6---
|-----E2-----|

121 ADAQSLHFTF NQFSQSPKDL ILQGDASTDS DGNLQLTRVS NGSPQNSVSG RALYYAPVHV
-----γ-chain-----
-----T7-----| |-----T8-----| |-----T9-----| |---T10-
-----| |---Q7---| |-----Q8-----|
|-----Q9-----|

181 WDKSAVASF DATFTFLIKS PDSDPADGIA FFIANTDSSI PHGSGGRLLG LFPDAN
-----γ-chain-----|
--| |-----T11-----| |-----T12-----| |---T13---|
|-----Q10-----| |-----Q11-----| |-----Q12-----|

```

**Figura 11. Sequência de aminoácidos da CPL.** A sequência de aminoácidos da CPL montada por sobreposição do peptídeos gerados por clivagem com tripsina (T), quimiotripsina (Q) e endopetidase Glu-C (E).

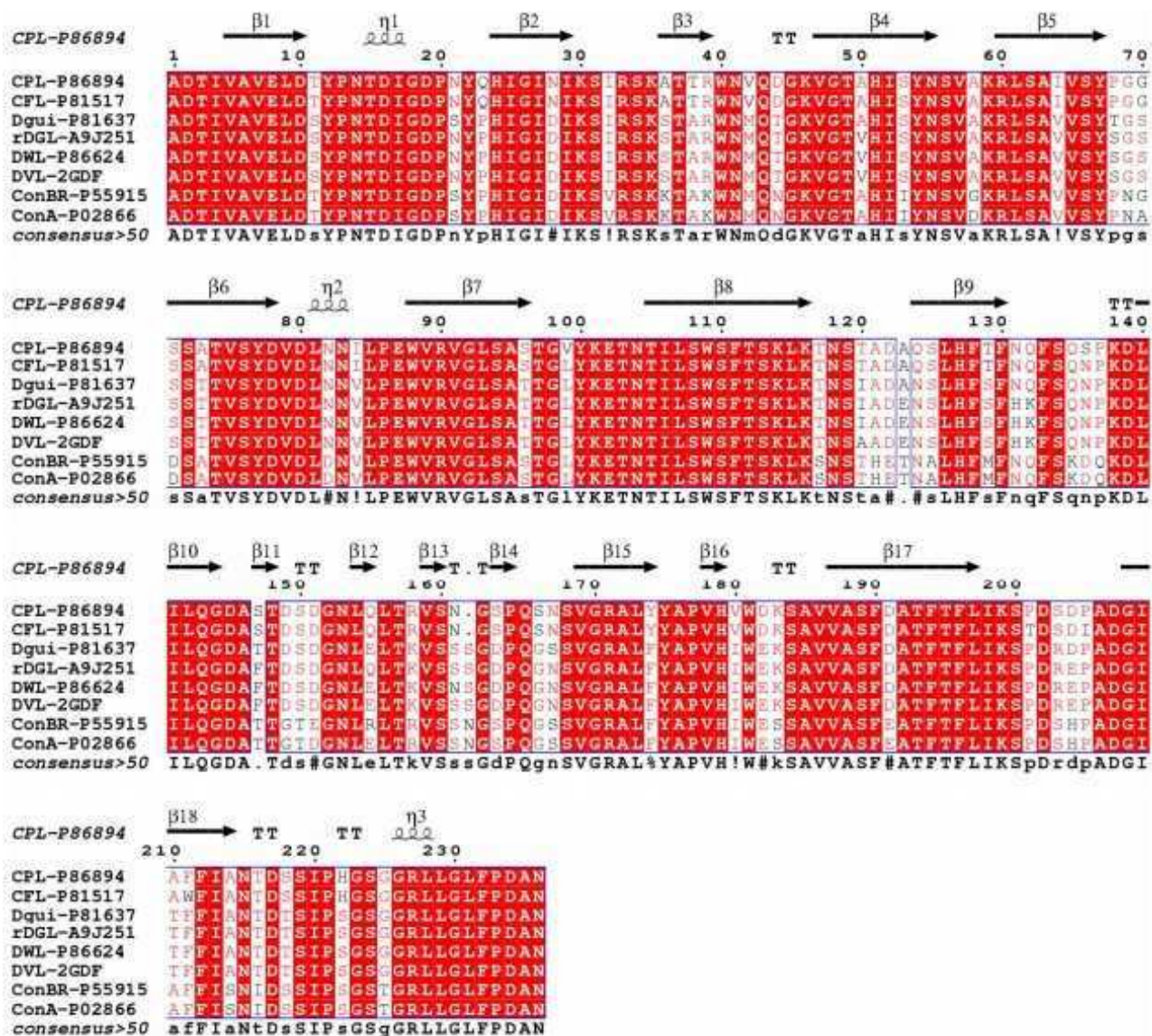
A média das massas moleculares dos isótopos calculada para a cadeia  $\alpha$  (25,297 Da) e seus fragmentos derivados cadeias  $\beta$  e  $\gamma$  (12,835 Da and 12,480 Da) respectivamente, estão de acordo com a massa determinada experimentalmente. A (Tabela 4) mostra todas as sequências dos peptídeos e suas respectivas massas moleculares. O ponto isoelétrico teórico baseado nos dados da sequência primária foi de 5,4. Os dados da sequência primária da CPL está depositada no banco de dados do UniProt com o número de acesso: P86894.

**Tabela 4.** Sequência de peptídeos da CPL e suas respectivas massas moleculares. As enzimas tripsina, quimiotripsina e endopeptidase Glu-C estão representadas por T, Q e E respectivamente.

PEPTÍDEO	MASSA EXPERIMENTAL (Da)	SEQUÊNCIA
T1	3298,6201	ADTIVAVELDTYPNTDIGDPNYQHIGINIK
T2	845,3484	WNVQDGK
T3	1345,6206	VGTAHISYNSVAK
T4	3221,5845	LSAIVSYPPGGSSATVSYDVDLNNILPEWVR
T5	1080,5006	VGLSASTGVYK
T6	1512,7410	ETNTILSWSFTSK
T7	2454,8164	TNSTADAQSLHFTFNQFSQSPK
T8	2131,0090	DLILQGDASTDSDGNLQLTR
T9	1287,5410	VSNGSPQSNVGR
T10	1460,7125	ALYYAPVHVWDK
T11	1715,9370	SAVVASFDFATFTFLIK
T12	2786,9766	SPDSDPADGIAFFIANTDSSIPHGSGGR
T13	958,5063	LLGFPDAN
Q1	3573,7866	DTYPNTDIGDPNYQHIGINIKSIRSKATTRW
Q2	1487,6444	NVQDGKVGTAHISY
Q3	1544,8213	SAIVSYPPGGSSATVSY
Q4	1850,9443	DVDLNNILPEWVGL
Q5	1090,3043	KETNTILSW
Q6	1847,8566	TSKLTNSTADAQSLHF
Q7	1162,5244	NQFSQSPKDL
Q8	3356,6765	ILQGDASTDSDGNLQLTRVSNGSPQSNVGRAL
Q9	1891,8994	TKVSNSGSPQGNVGRAL
Q10	1356,6044	DKSAVVASFDFATF
Q11	1431,6642	IKSPDSDPADGIAF
Q12	1840,7765	FIANTDSSIPHGSGGRL
E1	816,9266	ADTIVAVE
E2	1651,8729	WVRVGLSASTGVYKE

A estrutura primária da CPL apresenta similaridade com outras lectinas da subtribo *Diocleinae* tais como: *Dioclea guianensis* (Dgui) (86%) (SwissProt código de acesso: P81637), *D. grandiflora* (DGL) (85%) (SwissProt código de acesso: A9J251), *D. violacea* (DVL) (81%) (PDB código de acesso: 2GDF), *D. wilsonii* (DwL) (85%) (SwissProt código de acesso: P86624), *Canavalia ensiformis* (ConA) (81%) (SwissProt código de acesso: P02866), *C. brasiliensis* (ConBr) (81%) (SwissProt código de acesso: P55915) e *Cratylia floribunda* (CFL) (97%) (SwissProt código de acesso: P81517). CFL difere da CPL por apenas 5 resíduos: 99 (V/L), 136 (S/N), 201 (P/T), 205 (P/I) and 211 (F/W) (Figura 12).



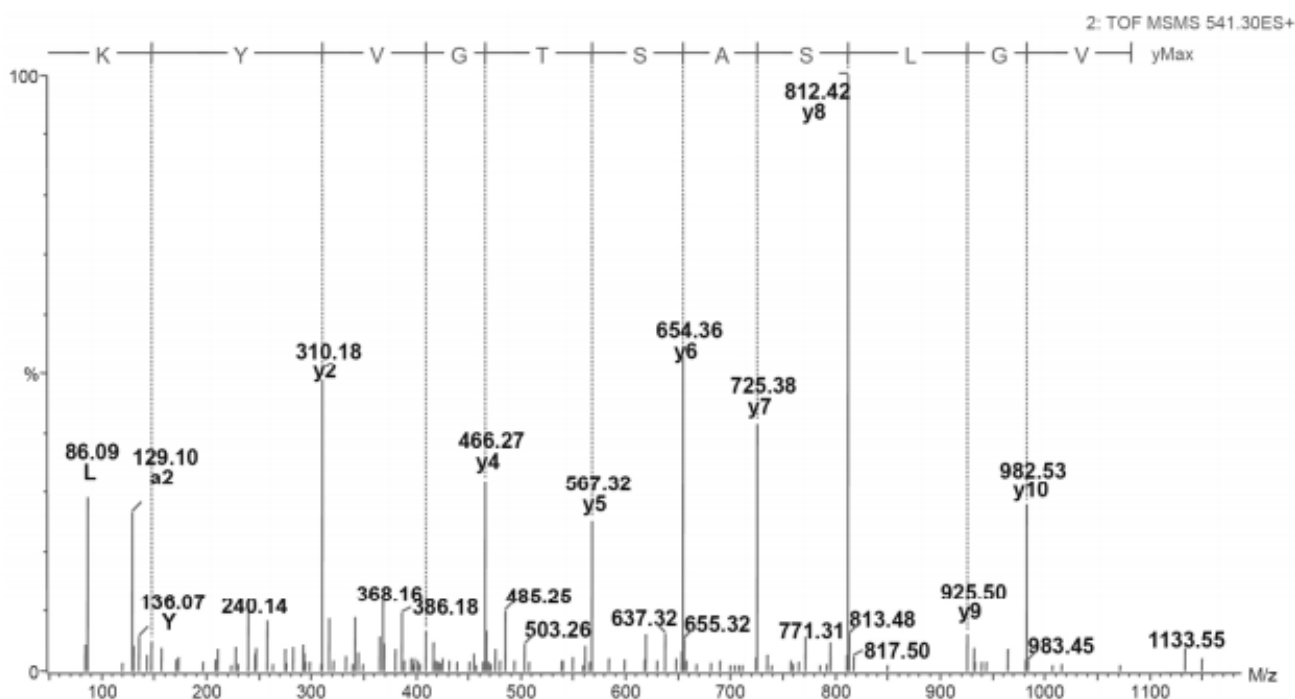


**Figura 12.** Alinhamento múltiplo de sequências de lectinas ConA-like. Alinhamento da CPL, CFL, Dgui, rDGL, DWL, DVL, ConBr e ConA revela alta similaridade na sequência primária e secundária. Setas e espirais na figura correspondem as estruturas secundárias fitas  $\beta$  e  $\alpha$  hélices respectivamente.

Algumas lectinas da subtribo *Diocleinae* exibem equilíbrio dímero-tetrâmero pH-dependente. A relevância fisiológica deste fenômeno ainda é pouco conhecida; contudo, o equilíbrio dímero-tetrâmero pH-dependente tem sido reportado em outros trabalhos como um considerável fator no reconhecimento do oligossacarídeo e atividade biológica (NAGANO *et al.*, 2008). Estes estudos sugerem que os resíduos 123 (A/E), 131 (N/H) e 132(Q/K) apresentam um importante papel nas propriedades do equilíbrio dímero-tetrâmero. Dgui (WAH *et al.*, 2001) e CFL (GALLEGO DEL SOL *et*

*al.*, 2007) exibem equilíbrio da transição dímero-tetrâmero pH-dependente devido a presença de que resíduos nas posições descritas, CPL apresenta os resíduos Ala123, Asn131 e Gln132, podendo estar envolvidos também na oligomerização da lectina em diferentes pH.

A diferença mais relevante na estrutura primária da CPL em relação a outras lectinas *Diocleinae* está na substituição da Val99 por Leu99. A substituição conservativa deste resíduo foi confirmada por análises dos peptídeos T5 e E2 por MS/MS (Figura 13).



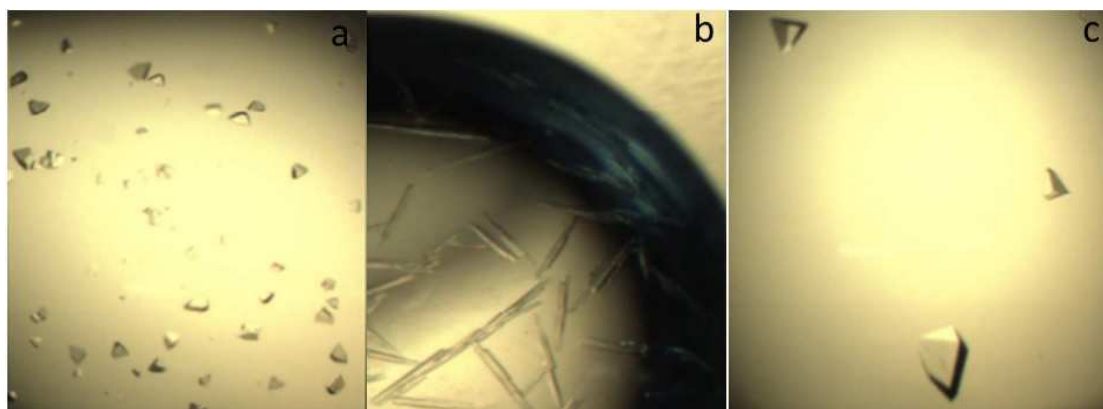
**Figura 13. Espectrometria de massa em *tandem* do peptídeo T5 da CPL.** Espectro de fragmentação do peptídeo T5. Fragmentação induzida por colisão (CID) do íon  $m/z$  541,3 duplamente carregado correspondente ao peptídeo T5 (1080,5006 Da).

O espectro gerado com íons duplamente carregado com razão  $m/z = 541,30$  do peptídeo tríptico (T5) e dos íons monocarregados com uma razão  $m/z=1651,87$  (E2) produziu um sequência comum a uma série de íons y, o qual foi interpretado como sequência primária VGLSASTGVYK e WVRVGLSASTGVYKE, respectivamente. Em outras lectinas *Diocleinae* sequenciadas por MS/MS, foi obtido um peptídeo tríptico com valina 91 na posição N-terminal e Arginina 101 na posição C-terminal. (DELATORRE *et al.*, 2011; RANGEL *et al.*, 2011; CORREIA *et al.*, 2011)

apresentando uma razão  $m/z$  de 555,45 ( $M+H^+$  1108,90), correspondendo ao peptídeo com Leu99. Lectinas de leguminosas são conhecidas por apresentar três tipos de subsítios hidrofóbicos baseados em diferentes afinidades a ligantes. O fato que um destes subsítios está adjacente ao domínio conservado de ligação a carboidrato explica o porquê dos açúcares hidrofóbicos derivados a partir da glicose/manose ligarem-se mais fortemente no sítio de reconhecimento a carboidrato do que os seus análogos não hidrofóbicos (SHARON; LIS, 1990). Kanellopoulos, *et al.*, 1996 determinaram a estrutura de duas ConA complexadas, ConA/4'-nitrofenil  $\alpha$ -D-manopiranosídeo e ConA/4'-nitrofenil  $\alpha$ -D-glicopiranosídeo, e mostrou que estes ligantes interagem com o subsítio hidrofóbico formado pelos os resíduos Tyr12, Leu99 e Tyr100. Estes resíduos de aminoácidos e suas interações são conservados em todas as lectinas da subtribo *Diocleinae* (BEZERRA *et al.*, 2007, OLIVEIRA *et al.*, 2008; BEZERRA *et al.*, 2011), porém a CPL é a primeira lectina descrita de *Diocleinae* que apresenta uma substituição conservativa no subsítio hidrofóbico com a substituição da Leu99 por Val99. Esta diferença traduz diretamente uma possível diferença no reconhecimento específico por carboidrato e conseqüentemente em atividades biológicas.

### 4.3 Estrutura completa da CPL complexada com X-Man

Cristais irregulares de CPL cresceram após uma semana na condição N° 39 do *Screen I* do kit de cristalização da Hampton Research que continha 100 mM de HEPES pH 7,5 com 2,0 M de sulfato de amônio e 2% de PEG 400 (Figura 14a), enquanto que cristais em forma de placas foram obtidos na condição N° 15 do *Screen I* que continha 0,2 M de sulfato de amônio, 100 mM de cacodilato pH 6,5 e 30% de PEG 8000 (Figura 14b), porém, estes cristais não foram suscetíveis aos experimentos de difração de raios X. Várias etapas de otimizações foram realizadas combinando as duas condições, variando pH e concentração de precipitante. Cristais de CPL cresceram em 100 mM de HEPES (pH 8,5) contendo 2,0 M de sulfato de amônio e 2% de PEG 400 com dimensões máxima de 0,1 x 0,2 x 0,15 mm (Figura 14c).



**Figura 14. Cristais de CPL.** (a) Pequenos cristais obtidos com a condição N° 39 do *Screen I*. (b) Várias placas obtidas na condição N° 15 do *Screen I*. (c) Cristais após otimização .

Cristais de CPL difrataram a uma resolução máxima de 2,15 Å usando radiação Síncroton na estação MX1 do Laboratório Nacional de Luz Síncroton (LNLS) Campinas – SP. O conjunto de 180 imagens foi indexado, integrado e escalonado a 2,15 Å. O cristal apresentou grupo espacial ortorrômbico I222 com parâmetros de cela  $a=65,9$  Å,  $b=66,7$  Å,  $c=107,7$  Å,  $\alpha=90^\circ$   $\beta=90^\circ$   $\gamma=90^\circ$ . O coeficiente de Matthews 2,37 Å<sup>3</sup>Da<sup>-1</sup>(MATTHEWS, 1968) foi calculado baseado no peso molecular da proteína  $25.298 \pm 2$  Da e implica que o cristal continha 48,18% de solvente, indicando a presença de um monômero por unidade assimétrica.

Os dados foram escalados com um total de 85234 reflexões, sendo 12072 reflexões únicas, revelando 96,4 % (96.9 %) de completude,  $R_{merge}$  a 14.7 (25.0) e  $(I/\sigma)$  em 3.4 (2.8), sendo os valores em parêntesis referentes a camada de maior resolução (2.2 – 2.1 Å). Todas as estatísticas da coleta de dados estão descritas na (Tabela 5).

**Tabela 5.** Estatística da coleta de dados de difração de raios x, refinamento e qualidade da estrutura.

Parâmetros	Valores
<b>Coleta de dados</b>	
Comprimento de onda	1,42 Å
Grupo espacial	I222
Parâmetros da cela unitária Å	
A	65,9
B	66,7
C	107,7
Número de reflexões torais	85239
Número de reflexões únicas	12072
Número de moléculas por unidade assimétrica	1
Limite de resolução	21,51 – 2,1
Rmerge (%)	14,7 (25,0) <sup>d</sup>
Completeza (%)	96,4 (96,9) <sup>d</sup>
Multiplicidade	6,7 (6,2) <sup>d</sup>
I/(σ)	3,4 (2,8) <sup>d</sup>
<b>Substituição molecular</b>	
Coeficiente de correlação	62,4
Rfactorb (%)	42,1
<b>Refinamento</b>	
Faixa de resolução (Å)	21,51 – 2,1
Rfactorb (%)	19,0
Rfreec (%)	23,6
Número de resíduos em unidade assimétrica	236
Número de moléculas de água	77
<b>R.M.S.D</b>	
Comprimento de ligação (Å)	0,019
Ângulos de ligação (graus)	1,924
<b>Média do B fator para toda a cadeia da proteína (Å)</b>	17,84
<b>Gráfico de Ramachandran</b>	
Resíduos em regiões mais favorecidas (%)	95,7
Resíduos em regiões adicionalmente permitidas (%)	4,3
Resíduos em regiões não permitidas (%)	0

<sup>a</sup>  $R_{merge} = \frac{\sum_{hkl} \sum_i |I(hkl) - \langle I(hkl)i \rangle|}{\sum_{hkl} \sum_i \langle I(hkl)i \rangle}$  where  $I(hkl)i$  is the intensity of  $i^{\text{th}}$  measurement of the reflection  $h$  and  $I(hkl)$  is the mean value of the  $I(hkl)i$  for all  $I$  measurements.

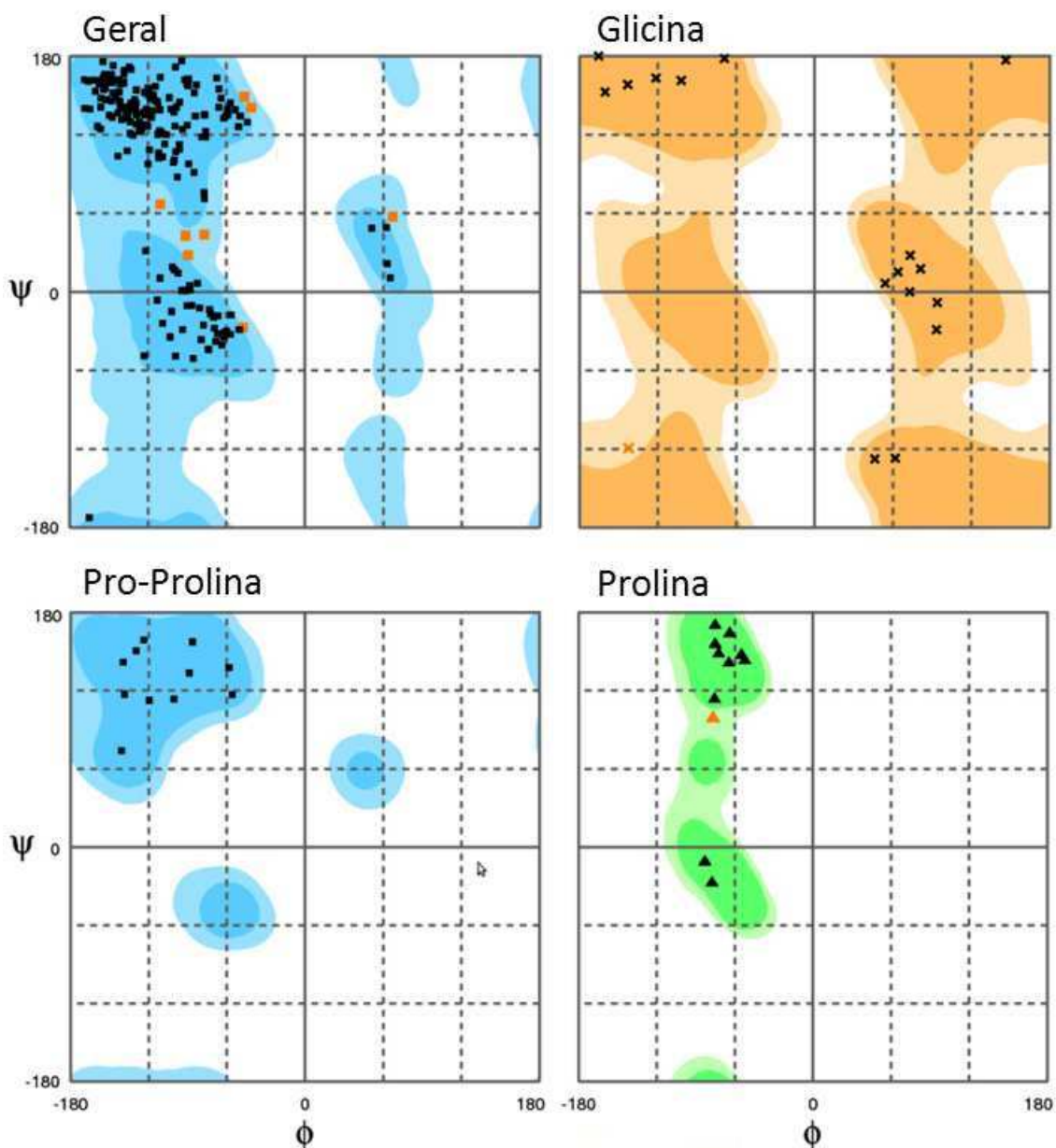
$$^b R_{factor} = \frac{|F_{obs}| - |F_{calc}|}{|F_{obs}|}$$

<sup>c</sup>Calculada com 5% dos valores omitidos no refinamento.

<sup>d</sup>Valores em parenteses representam a resolução da ultima camada.

O refinamento de corpo rígido da estrutura da CPL apresentou um *Rfactor* de 35,96 e *Rfree* de 35,95 seguido de 5 ciclos de refinamento de restrições com *Rfactor* 23,7, *Rfree* 27,4 e FOM de 79,3%. Após a adição das moléculas de água e do ligante do X-Man, a estrutura foi novamente submetida a um refinamento de restrição com um *Rfactor* e *Rfree* finais de 19,0 e 23,6 respectivamente e o FOM de 83%. O ligante X-Man apresentou um B-fator de 9,04 Å, o que está dentro dos valores esperados para vibrações térmicas de ligantes em estrutura de proteínas que vai até um máximo de 30 Å. A qualidade estereoquímica da estrutura foi avaliada no programa PROCHECK baseado nas torções psi ( $\psi$ ) e phi ( $\psi$ ) das ligações do carbono- $\alpha$ , com 95,7% dos resíduos em regiões mais favorecidas, 4,3% dos resíduos em regiões adicionalmente permitidas e nenhum resíduo em regiões não permitidas, mostrando que a estrutura final da CPL apresentou bons parâmetros estereoquímicos vistos na Figura 15 e fatores de vibração térmica (tabela 5).



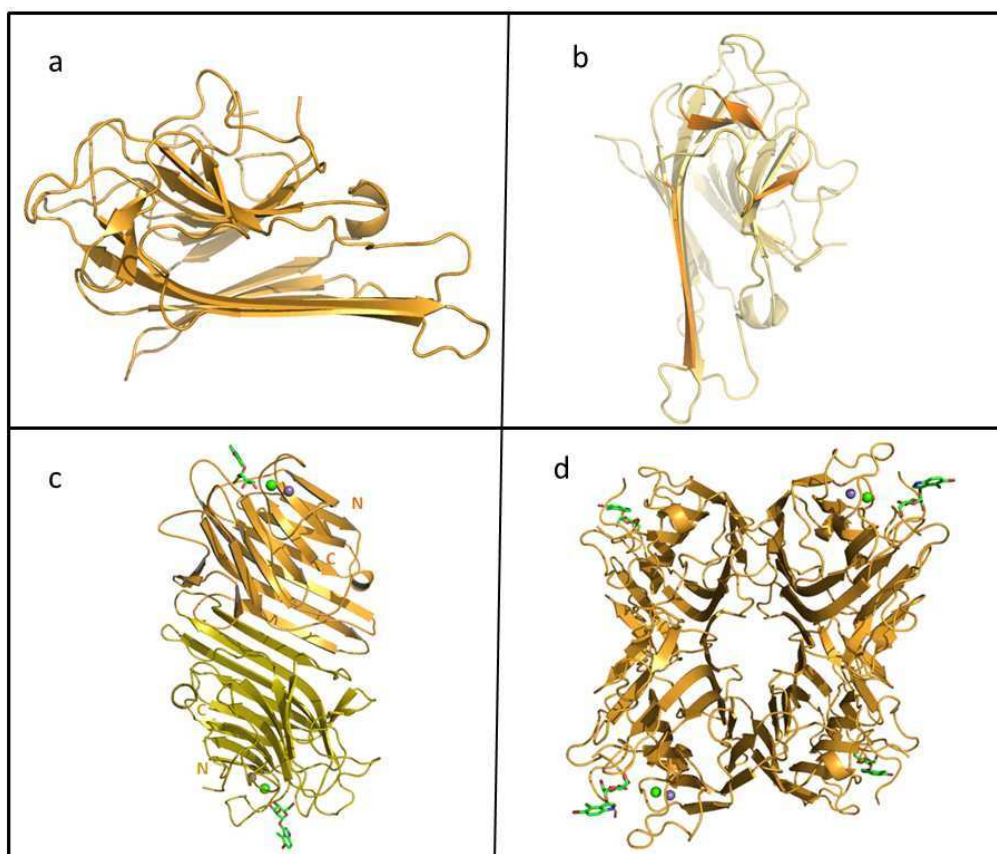


**Figura 15. Gráfico de Ramachandran das coordenadas da CPL.** Nos gráficos as regiões em azul, verde, e laranja escuro e claro, representam regiões permitidas e generosamente permitidas de ocupação por resíduos discriminados em cada gráfico, respectivamente.

Além da visão geral dos resíduos, o gráfico de Ramachandran mostra em detalhes regiões permitidas e generosamente permitidas, para os resíduos de glicina, apesar de este resíduo apresentar uma mobilidade devido a sua pequena cadeia lateral, (um átomo de hidrogênio). Como observado na figura 15, os resíduos também está dentro das posições esperadas para uma boa estrutura predominando em regiões permitidas. Outros dois parâmetros avaliados no gráfico de Ramachandran são o

posicionamento dos resíduos de prolina que geralmente forma uma curvatura na estrutura tridimensional, estando geralmente em regiões mais flexíveis como nos *loops*, bem como o posicionamento dos demais resíduos que antecedem as prolinas que também apresentam uma flexibilidade devido ao dobramento na estrutura pela presença da prolina. Ambos estão em regiões estimadas, validando assim a estrutura da CPL (Figura 15).

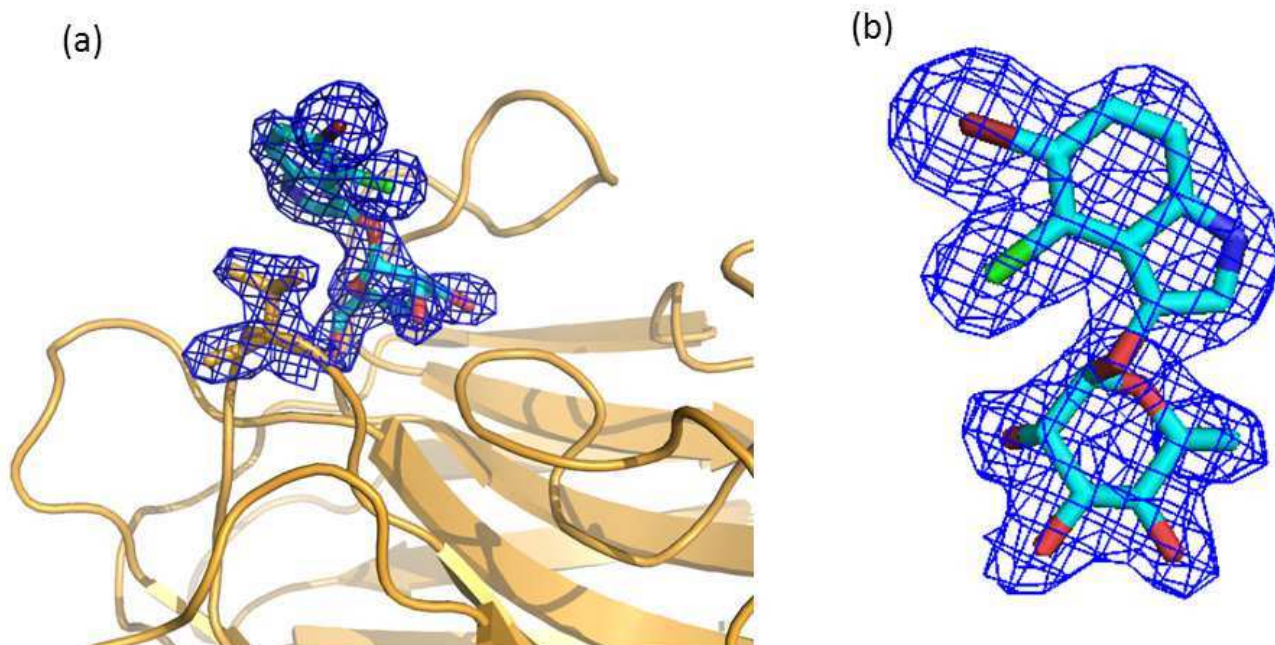
A estrutura do monômero da CPL complexado com X-Man consiste de 236 resíduos de aminoácidos apresentando os domínios  $\beta$ -sanduíche/*Jelly-roll* como observado em várias lectinas de leguminosas (Figura 16a e b). O tetramero da CPL consiste de dois dímeros canônicos ligados por interações iônicas entre as folhas beta (Figura 16c e d).



**Figura 16. Estrutura Completa da CPL complexada com X-Man.** (a) Representação do motivo  $\beta$ -sanduíche. (b) Representação do motivo *Jelly-roll*. (c) Representação do dímero canônico (d) Representação tetramérica da CPL. Esferas verde e lilás representam íons de  $\text{Ca}^{2+}$  e  $\text{Mn}^{2+}$  respectivamente, molécula de X-Man representada em verde.

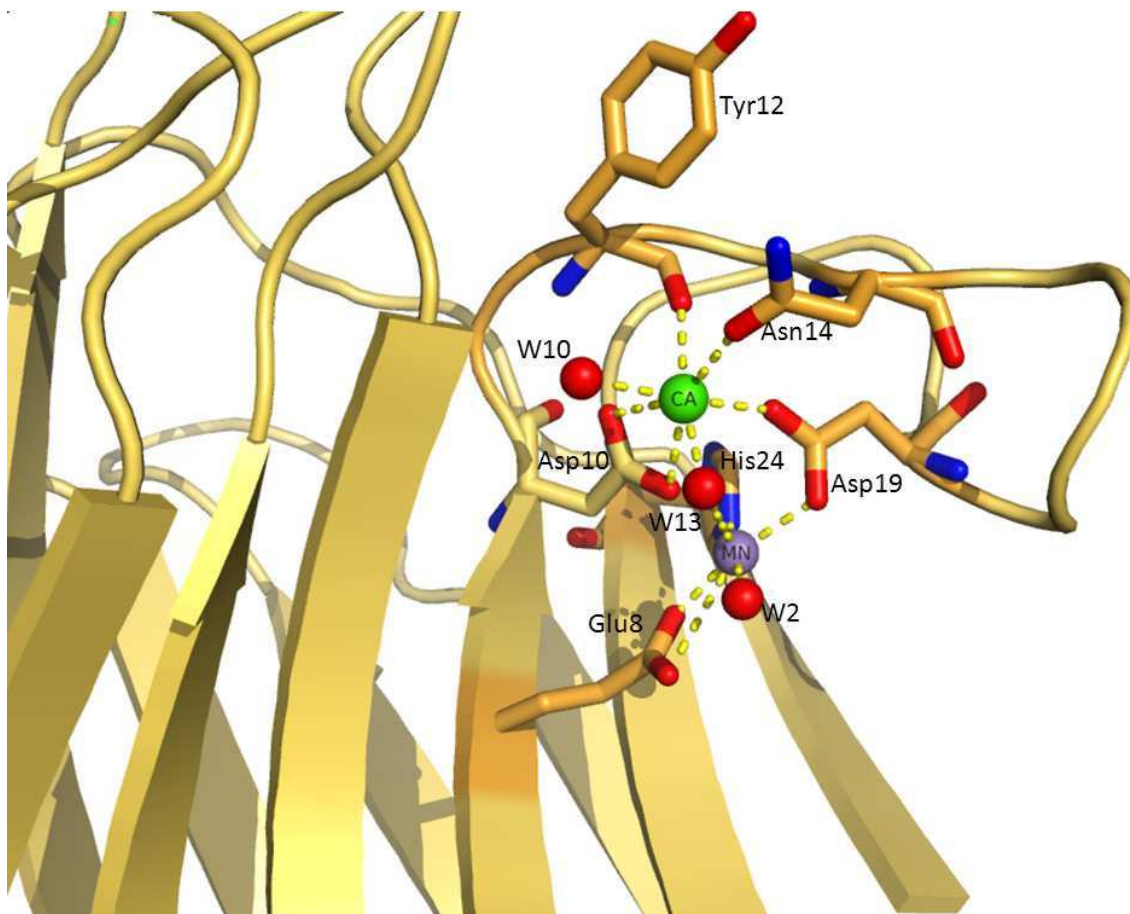


Como já descrito nos experimentos de espectrometria de massas, a estrutura da CPL apresentou um resíduo de val99, confirmada a partir do mapa de densidade (2Fo-Fc) do resíduo (Figura 17). A substituição da Leu99 por Val99 no sítio de reconhecimento a carboidrato é o primeiro descrito dentre as lectinas da tribo *Diocleinae*, até então esse sítio era tido como totalmente conservado para esse grupo de proteínas (SHARON; LIS, 1990).



**Figura 17: Mapa da densidade eletrônica para a Val99 e X-Man.** (a) Mapa de densidade eletrônica (2Fo-Fc) do resíduo de Val 99 e do X-Man no sítio de reconhecimento a carboidrato contornado a  $1\sigma$ . (b) Mapa de densidade eletrônica (2Fo-Fc) do X-Man contornado a  $1\sigma$ .

A CPL apresenta um sítio de ligação a metal conservado com os resíduos Asn14 e Tyr12 interagindo com o cálcio, Glu8 e His24 (interagindo com manganês), Asp10 e Asp19 interagindo com ambos (Figura 18). Em adição, os resíduos Ala206 e Asn207 estão isomerizados na configuração *cis* devido à presença dos íons divalentes que mudam a orientação da cadeia lateral dos resíduos Asn14 e Asp207, uma observação comum em lectinas de leguminosas. A substituição de um único resíduo neste sítio pode afetar significativamente o efeito da atividade biológica, como observado em outras lectinas com mutagênese sítio-dirigida (HIRABAYASHI; KASAI, 1991).

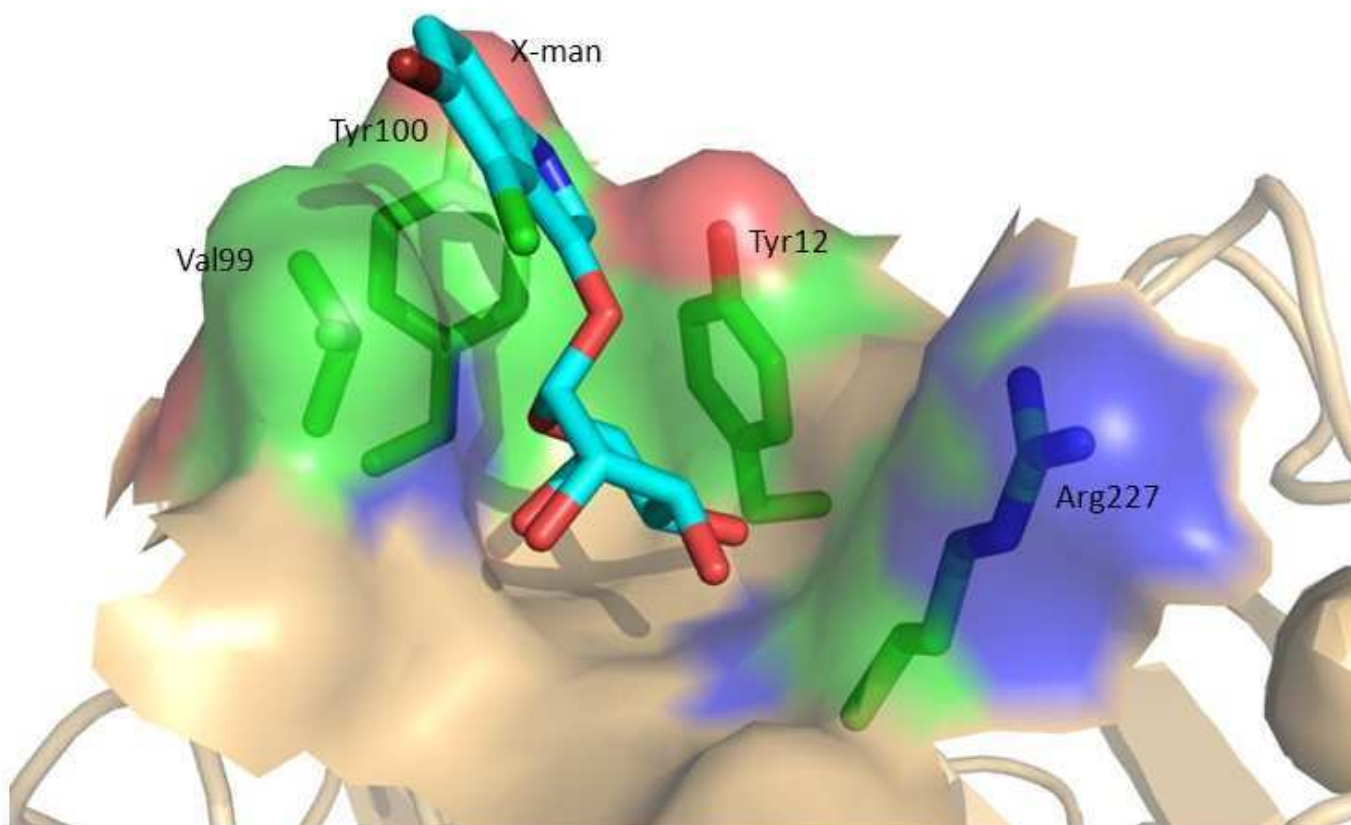


**Figura 18. Representação da interação a metais da CPL.** Interações com os metais de  $\text{Ca}^{2+}$  e  $\text{Mn}^{2+}$  (esferas verde e lilás respectivamente) estabilizando o sítio de ligação ao carboidrato na proteína. Linhas pontilhadas representam coordenações realizadas entre os metais, moléculas de água e resíduos de aminoácidos.

O domínio de reconhecimento a carboidrato está ocupado por X-Man (Figura 19). O complexo envolve 10 interações intermoleculares por pontes de hidrogênio entre O3, O4, O5 e O6 do X-Man e as cadeias laterais dos resíduos de aminoácidos e 8 interações de Van der Waals com resíduos de aminoácidos do subsítio hidrofóbico (Tabela 6).

**Tabela 6.** Interações de Van der Waals e contatos polares entre a lectina de sementes de *C. pedicellatume* X-Man

Aminoácidos	X-Man	Distâncias (Å)
<i>Interações de Van der Waals</i>		
Tyr12 OH	C11	3,38
Asp207 OD1	C4	3,39
Asp207 CG	O4	3,48
Arg227 CB	O4	3,40
Val99 CB	O5	3,44
Gly98 CA	O6	3,21
Gly98 C	O6	3,43
Ala206 CB	O6	3,27
<i>Contatos Polares</i>		
Arg227 N	O3	2,93
Asn14 ND2	O4	3,01
Asp207 OD1	O4	2,60
Arg227 N	O4	3,19
Val99 N	O5	3,06
Gly98 N	O6	3,33
Val99 N	O6	3,01
Tyr100 O	O6	3,36
Tyr100 N	O6	3,01
Asp207 OD2	O6	2,97

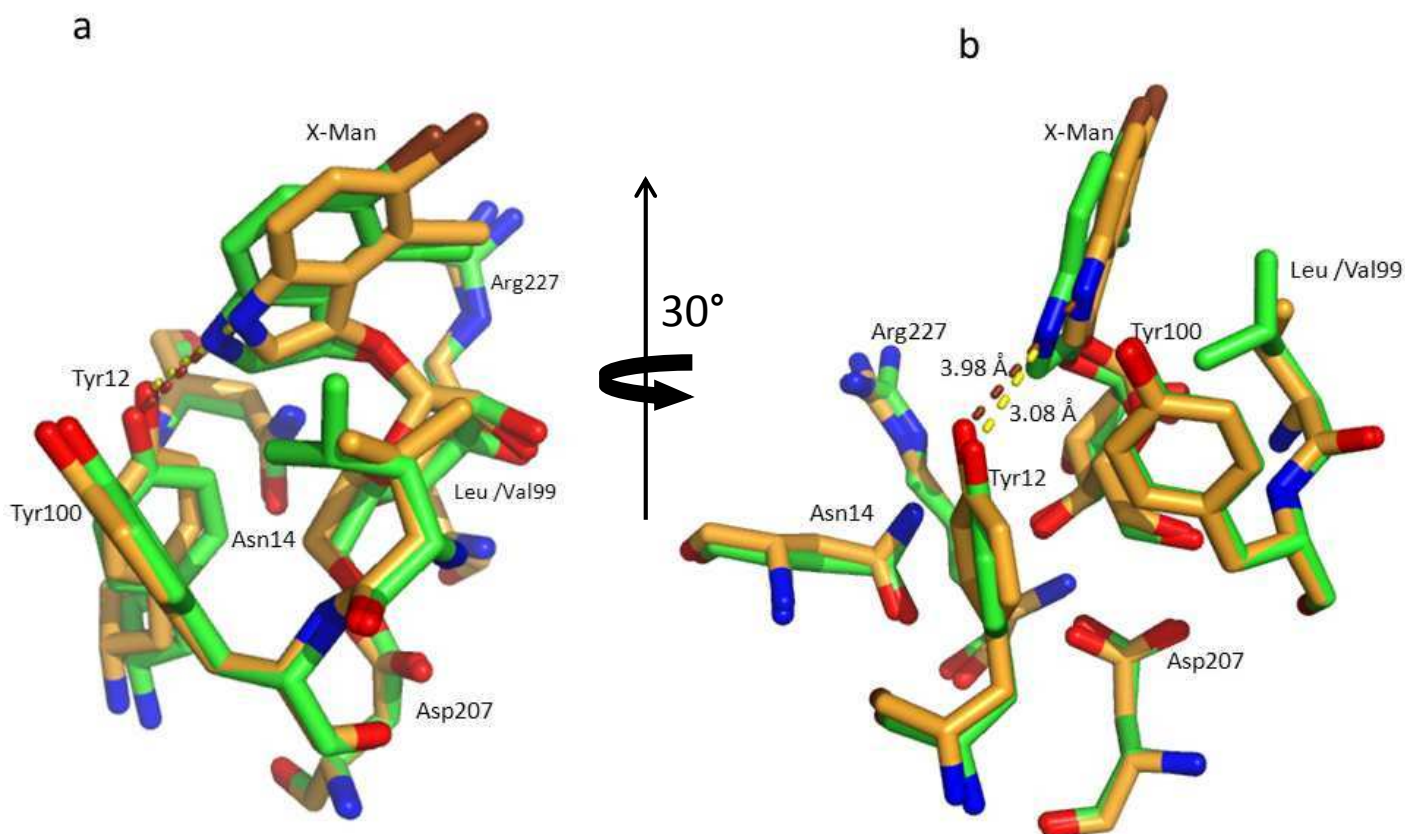


**Figura 19. Representação da área de superfície do sítio de reconhecimento a carboidrato da CPL.**

As interações de Van der Waals entre os resíduos de aminoácidos e o ligante, neste caso o X-Man, são basicamente estabelecidas por resíduos que compõem uma porção do sítio, chamado de subsítio hidrofóbico composto por Tyr12, Leu99 e Tyr100. (BEZERRA *et al.*, 2007). A estrutura primária da CPL apresenta uma substituição conservativa neste sítio. Esta mudança no subsítio hidrofóbico teoricamente reduziria as interações hidrofóbicas entre o X-Man e a lectina; porém, devido à ausência de um grupamento metil na Val99 em relação a Leu99, o grupo indol pode adentrar

mais no sítio, deixando a Val99 1,0 Å mais próximo ao grupo indol do X-Man em relação a Leu99, garantindo assim, uma interação mais efetiva deste resíduo com o ligante através do subsítio hidrofóbico (SHARON; LIS 1990).

A presença do aminoácido hidrofóbico na posição 99 não apenas aumenta a interação lectina e carboidrato, ele também afeta a orientação do X-Man no domínio de reconhecimento a carboidrato. A presença da Val99 causa uma mudança no posicionamento do ligante no sítio, deixando o grupo indol do X-Man a uma distância de 3,98 Å da OH da Tyr12 (Figura 20a) não permitindo assim o estabelecimento da ponte de hidrogênio entre o N1 do X-Man e OH da Tyr12 (Figura. 20b). Em outras lectinas *Diocleinae* complexadas com X-Man, a distância do ligante para a OH da Tyr12 é reduzido e a interação é facilitada devido à presença de Leu99 no subsítio hidrofóbico.



**Figura 20. Superposição do domínio de reconhecimento a carboidrato da CPLE DVL. (a) Diferença posicional do X-Man na CPL (laranja) e DVL (verde) devido a substituição da Leu por Val na posição 99. (b) Formação de uma ponte de hidrogênio entre N1 do X-Man e OH da Tyr12 orientado por Leu99 em DVL. A presença de Val99 em CPL não permite a formação dessa ponte de hidrogênio devido a uma reorientação**



no DRC. O ligante está posicionado a uma distância 3,98Å do OH da Tyr12 (linhas marron).

**Tabela 7.** Distâncias entre OH Tyr12 e N1 do X-Man no domínio de reconhecimento a carboidrato de lectinas de Diocleinae.

<b>Lectinas Diocleinae</b>	<b>Resíduo 99</b>	<b>Distância (Å)</b>	<b>PDB Código</b>
<i>Camptosema pedicellatum</i>	Valina	3,98 Å	3U4X
<i>Dioclea violacea</i>	Leucina	3,08 Å	3AX4
<i>Dioclea wilsonii</i>	Leucina	3,20 Å	3SH3
<i>Dioclea guianensis</i>	Leucina	3,19 Å	2JDZ

A análise da sobreposição do domínio de reconhecimento da DVL (PDB código: 3AX4) e CPL (PDB código 3U4X) (Figura 18 b) revelou diferenças na posição do grupo indol e OH da Tyr12. Gadelha e colaboradores demonstraram a importância da ponte de hidrogênio estabelecida entre a hidroxila da Tyr12 e o dissacarídeo na estrutura da lectina de *Canavalia maritima* (ConM PDB: ). A substituição do resíduo de prolina por serina na posição 202 aproxima a Tyr12 dos carboidratos no DRC, favorecendo assim a formação da ponte de hidrogênio entre a OH da Tyr12 e glicanos no sítio. Esta observação na estrutura da ConM foi suficiente para mais tarde explicar o porquê da ConM interagir com glicanos das células endoteliais e estimular a produção de NO pela ativação da enzima óxido nítrico sintase endotelial (GADELHA *et al.*, 2005). Logo, a diferença pontual na Tyr12 da ConM em relação a outras lectinas tipo ConA foi suficiente para aumentar a produção de NO em relação a outras lectinas *Diocleinae*, mesmo essas compartilhando cerca de 96% de identidade nas suas sequências primárias (DELATORRE *et al.*, 2006).

A posição do grupo indol na CPL em comparação a outras estruturas de lectinas semelhantes a ConA complexadas com este ligante difere 0,9 Å, o bastante para não permitir a formação da ponte de hidrogênio entre a OH da Tyr12 e N1 do X-Man. De fato, a interação hidrofóbica da Val/Leu com o ligante aumenta significativamente a constante de associação do ligante no domínio de reconhecimento a carboidrato. Foi observado, portanto, a importância do resíduo 99 no posicionamento do ligante no domínio de reconhecimento a carboidrato por interação hidrofóbica, possibilitando a modificação da afinidade das lectinas com Val99 e Leu99 por carboidratos complexos.

Esta variação é pequena, porém importante dentro dessas moléculas que apresentam uma alta similaridade.

Vários fatores tem sido relacionados à diversidade das atividades biológicas induzidas pelas lectinas *Diocleinae* por, isolados ou em combinação, incluindo arranjo biológico e equilíbrio dímero-tetrâmero pH dependente (GALLEGO DEL SOL *et al.*, 2007), mudanças na configuração do domínio de reconhecimento a carboidrato induzido por modificações pontuais (DELATORRE *et al.*, 2006), volume do domínio de reconhecimento a carboidrato (BEZERRA *et al.*, 2011) e como mostrado no presente estudo, substituição de resíduo no subsítio hidrofóbico.

## 5. CONCLUSÃO

A estrutura tridimensional da CPL demonstrou uma mutação conservativa da Leu99 por Val99 no subsítio hidrofóbico, um constituinte do domínio de reconhecimento a carboidrato. Esta modificação mostrou a relevância das interações hidrofóbicas no estabelecimento das interações entre o carboidrato e o substituinte e como esta interação determina a orientação espacial do ligante no domínio de reconhecimento a carboidrato, modificando a especificidade fina da interação. A substituição e a análise das interações da CPL com o ligante X-Man revelou o efeito causado pela presença da Val99 no sítio impedindo a formação de ponte de hidrogênio entre o ligante e o N1 da Tyr12 devido a orientação do grupo indol no sítio de reconhecimento a carboidrato. Estes aspectos estruturais são responsáveis por parte da variabilidade na atividade biológica observado em outras lectinas *Diocleinae* que apresentam estrutura primária e quaternária similar a CPL.

Como perspectivas para trabalhos futuro, deverá ser realizado análise da interação da CPL com açúcares complexos por microclorimetria, bem como a determinação de atividades biológicas já descritas para outras lectinas da mesma tribo correlacionando assim, a modificação pontual do resíduo 99 com os resultados dos ensaios, a fim de esclarecer a importância dessa modificação no mecanismo de reconhecimento a carboidratos nas lectinas *Diocleinae*.



## 6. REFERÊNCIAS:

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

ANDRADE, J. L., ARRUDA, S., BARBOSA, T., PAIM, L., RAMOS, M. V., CAVADA, B.S., BARRAL-NETTO, M. Lectin-Induced No Production. **Cellular Immunology**, v. 194, n. 1, 98-102, 1999.

ASSREUY, A. M. S., SHIBUYA, M. D. MARTINS, G. J., de SOUZA, M. L. P., CAVADA, B. S., MOREIRA, R. A., OLIVEIRA, J. T. A., RIBEIRO, R. A., FLORES, C. A. Antiinflammatory effect of glucose-mannose binding lectins isolated from brazilian beans. **Mediators of inflammation**.6, 1-10, 1997.

BALZARINI, J. Targeting the glycans of glycoproteins: a novel paradigm for antiviral therapy. **Nature Reviews**, v. 5, p. 583-597, 2007.

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.

BARBOSA, T., ARRUDA, S., CAVADA, B.S., 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.

BARRAL-NETTO, M., SANTOS, S.B., BARRAL, A., MOREIRA, L. I. M., SANTOS, C. F., MOREIRA, R. A., OLIVEIRA, J. T. A., CAVADA, B.S. Human lymphocyte stimulation by legume lectins from the Diocleae tribe. **Immunol. Invest.** 21, 297-303, 1992.

BENEVIDES, R.G. Caracterização bioquímica e estrutural de uma lectina recombinante de sementes de *Platypodium elegans* Vogel. **Tese de Doutorado**, Universidade Federal do Ceará, Fortaleza. 2011

BENTO, C.A.M., CAVADA, B.S., OLIVEIRA, J.T.A., FIDALGO, C. B Rat paw edema and leukocyte immigration induced by plant-lectins. **Agent Actions**.38, 48-54, 1992.

BEZERRA, E.H.S., ROCHA, B.A.M., NAGANO, C.S., BEZERRA, G.A.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 endothelial no synthase activation. **Biochem.Bioph.Res.** 408, 566 - 570, 2011.

BEZERRA, G.A., OLIVEIRA, T.M., MORENO, F.B.M.B., SOUZA, E.P., ROCHA, B.A.M., BENEVIDES, R.G., DELATORRE, P., AZEVEDO JR., 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. **J. Struct. Biol.** 160, 168-176, 2007

BRINDA, K. V., MITRA, N., SUROLIA, A., VISHVESHWARA, S.. Determinants of quaternary association in legume lectins. **Protein Science**.13, 1735-1749, 2004.

CALVETE, J.J., THOLE, H.H., RAIDA, M.,URBANKE, C. ROMERO, A. GRANGEIRO, T. B. RAMOS, M. V. ALMEIDA, D. A. ROCHA, I. M. GUIMARÃES, F. N. CAVADA, B. S. Molecular characterization and crystallization of *Diocleinae*lectins. **Biochim. Biophys.Acta**.1430, 367–375, 1999.

CARRINGTON, D.M., AUFFRET, A., HANKE, D. E. Polypeptide ligation occurs during post-translational modification of Concanavalin A. **Nature**. 313, 3, 64-67, 1985.

CAVADA, B. S., BARBOSA, T., ARRUDA, S., GRANGEIRO, T. B., BARRAL NETTO, M. Revisiting proteus: Do minor changes in lectin structure matter in biological activity? Lessons from and potential biotechnological uses of the *Diocleinae* subtribe lectins. **Curr. Protein Pept.** 2, 123–135, 2001.

CAVADA, B. S., GRANGEIRO, T. B., RAMOS, M. V., CORDEIRO, E. F., OLIVEIRA, J. T. A., MOREIRA, R. A. Isolation and partial characterization of a lectin from *Diocleastrostrata* Benth seeds. **Revista Brasileira de Fisiologia Veget al.** 8, 1, 31-36, 1996a.

CAVADA, B. S., RAMOS, M. V., CORDEIRO, E. F., GRANGEIRO, T. B., OLIVEIRA, J. T. A., CARVALHO, A. F. F. U., MOREIRA, R. A. Purification and partial characterization of a lectin from *Dioclea virgata* Benth seeds. **Revista Brasileira de Fisiologia Vegetal** 8,1, 37-42, 1996b.

CECCATTO, V. M., CAVADA, B. S., NUNES, E. P., NOGUEIRA, N. A. P., GRANGEIRO, M. B., MORENO, F. B. M. B., TEIXEIRA, E. H., SAMPAIO, A. H., ALVES, M. A. O., RAMOS, M. V., CALVETE, J. J., GRANGEIRO, T. B. Purification and partial characterization of a lectin from *Canavalia grandiflora* Benth. seed. **Protein and Peptide Letters**, Holanda. 9, 1, 67-73, 2002.

COOLINGE, D. B., KRAGH, K. M., MIKKELSEN, J.D., NIELSEN, K.K., RASMUSSEN, U., VAD, K. Plant chitinases. **Plant Journal**.3, 31-40, 1993.

COLLABORATIVE COMPUTATIONAL PROJECT, NUMBER 4. The CCP4 suite: programs for crystallography. **Acta Crystallogr. D. Biol. Crystallogr.**50, 760–763 1994.

CORREIA, J.L.A., NASCIMENTO, A.S.F., CAJAZEIRAS, J.B., GONDIM, A.C. S., PEREIRA, R.I., SOUSA, B.L., SILVA, A.L.C., GARCIA, W., TEIXEIRA, E.H., NASCIMENTO, K.S., ROCHA, B.A.M., NAGANO, C.S., SAMPAIO, A.H., CAVADA, B.S. Molecular Characterization and Tandem Mass Spectrometry of the Lectin Extracted from the Seeds of *Dioclea sclerocarpa* Ducke. **Molecules**.16, 9077-9089, 2011.

DAVIS, I.W.; LEAVER-FAY A.; CHEN, V. B.; BLOCK, J. N., KAPRAL, G. J. WANG, X.; MURRAY, L.W.; ARENDALL, W.B., R. D. SNOEYINK, J. RICHARDSON, J. S. RICHARDSON, D. C. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. **Nucleic Acids Res.** 35 375–383 2007.

DAYHOFF, J.E., SHOEMAKER, B.A., BRYANT, S.H., AND PANCHENKO, A.R. Evolution of Protein Binding Modes in Homooligomers. **J. Mol. Biol.** 395, 860–870, 2009.

DELANO, W.L. The Pymol Molecular Graphics System. DeLano Scientific, San Carlos, CA. 2002.

DELATORRE, P., ROCHA, B.A.M., GADELHA, C.A.A., SANTI-GADELHA, T., CAJAZEIRAS, J.B., SOUZA, E.P., NASCIMENTO, K.S., FREIRE, V.N., SAMPAIO, A.H., AZEVEDO JR., W.F. CAVADA, B. S. Crystal structure of a lectin from *Canavalia maritima* (ConM) in complex with threose and maltose reveals relevant mutation in ConA-like lectins. **J. Struct. Biol.** 154, 280-286, 2006.

DELATORRE, P., ROCHA, B.A.M., SIMÕES, R.C., PEREIRA-JÚNIOR, F.N., SILVA, H.C., BEZERRA, E.H.S., BEZERRA, M.J.B., MARINHO, E.S., GADELHA, C.A.A., SANTI-GADELHA, T., FARIAS, D.L., ASSREUY, A.M.S., MARQUES-DOMINGOS, G.F.O., NAGANO, C.S., CAVADA, B.S. Mass Spectrometry and X-ray Diffraction Analysis of Two Crystal Types of *Dioclea virgata* Lectin: An Antinociceptive Protein Candidate to Structure/Function Analysis. **Appl. Biochem. Biotechnol.** 164, 741–754, 2011.

DELATORRE, P., ROCHA, B. A., SOUZA, E. P., OLIVEIRA, T. M., BEZERRA, G. A., MORENO, F. B. M. 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.

DENNIS, J. W., LAFERTE, S., WAGHORNE, C., BREITMAN, M. L., KERBEL, R. S. J. Roth, **Chem. Rev.** 102, 285, 2002

EDELMAN, G. M., WANG, J. L. Binding and functional properties of Concanavalin A and its derivatives. III interactions with indoleacetic acid and other hydrophobic ligands. **J. Biol. Chem.** 253, 3016–3022, 1978.

EVANS, P. R. Scala. Joint CCP4 and ESF-EACBM. **Newsletter.** 33, 22-24, 1997.

FERRIGE, A.G., SEDDON, M.J., GREEN, B.N., JARVIS, S.A., SKILLING, J. (Disentangling electrospray spectra with maximum entropy. **Rapid Commun. Mass Spectrom.** 6, 707–711, 1992.

GABIUS, H. J. Biological information transfer beyond the genetic code: the sugar code. **Naturwissenschaften**, 87, 108–121, 2000.

GADELHA, C.A.A., MORENO, F.B.M.B., SANTI-GADELHA, T., CAJAZEIRAS, J. B., RODRIGUES, J. K., TUPINAMBA, B. F., CANDURI, F., DELATORRE, P., FILGUEIRA, A. W., CAVADA, B.S. Crystallization and preliminary x-ray diffraction

analysis of a lectin from *Canavalia maritima* seeds. **Acta Crystallogr.** 61, .87–89, 2005.

GADELHA, C.A.A., MORENO, F.B.M.B., SANTI-GADELHA, T., CAJAZEIRAS, J.B., ROCHA, B.A.M., ASSREUY, A.M.S., MOTA, M.R.L., PINTO, N.V., MEIRELES, A.V.P., BORGES, J.C., FREITAS, B.T., CANDURI, F., SOUZA, E. P., DELATORRE, P., CRIDLLE, D. N. AZEVEDO JR., W. F., CAVADA, B. S. Native crystal structure of a nitric oxide releasing lectin from the seeds of *Canavalia maritima*. **J. Struct. Biol.** 152, 185-195 2005.

GALLEGO DEL SOL, F., NAGANO, C.S., MORENO, F.B.M.B. Diocleinae lectins: clues to delineate structure/function correlations. **Methods In Proteome And Protein Analysis**, p.81–91, 2004.

GALLEGO-DEL-SOL, F., NAGANO, C. S., CAVADA, B. S. Lectinas: proteínas traductoras de glicocódigos y herramientas bioquímicas. **Investigación Y Ciencia**, 2006.

GALLEGO DEL SOL, F., CAVADA, B.S., CALVETE, J.J. Crystal structures of *Cratylia floribunda* seed lectin at acidic and basic pHs. Insights into the structural basis of the pH-dependent dimer–tetramer transition. **J Struct.Biol.** 158, 1-9,2007.

GOLDSTEIN, I. J. My Favorite Enzyme: For Love of Lectins. **IUBMB Life**, 623, 247–250, 2010.

GOMES, C., ROSSI, R. R., CAVADA, B. S. MOREIRA, R. A.; OLIVEIRA J. T.. Histamine release induced by glucose (mannose)-specific lectins isolated from Brazilian beans. Comparison with concanavalin A. **Agents Actions.** 41, 132-135, 1994.

HAMELRYCK, T. W. POORTMANSI, F. GOOSSENS, A. ANGENON, G. MONTAGU, M. V. WYNS, L. LORIS, R. Crystal Structure of Arcelin-5, a Lectin-like Defense Protein from *Phaseolus vulgaris*. **The journal of biological chemistry.** 20. 32796 –32802. 1996

HAMELRYCK, T. W., LORIS, R., BOUCKAERT, J., DAO-THI, M., STRECKER, G., IMBERTY, A., FERNANDEZ, E., WYNS, L., ETZLER, M. E. Carbohydrate Binding, Quaternary Structure and a Novel Hydrophobic Binding Site in Two Legume Lectin Oligomers from *Dolichos biflorus*. **J. Mol. Biol.** 286, 1161-1177, 1999.

HARDMAN, K. D. A. I. N., SWORTH, C. F. Structure of Concanavalin A at 2.4 Å resolution. **Biochemistry**.11, 4910–4919, 1972.

HIRABAYASHI, J., KASAI, K. Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin. **J. Biol.Chem.** 266, 23648-23653, 1991.

KANELLOPOULOS, P.N., PAVLOU, K., PERRAKIS, A., AGIANIAN, B., VORGIAS, C.E., MAVROMMATIS, C., SOUFI, M., TUCKER, P.A., HAMODRAKAS, S.J. The crystal structure of the complexes of concanavalin A with 40-nitrophenyl-alpha-D-mannopyranoside and 40-nitrophenylalpha-D-glucopyranoside. **J. Struct. Biol.** 116, 345–355, 1996

Komath, S. S., Kavithab, M., Swamy, M. J. Beyond carbohydrate binding: new directions in plant lectin research. **Organic & Biomolecular Chemistry**. 4, 973–988. 2006

LAEMLI, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. **Nature**, 227, 680-685, 1970.

LESLIE, A.G.W. MOSFLM User's Guide: MOSFLM Version 550. MRC Laboratory of Molecular Biology, Cambridge, England. 1997

LORIS R., HAMELRYCK T., BOUCKAERT J. Legume lectin structure, **Biochim Biophys Acta**.1383, 9–36, 1998.

LORIS, R. Principles of structures of animal and plant lectins. **Biochim Biophys Acta**. 1572: 198–208, 2002

MATTHEWS, B.W. Solvent content of protein crystals. **J. Mol. Biol.** 33, 491–497 1968.

MITCHELL, S. W., REICHERT, E. T. Researches upon the venoms of poisonous serpents, Smithsonian. **Contrib. Knowl.** XXVI, 155p, 1886.

MITCHELL, S.W. Researches upon the venom of the rattlesnake, Smithsonian. **Contrib. Knowl.** XII. 89–90, 1860.

- MOREIRA, R.A., CAVADA, B.S. Lectin from *Canavalia brasiliensis* (mart.). isolation, characterization and behavior during germination, **Biol. Plant.**, v.28, p.113-120,1984.
- MOREIRA, R. A. CORDEIRO, E. F., GRANGEIRO, T. B., RAMOS, M. V., MARTINS, J. L., CAVADA, B. S., OLIVEIRA, J. T. A. Isolation and partial characterization of a lectin from *Dioclea violacea* Benth seeds. **Revista Brasileira de Fisiologia Vegetal** 8, 1, 23-29, 1996.
- MOREIRA, R. A., MONTEIRO, A. C. O., HORTA, A. C. G., OLIVEIRA, J. T. A., CAVADA, B. S. Isolation and characterization of alectin from *Dioclea altissima* var Megacarpa seeds lectin. **Phytochemistry** 46, n. 1, 139-144, 1997.
- MOREIRA, R. A. & PERRONE, J. C. Purification and partial characterization of a lectin from *Phaseolus vulgaris*. **Plant Physiology**. 59, 783-787. 1977.
- MORENO, F., DE OLIVEIRA, T., MARTIL, D., VICOTI, M., BEZERRA, G., ABREGO, J., CAVADA, B., AZEVEDOJR, W. F. Identification of a new quaternary association for legume lectins. **Journal of Structural Biology**. 161, 133-143, 2008.
- MURDOCK, L.L., SHADE, R.E. Lectins and protease inhibitors as plant defenses against insects. **J Agric Food Chem.**, 50: 6605-11. 2002.
- NASI, A. PICARIELLO, G. FERRANTI, P. Proteomic approaches to study structure, functions and toxicity of legume seeds lectins. Perspectives for the assessment of food quality and safety. **Journal of proteomics** 72 527 – 538. 2009.
- 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.
- 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.

OKUYAMA, S., NAKAMURA-TSURUTA, S., TATENO, H., HIRABAYASHI, J., MATSUBARA, K., HORI, K. Strict Binding Specificity of Small-Sized Lectins from the Red Alga *Hypnea japonica* for Core ( $\alpha$ 1-6) Fucosylated N-Glycans. **Biosci. Biotechnol. Biochem.**, 73 4, 912–920, 2009.

OLIVEIRA, J. T. A., CAVADA, B. S., MOREIRA, R. A. Isolation and partial characterization of a lectin from *Cratylia floribunda* Mart. seeds. **Revista Brasileira de Botânica**. 14, 61-66, 1991.

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.M.B., FREIRE, V.N., AZEVEDO JR., W.F., 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. **J. Struct. Biol.** 164, 177 - 182, 2008.

PEUMANS, W. J., BARRE, A., HAO, Q., ROUGÉ, P., VAN DAMME, E. J. M. Higher plants developed structurally different motifs to recognize foreign glycans. **Trends in Glycoscience and Glycotechnology**, 12 64 83-101, 2000.

PEUMANS, W.J., STINISSEN, H.M. Gramineae lectins: occurrence, molecular biology, and physiological function. In: **Chemical Taxonomy, Molecular Biology and Function of Plant Lectins**. 99-138. 1983

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

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.

PEUMANS, W.J., VAN DAMME, E.J. Plant lectins: specific tools for the identification, isolation, and characterization of O-linked glycans. **Crit.Rev. Biochem. Mol. Biol.** 33, 209-58, 1998.

PEUMANS, W.J., HAUSE, B. VAN DAMME, E.J. The galactose-binding and mannose-binding jacalin-related lectins are located in different sub-cellular compartments. **FEBS Letters** 477 186-192. 2000



PUSZTAI, A. Plant lectins. **Cambridge University Press**, Cambridge, 1991.

RAMOS, M.V., MOREIRA, R.A., CAVADA, B.S., OLIVEIRA, J.T.A., ROUGE, P. Interaction of lectins from the sub tribe Diocleinae with specific ligands. **Rev.Bras.Fisiol.Veg.** 8, 193-199, 1996.

RANGEL, T.B.A., ASSREUY, A.M.S., PIRES, A.F., CARVALHO, A.U., BENEVIDES, R.G., SIMÕES, R.C., SILVA, H.C., BEZERRA, M.J.B., NASCIMENTO, A.S.F., NASCIMENTO, K.S., NAGANO, C.S., SAMPAIO, A.H., DELATORRE, P., ROCHA, B.A.M., FERNANDES, P.M.B., CAVADA, B.S. Crystallization and Characterization of an Inflammatory Lectin Purified from the Seeds of *Dioclea wilsonii*. **Molecules.** 16, 5087-5103 2011.

RODRIGUEZ, D., CAVADA, B.S., OLIVEIRA, J.T.A., *et al.* Differences In Macrophage Stimulation And Leukocyte Accumulation In Response To Intraperitoneal Administration Of Glucose Mannose-Binding Plant-Lectins. **Braz. J. Med. Biol. Res.** 25, 823-826, 1992.

ROUGÉ, P., RICHARDSON, M., RANFAING, P., YARWOOD, A., CAVADA, B. S. Single- and two-chain legume lectins as phylogenetic markers of speciation. **Biochemical Systematics and Ecology.** 15, 341- 348, 1987.

SHARON, N. and LIS, H. Lectins as cell recognition molecules. **Science**, 46, 227- 234, 1989.

SHARON, N., LIS, H. History of lectins: from hemagglutinins to biological recognition molecules; **Glycobiology.** 14 11: 53-62 2004.

SHARON, N., LIS, H., Legume lectins- A large family of homologous proteins. **FASEB J.** 4, 3198–3208, 1990.

SHARON, N. Lectins: Carbohydrate-specific Reagents and Biological Recognition Molecules. **The Journal of Biological Chemistry**, 282, 5, 2753-2764, 2007.

SINHA, S., GUPTA, G., VIJAYAN, M., SUROLIA, A. Subunit assembly of plant lectins. **Current Opinion in Structural Biology.** 17, 5, 498-505, 2007

SRINIVAS, V. R., ACHARYA, S., RAWAT, S., SHARMA, V., SUROLIA, A.. The primary structure of the acidic lectin from winged bean *Psophocarpus tetragonolobus*:

insights in carbohydrate recognition, adenine binding and quaternary association. **FEBS Lett.** 474, 76, 2000.

SRINIVAS, V. R., REDDY, G. B., AHMAD N. Legume lectin family, the 'natural mutants of the quaternary state', provide insights into the relationship between protein stability and oligomerization. **Biochimica Et Biophysica Acta - General Subjects.** 1527, 3, 102-111, 2001.

STILLMARK, H. Uber Ricin, ein giftiges Ferment aus den Samen von *Ricinus communis* L.und einigen anderen Euphorbiaceen, **Tese de Doutorado**, Universidade de Dorpat, Estônia, 1888.

STIRPE, F., BATTELLI, M.G. . Ribosome-inactivating proteins: progress and problems. **Cell Mol. Life Sci.** 63: 1850-66, 2006.

SUMNER, J.B. The globulins of the jack bean, *Canavalia ensiformis*, **J. Biol. Chem.** 37, 137-42, 1919.

TRANSUE, T.R., SMITH, A.K., MO, H., GOLDSTEIN, I.J., SAPER, M. A. . Structure of benzyl T-antigen disaccharide bound to *Amaranthus caudatus agglutinin*. **Nat. Struct. Biol.** 4:779-83, 1997

TULLIS, R. H., DUFFIN, R. P.,HANDLEY, H. H., SODHI, P., MENON, J., JOYCE, J. A., KHER, V. Reduction of Hepatitis C Virus Using Lectin Affinity Plasmapheresis in Dialysis Patients. **Blood Purif.** 27 64-69, 2009.

VAN DAMME, E.J.M., PEUMANS, W.J., BARRE, A. & ROUGÉ, P. Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. **Crit. Rev. Plant Sci.** 17 575-692, 1998.

VARGIN, A., TEPLYAKOV, A. MOLREP: an automated program for molecular replacement. **J. Appl. Crystallogr.** 30, 1022-1025, 1997.

WAH, D. A., ROMERO, A., GALLEGO DEL SOL, F., CAVADA, B. S., RAMOS, M. V., GRANGEIRO, T. B.,SAMPAIO, A. H. AND CALVETE, J. J. 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.

WERZ, D., SEEBERGER, P. Total synthesis of antigen Bacillus anthracis tetrasaccharide Creation of an anthrax vaccine candidate. **Angew. Chem. Int. Ed.** 44, 6315-6318. 2005

WILLIAMS, A., WESTHEAD D.R. Sequence relationships in the legume lectin fold and other jelly rolls. *Protein Engineering* vol.15 no.10 pp.771–774, 2002.

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**, v. 12, p. 135-141. 1995.

# **ANEXO**

**Crystal structure of the lectin of *Camptosema  
pedicellatum*: implications of a conservative  
substitution at the hydrophobic subsite  
(paper submetido ao *The Journal of Biochemistry*)**

structure (6, 7). Nevertheless, Diocleinae lectins may differ substantially in biological activities, as shown in studies on human lymphocyte proliferation and interferon  $\gamma$  production (8), induction of paw oedema and vasodilator effects in rats (9), nitric oxide production (10, 11), histamine release from rat peritoneal mast cells (12), *in vivo* lymphocyte activation and apoptosis (13), antidepressant-like effect in mice (14) and differentiation of human colon cancer cells (15), among others.

Differences in biological activity are associated with factors such as pH-dependent oligomerization, differences in the position of amino acids in the carbohydrate-binding site (6, 7, 16) and minor changes in residues in key positions of the quaternary structure (17, 18).

Structural studies can help clarify the molecular mechanisms involved in the biological activities of different lectins, identify important aspects of protein-carbohydrate interaction and show how lectins decode the large amounts of information stored in glycan structures.

Thus, the purpose of the present study was to purify and solve the complete primary and crystal structure of the lectin of *Camptosema pedicellatum* (CPL) complexed with 5-bromo-4-chloro-3-indolyl- $\alpha$ -D-mannose (X-Man) using tandem mass spectrometry.

## Materials and Methods

### Plant material

Seeds of CPL were collected from plants grown in Northeastern Brazil (Crato, Ceará). The botanical identification was carried out at the Department of Biology, Federal University of Ceará (UFC).

### Lectin purification

Mature seeds of *C. pedicellatum* were ground into a fine powder using a coffee mill. The flour was stirred with 0.15 M NaCl (1:10, m/v) for 3 h at room temperature. The mixture was centrifuged at 10,000  $\times g$  for 20 min at 4°C. The clear supernatant (crude extract) was applied to a Sephadex G-50 column (a cross-linked dextran) (15 cm  $\times$  2 cm) equilibrated with 150 mM NaCl pH 7.0 containing 5 mM CaCl<sub>2</sub> and 5 mM MnCl<sub>2</sub>. After washing of the unbound material in the equilibrium solution, the lectin was eluted from the gel with 100 mM glucose, pooled, exhaustively dialysed against distilled water and freeze-dried. Absorbance at 280 nm was used to estimate the protein concentration in all chromatographic fractions. The purification process was monitored by SDS-PAGE as described elsewhere (19) and the purified lectin was used for mass spectrometry analysis and crystallization trials.

### Hemagglutination activity and inhibition assays

Hemagglutination assays were carried out as described elsewhere (20) using serial dilutions with rabbit erythrocytes, either native or treated with proteolytic enzymes (trypsin or papain). Results were expressed in hemagglutinating units (HU), with one HU being defined as the smallest amount (mg) of protein per

millilitre capable of inducing visible agglutination. The lectin carbohydrate-binding specificity was defined as the smallest sugar concentration capable of fully inhibiting agglutination. Two-fold serial dilutions (initial concentration: 100 mM) of D-glucose, D-galactose, D-mannose, D-arabinose, D-fructose, D-fucose, D-xylose, *N*-acetyl-D-glucosamine,  $\alpha$ -methyl-D-galactopyranoside,  $\alpha$ -methyl-D-mannopyranoside, lactose and carrageenan were prepared in 150 mM NaCl. Lectin (4 HU) was added to each dilution.

### MW determination by mass spectrometry

The molecular mass of CPL was determined by electrospray ionization mass spectrometry (ESI-MS) using a hybrid mass spectrometer (the Synapt HDMS system, Waters Corp., Milford, USA) operating in positive ion mode at 10,000 resolution. Protein solution (10 pmol/ $\mu$ l) was infused into the system using the built-in syringe drive at a flow rate of 10  $\mu$ l/min. The capillary voltage and the cone voltage were set at 3 kV and 40 V, respectively. The source temperature was maintained at 100°C and nitrogen was used as a drying gas (flow rate: 150 l/h). Data acquisition was done with the software Mass Lynx 4.0 and the multiply charged spectra were deconvoluted using maximum entropy techniques (21).

### Protein digestion and tandem mass spectrometry analysis

The protein digestion and tandem mass spectrometry analysis were carried out as described by Shevchenko *et al.* (22). Thus, protein was submitted to SDS-PAGE and the bands were excised and bleached in a solution of 50 mM ammonium bicarbonate in 50% acetonitrile. The bands were then dehydrated in 100% acetonitrile and dried in a speedvac (LabConco). The gel was rehydrated with a solution of 50 mM ammonium bicarbonate containing trypsin (Promega) or chymotrypsin (Sigma) (1:50 w/w; enzyme:substrate) at 37°C overnight. Subsequently, the peptides were extracted in a solution of 50% acetonitrile with 5% formic acid and concentrated in a speedvac. In addition, a lectin aliquot was dissolved in 2.0 M urea in 50 mM sodium phosphate (pH 6.8) and digested for 24 h at 37°C with endoproteinase Glu-C (Roche) obtained from *Staphylococcus aureus* (1:100 w/w; enzyme:substrate). The peptides were separated with a C18 chromatography column (75  $\mu$ m  $\times$  100 mm) using a nanoAcquity<sup>TM</sup> system and eluted with acetonitrile gradient (10–85%) containing 0.1% formic acid. The liquid chromatograph was connected to a nanoelectrospray mass spectrometer source (Synapt HDMS system, Waters Corp., Milford, USA). The mass spectrometer was operated in positive mode, using a source temperature of 80°C and a capillary voltage of 3.5 kV. The instrument was calibrated with double-protonated [Glu1]-fibrinopeptide B ions ( $M + 2H^+$ )  $2^+ = 785.84$ . The LC-MS/MS experiment was done with the DDA (data-dependent acquisition) function selecting for experiments with double or triple-charged precursor ions, which were fragmented by collision-induced dissociation (CID) with the ramp collision energy level adjusted according to the charge state of the precursor

ion. The data were managed and analysed with the software Proteinlynx (Waters), using 'peptide fragmentation pattern' as search parameter. Some peptide sequences were obtained by *de novo* manual sequencing followed by manual interpretation of CID spectra.

### Primary structure analysis

The primary sequence alignments and secondary structure predictions were made with the software ESPript 2.2 (23). The theoretical pI was calculated with ProtParam (24).

### Crystallization and data collection

The freeze-dried purified lectin was resuspended in Milli-Q<sup>TM</sup> water containing 3 mM X-Man at a final concentration of 10 mg ml<sup>-1</sup> and incubated at 310 K for 1 h prior to the crystallization experiments. The crystals grew on 24-well Linbro<sup>TM</sup> plates at room temperature (293 K) by the vapour-diffusion method (25) in hanging drops using Crystal Screens I and II (Hampton Research, Riverside, CA, USA). Each well contained 300  $\mu$ l reservoir solution and the drops were composed of equal amounts (2  $\mu$ l) of protein solution and reservoir solution.

X-ray data were collected from a single crystal cooled to 100 K. Crystals were previously soaked in a cryoprotectant solution of water (70%) and PEG 400 (30% v/v) to avoid ice formation and submitted to X-ray diffraction at a wavelength of 1.42 Å using a synchrotron-radiation source (MX1 station, Laboratório Nacional de Luz Síncrotron, Campinas, Brazil). A complete data set was obtained using a MarCCD 165 mm (MAR Research) in 180 frames with an oscillation range of 1°. The data set was indexed and integrated using MOSFLM (26). Intensities were reduced through SCALA (27).

### Molecular replacement and refinement

The crystal structure of CPL was determined by molecular replacement using the program MolRep (28). The atomic coordinates used as model were those of recombinant *Dioclea grandiflora* (lectin of DGL) complexed with X-Man (PDB code 2JEC) (7), with a final correlation coefficient of 62.4% and an  $R_{\text{factor}}$  of 42.1% after calculation of rotation and translation functions.

The initial structure was submitted to rigid body and restrained refinement with the program REFMAC5 (29).  $R_{\text{factor}}$  and  $R_{\text{free}}$  converged to 35.96 and 36.02%, respectively. Subsequently, the structure was modelled using WINCOOT. The loop region (117–123) was adjusted to satisfy the electron density map and 77 water molecules were added to the structure. The second restrained refinement yielded an  $R_{\text{factor}}$  of 19.0% and an  $R_{\text{free}}$  of 23.6%. The stereochemical quality of the model was inspected by Ramachandran plot using PROCHECK program (29), Van der Waals contacts, polar contacts and hydrophobic interaction were analysed with the CCP4 software CONTACT, adopting the cut-off distances 3.5, 3.5 and 5.0 Å, respectively. All figures and superpositions were performed with the program

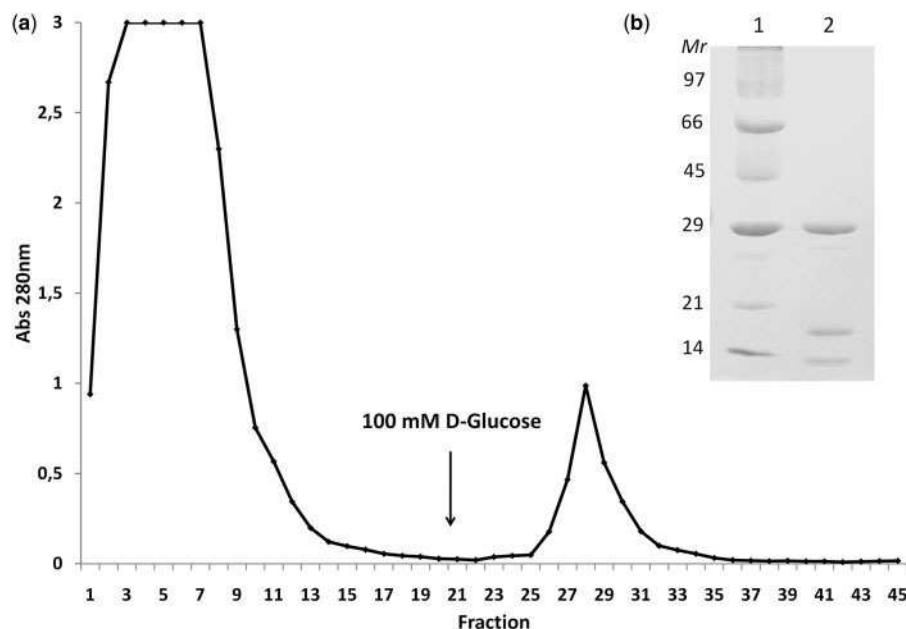
PyMOL (30). The atomic coordinates for the structure were deposited in PDB with access code 3U4X.

## Results and Discussion

CPL was purified in a single step using a Sephadex G-50 affinity chromatography column, in which the lectin was quantitatively retained in the cross-linked dextran gel column and desorbed with D-glucose, providing strong evidence of carbohydrate-binding properties (Fig. 1a). This procedure has been widely used for the purification of Diocleinae lectins (16, 31–33). Hemagglutinating activity towards native and enzyme-treated rabbit erythrocytes was fully inhibited by D-mannose (12.5 mM), D-glucose and  $\alpha$ -methyl-D-mannopyranoside (3.16 mM), but not by as much as 100 mM D-galactose, D-arabinose, D-fructose, D-fucose, D-xylose, N-acetyl-D-glucosamine,  $\alpha$ -methyl-D-galactopyranoside, lactose or carrageenan (Table I). Similar carbohydrate-binding specificities have been reported for other Diocleinae lectins (16, 34). Furthermore, CPL displayed greater affinity for  $\alpha$ -methyl-D-mannopyranoside than for D-mannose, indicating the presence of a more strongly hydrophobic substituent at C-1, conceivably allowing for additional interactions with hydrophobic regions of the lectin-binding site (34).

The electrophoretic profile obtained with SDS-PAGE of the affinity-purified CPL, both in the presence and absence of  $\beta$ -mercaptoethanol, revealed a major band of 30 kDa ( $\alpha$ -chain) and two minor components ( $\beta$ -chain and  $\gamma$ -chain) of 16 and 13 kDa, respectively (Fig. 1b), suggesting CPL is a typical ConA-like lectin subject to the posttranslational process of circular permutation described by Carrington *et al.* (35). Thus, CPL is expressed as a pre-pro-protein ( $N_{\text{term}}$ signal peptide +  $\gamma$ -chain + linker peptide +  $\beta$ -chain +  $C_{\text{term}}$ signal peptide) cleaved into a  $\gamma$  product and a  $\beta$  product. The active protein is a final fused product ( $\alpha$ -chain) with the two fragments in inverse order and no signal or linker peptides (35–37). The  $\beta$ - and  $\gamma$ -fragments observed on SDS-PAGE (Fig. 1b) are unlinked products of this process while the  $\alpha$ -chain is the mature protein. Electrospray ionization mass spectrometry confirmed purified CPL consists of a combination of chains weighing  $25,298 \pm 2$  ( $\alpha$ -chain),  $12,835 \pm 2$  ( $\beta$ -chain) and  $12,481 \pm 2$  Da ( $\gamma$ -chain) (Fig. 2).

The complete CPL protein sequence was obtained from overlaps of 27 digested peptides sequenced by tandem mass spectrometry, resulting in 236 amino acid residues ( $\beta$ -chain: 1–118.  $\gamma$ -chain: 119–236) (Fig. 3). The isotope-averaged molecular masses calculated for the full-length  $\alpha$ -chain (25,297 Da) and its derived fragments  $\beta$  and  $\gamma$  (12,835 and 12,480 Da, respectively) are in agreement with the experimentally determined mass. Table II shows all sequenced peptides and their respective molecular mass. The theoretical pI based on the final sequence was 5.4. The protein sequencing data reported in this article are deposited in the UniProt Knowledgebase under the accession number P86894.



**Fig. 1 Purification of CPL.** (a) Elution profile of Sephadex G-50 affinity chromatography. Approximately 10 ml crude extract was applied to the Sephadex G-50 column (15 cm × 2 cm) equilibrated with 150 mM NaCl containing 5 mM CaCl<sub>2</sub> and 5 mM MnCl<sub>2</sub>. The lectin was eluted with 100 M D-glucose included in the above buffer at a flow rate of 1 ml/min. Fractions (2.0 ml) were collected and monitored for protein content at 280 nm. (b) SDS-PAGE. Lane 1: Molecular mass markers (phosphorylase b, 97 kDa; bovine serum albumin, 66 kDa; ovalbumin, 45 kDa; carbonic anhydrase, 29 kDa; trypsin inhibitor, 20.1 kDa and  $\alpha$ -lactalbumin, 14.4 kDa); Lane 2: CPL.

**Table I. Inhibitory effect of monosaccharides and disaccharides on haemagglutinating activity of CPL.**

Sugar	MIC (mM)
D-Glucose	50
D-Galactose	NI
D-Mannose	12.5
D-Arabinose	NI
D-Fructose	NI
D-Xilose	NI
D-Fucose	NI
N-Acetyl-D-glucosamine	NI
$\alpha$ -methyl-D-galactopyranoside	NI
$\alpha$ -methyl-D-mannopyranoside	3.16
Lactose	NI
Carrageenan	NI

MIC, minimum inhibitory concentration; NI, sugar not inhibitory until a concentration of 100 mM.

The primary structure of CPL is quite homologous to that of other Diocleinae lectins, such as *D. guianensis* (Dgui) (86%) (SwissProt accession code: P81637), *D. grandiflora* (DGL) (85%) (SwissProt accession code: A9J251), *D. violacea* (DVL) (81%) (PDB accession code: 2GDF), *D. wilsonii* (DwL) (85%) (SwissProt accession code: P86624), *C. ensiformis* (ConA) (81%) (SwissProt accession code: P02866), *C. brasiliensis* (ConBr) (81%) (SwissProt accession code: P55915) and *Cratylia floribunda* (CFL) (97%) (SwissProt accession code: P81517). CFL differs from CPL by only five residues: 99 (V/L), 136 (S/N), 201 (P/T), 205 (P/I) and 211 (F/W) (Fig. 4).

Some Diocleinae lectins exhibit pH-dependent dimer-tetramer equilibrium. The physiological relevance of this phenomenon remains poorly understood.

However, pH-dependent dimer/tetramer equilibrium has been reported to exert a considerable influence on oligosaccharide recognition and biological activity (7), suggesting residues 123 (A/E), 131 (N/H) and 132(Q/K) play an important role in dimer/tetramer equilibrium properties. Similarly to Dgui (17) and CFL (38), which exhibit pH-dependent dimer-tetramer transition, CPL contains the residues Ala123, Asn131 and Gln132.

One of the most important primary structural differences between CPL and other Diocleinae lectins is the substitution of Val99 for Leu99. This conservative amino acid substitution was confirmed by MS/MS analyses of the peptides T5 (Fig. 5) and E2. The product-ion spectra of the doubly charged ion at  $m/z = 541.30$  and of the singly charged ion at  $m/z = 1651.87$  (E2) produced sequence-specific y-ion series, from which the primary structures were interpreted as VGLSASTGVYK and WVRVGLSASTGVYKE, respectively. The tryptic peptide which comprises amino acids from V91 to K101 of other Diocleinae lectins sequenced by MS/MS (39–41) presented an  $m/z$  ratio of 555.45 (M+H+ 1108.90), corresponding to the peptide with Leu99. Legume lectins are known to possess three types of hydrophobic subsites based on different ligand affinities. The fact that one of these subsites is adjacent to the conserved monosaccharide binding site explains why hydrophobic glyco/mannosides and other hydrophobic derivative monosaccharides bind more strongly (10–50 times) than their non-hydrophobic analogues (42). Kanellopoulos *et al.* (43) determined the structures of two ConA complexes, ConA/4'-nitrophenyl- $\alpha$ -D-mannopyranoside







Table II. Sequenced peptides of *C. pedicellatum* lectin and their respective molecular masses.

Peptide	Experimental mass (Da)	Sequence
T1	3,298.6201	ADTIVAVELDTYPNTDIGDPNYQHIGINIK
T2	845.3484	WNVQDGK
T3	1,345.6206	VGTAHISYNSVAK
T4	3,221.5845	LSAIVSYPGSSATVSYDVLNINILPEWVR
T5	1,080.5006	VGLSASTGVYK
T6	1,512.7410	ETNTILSWSFTSK
T7	2,454.8164	TNSTADAQSLHFTFNQFSQSPK
T8	2,131.0090	DLILQGDASTDSDGNLQLTR
T9	1,287.5410	VNSGSPQSNSVGR
T10	1,460.7125	ALYYAPVHVWDK
T11	1,715.9370	SAVVASFDFATFTFLIK
T12	2,786.9766	SPDSDPADGIAFFIANTDSSIPHGSGGR
T13	958.5063	LLGFDPAN
Q1	3,573.7866	DTYPNTDIGDPNYQHIGINIKSIRSKATTRW
Q2	1,487.6444	NVQDGKVGTAHISY
Q3	1,544.8213	SAIVSYPGSSATVSY
Q4	1,850.9443	DVLDNINILPEWVGL
Q5	1,090.3043	KETNTILSW
Q6	1,847.8566	TSKLLKTNSTADAQSLHF
Q7	1,162.5244	NQFSQSPKDL
Q8	3,356.6765	ILQGDASTDSDGNLQLTRVNSGSPQSNSVGRAL
Q9	1,891.8994	TKVNSGSPQGNVGRAL
Q10	1,356.6044	DKSAVVASFDFATF
Q11	1,431.6642	IKSPDSDPADGIAF
Q12	1,840.7765	FIANTDSSIPHGSGGRL
E1	816.9266	ADTIVAVE
E2	1,651.8729	WVRVGLSASTGVYKE

CPL yielded small irregular crystals in 100 mM HEPES (pH 7.5) containing 2.0 M ammonium sulphate and 2% PEG 400 after 1 week (Fig. 6a) (Hampton Research Crystal Screen I, condition No. 39), while plate crystals were obtained in 200 mM ammonium sulphate, 100 mM sodium cacodylate pH 6.5 and 30% PEG 8000 (Fig. 6b) (Hampton Research Crystal Screen I, condition No. 15). These crystals were not suitable for X-ray diffraction experiments. However, several optimization steps were performed combining these two conditions, changing the pH and precipitant concentration. CPL crystals grew in 100 mM HEPES (pH 8.5) containing 2.0 M ammonium sulphate and 2% PEG 400 to maximum dimensions of 0.1 mm × 0.2 mm × 0.15 mm (Fig. 6c).

CPL crystals diffracted to a maximum resolution of 2.15 Å using a Synchrotron radiation source (LNLS, Campinas, Brazil). The complete data set of 180 frames was indexed, integrated and scaled in the 2.15 Å resolution range. The crystals belonged to the orthorhombic space group I222 with cell parameters  $a = 65.9$  Å,  $b = 66.7$  Å,  $c = 107.7$  Å,  $\alpha = 90^\circ$ ,  $\beta = 90^\circ$  and  $\gamma = 90^\circ$ . The Matthews coefficient  $2.37$  Å<sup>3</sup> Da<sup>-1</sup> (47) was calculated based on a molecular weight of  $25,298 \pm 2$  Da and implies that the crystal contains 48.18% solvent (indicating the presence of a monomer in the asymmetric unit) and that the biological assembly is a 222 tetramer. The data collection statistics are shown in Table III.

The refined monomer structure of CPL complexed with X-Man consists of 236 amino acids folded as a  $\beta$ -sandwich, as observed in several other legume lectins (Fig. 7a). The CPL tetramer consists of two canonical

dimers linked by salt bridges between  $\beta$ -strands (Fig. 7b). CPL features a metal-binding site with the conserved residues Asn14 and Tyr12 (interacting with calcium), Glu8 and His24 (interacting with manganese) and Asp10 and Asp19 (interacting with both). In addition, the peptide bond of Ala206 and Asn207 in the *cis* configuration is isomerized due to the presence of divalent metals changing the side chain orientation of Asn14 and Asp207—a common observation in legume lectins. The substitution of a single residue can significantly affect biological activity, as observed with site-directed mutagenesis applied to other lectins (48).

The carbohydrate-binding site is occupied by X-Man (Fig. 8). The complex involves 10 intermolecular H bonds formed between O3, O4, O5 and O6 of the X-Man complex, amino acid side chains and 8 van der Waals interactions (Table IV). Legume lectin structures feature a conserved hydrophobic subsite composed of Tyr12, Leu99 and Tyr100 (44). However, the primary structure of CPL has a conserved substitution at this site (Val99 for Leu99) (Fig. 8). This change in the hydrophobic subsite reduces the hydrophobic interactions between X-Man and the lectin. Due to the absence of a methyl group, Val99 is  $\sim 1.0$  Å closer than Leu99 to the indolyl group of X-Man. As a result, the hydrophobic interactions in this subsite strengthen the carbohydrate association in the carbohydrate binding site (42).

The hydrophobic amino acid in position 99 not only enhances lectin-carbohydrate interaction, it also affects the orientation of X-Man in the carbohydrate recognition domain (CRD). By placing the ligand at a distance of 3.98 Å, Val99 causes a change in the

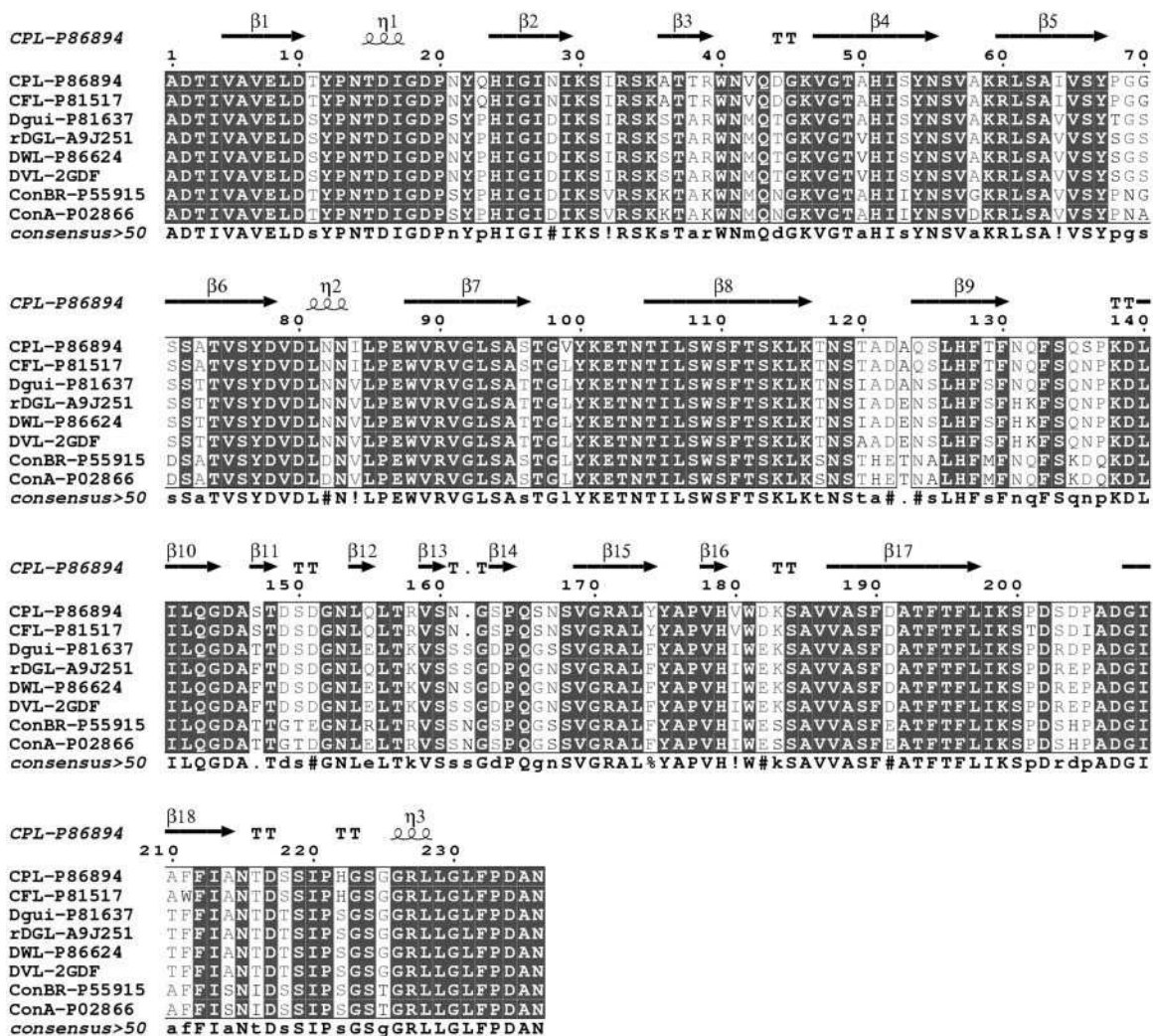


Fig. 4 Multiple sequence alignment of ConA-like lectins. Alignment of CPL, CFL, Dgui, DGL, DVL, DwL, ConBr and ConA reveals a highly conserved sequence in the secondary structure.

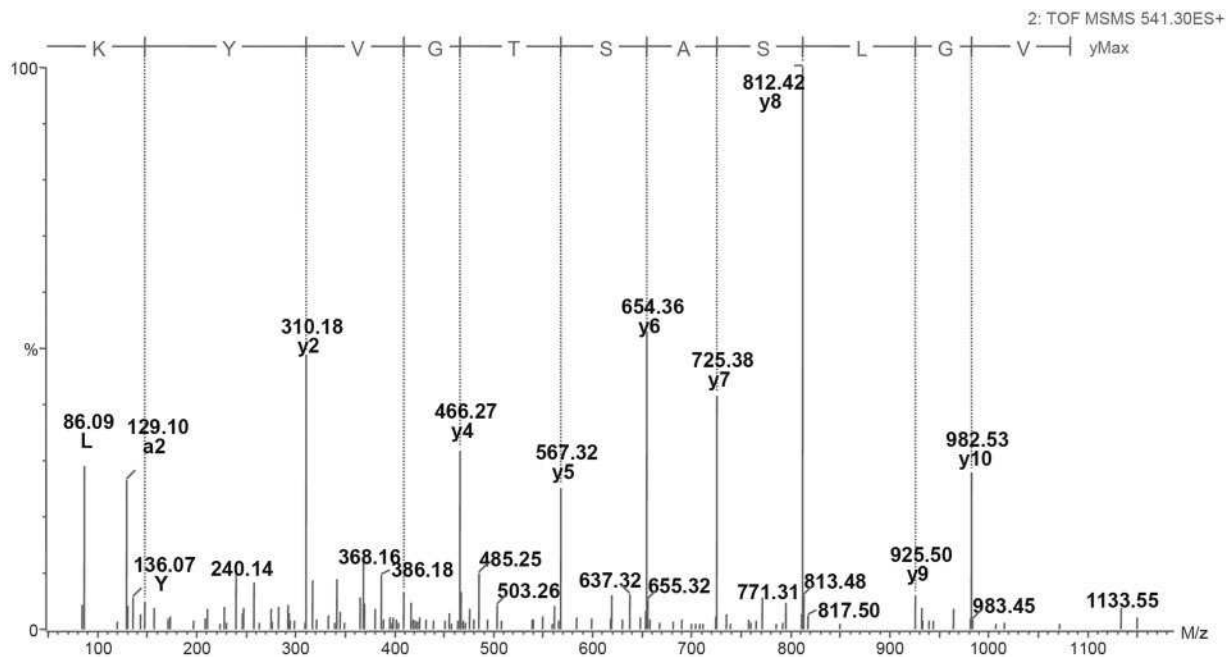
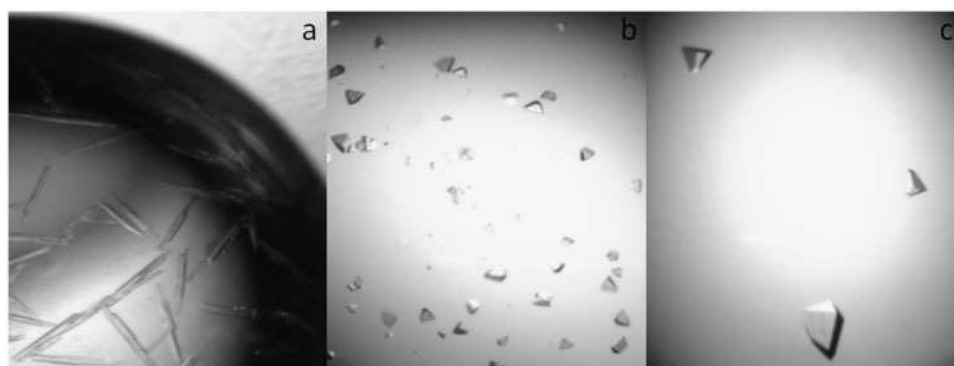


Fig. 5 Tandem mass spectrometry of T5 peptide of CPL. Collision-induced fragmentation of the doubly charged ion at  $m/z=541.3$  corresponding to the peptide T5 of CPL (see Fig. 2 and Table II). Sequence-specific y ions used for structure determination are indicated.



**Fig. 6** Crystals of CPL. (a) Plate crystals obtained with Crystal Screen I condition No. 15, (b) crystals obtained with Crystal Screen I condition No. 15 and (c) single crystal diffracted at 2.15 Å obtained with the optimization of condition No. 39.

**Table III.** Statistics of data collection, refinement and structure quality.

Parameter	Value
Data collection	
Beamline wavelength	1.42 Å
Space group	I222
Unit cell parameters (Å)	
a	65.9
b	66.7
c	107.7
Total reflections	85,239
Number of unique reflections	12,740
Molecules per asymmetric unit	1
$R_{\text{merge}}^a$ (%)	14.7 (25.0) <sup>d</sup>
Completeness (%)	96.4 (96.9) <sup>d</sup>
Multiplicity	6.7 (6.2) <sup>d</sup>
I/σ	3.4 (2.8) <sup>d</sup>
Molecular replacement	
Correlation coefficient	62.4
$R_{\text{factor}}^b$ (%)	42.1
Refinement	
Resolution range (Å)	21.5–2.1 (2.2–2.1) <sup>d</sup>
$R_{\text{factor}}^b$ (%)	19.0
$R_{\text{free}}^c$ (%)	23.6
Number of residues in asymmetric unit	236
Number of water molecules	77
RMS deviations from ideal values	
Bond lengths (Å)	0.019
Bond angles (degrees)	1.924
Temperature factor	
Average B value for whole protein chain (Å <sup>2</sup> )	17.84
Ramachandran plot	
Residues in most favoured regions (%)	95.7
Residues in additional allowed regions (%)	4.3
Residues in generously allowed regions (%)	0

<sup>a</sup> $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I(hkl)}$  where  $I(hkl)_i$  is the intensity of  $i$ -th measurement of reflection  $h$  and  $\langle I(hkl) \rangle$  is the mean value of the  $I(hkl)_i$  for all  $I$  measurements.

$$^b R_{\text{factor}} = \frac{\sum_h ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum_h |F_{\text{obs}}|}$$

<sup>c</sup>Calculated with 5% of the reflections omitted from refinement.

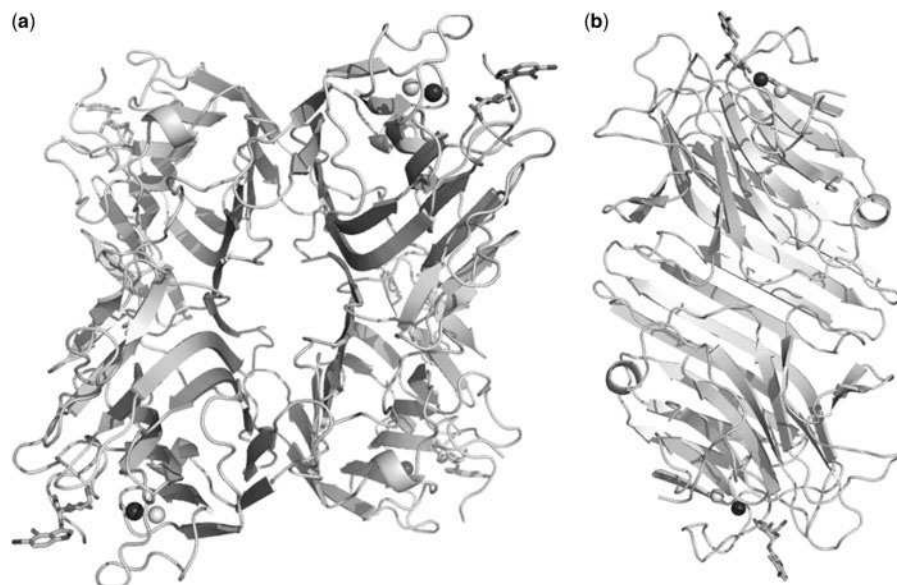
<sup>d</sup>Values in parentheses represent the high resolution shell.

position of the indolyl group in a way unfavourable to the establishment of a bond between the OH molecule in Tyr12 and the N1 component of X-Man (Fig. 9a), thereby making H-bonds with Tyr12 impossible (Fig. 9b). In other Diocleinae lectins complexed with X-Man, the distance to the hydroxyl group in Tyr12 is reduced and interaction is facilitated due to

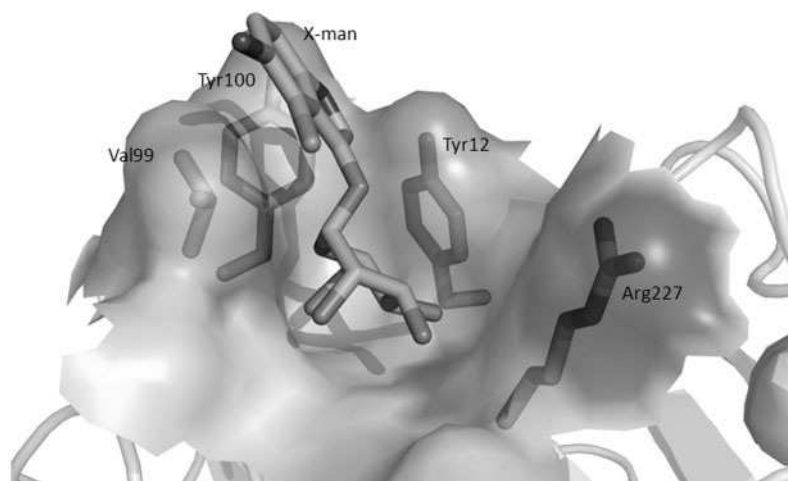
the presence of Leu99 in the hydrophobic subsite (Table V).

The CRD superposition analysis of DVL (PDB code: 3AX4) and CPL (PDB code: 3U4X) (Fig. 9b) revealed differences in the position of the indolyl group and the hydroxyls in Tyr12. Gadelha *et al.* demonstrated the importance of the hydroxyl-Tyr12





**Fig. 7** The canonical lectin dimer presented in the CPL tetramer. (a) Overall structure of CPL complexed with X-Man. (b) The two subunits of the canonical dimer.  $\text{Ca}^{2+}$  (black) and  $\text{Mn}^{2+}$  (white) are shown as spheres.



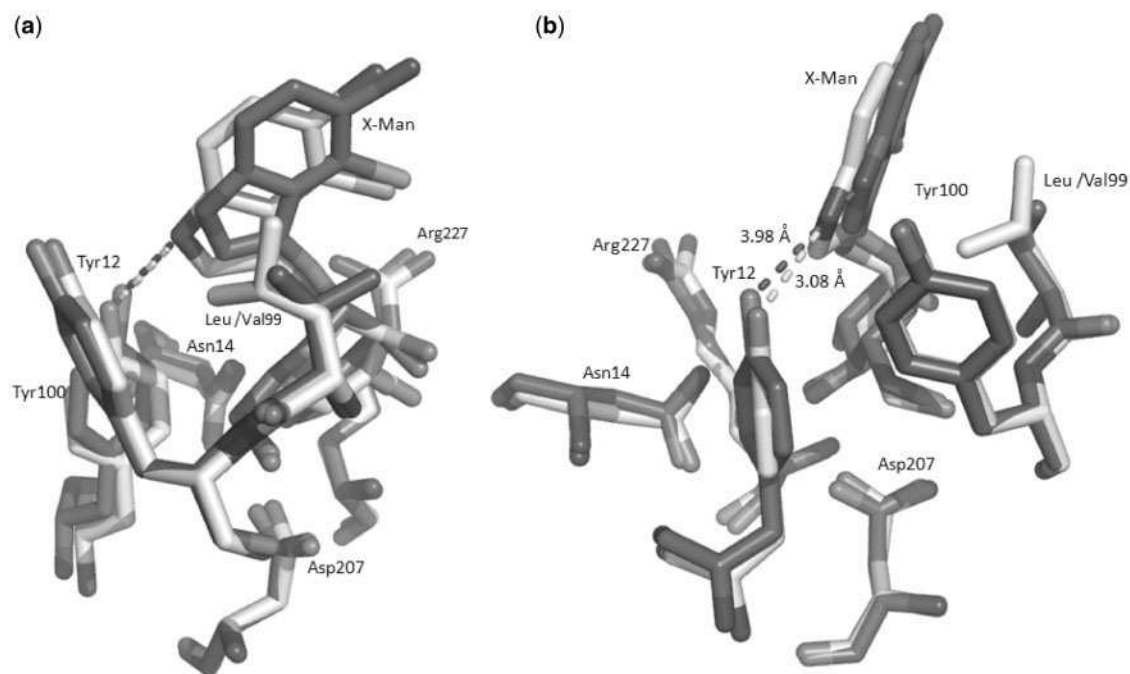
**Fig. 8** Representation of the surface area of the carbohydrate recognition domain of CPL complexed with X-Man. Amino acids that compose the hydrophobic subsite are labelled: Tyr12, Val99 and Tyr100.

**Table IV.** Van der Waals interaction and polar contacts between the lectin from *C. pedicellatum* seeds and X-Man.

Amino acids	X-Man	Distances (Å)
Van der Waals interactions		
Tyr12 OH	C11	3.38
Asp207 OD1	C4	3.39
Asp207 CG	O4	3.48
Arg227 CB	O4	3.40
Val99 CB	O5	3.44
Gly98 CA	O6	3.21
Gly98 C	O6	3.43
Ala206 CB	O6	3.27
Polar contacts		
Arg227 N	O3	2.93
Asn14 ND2	O4	3.01
Asp207 OD1	O4	2.60
Arg227 N	O4	3.19
Val99 N	O5	3.06
Gly98 N	O6	3.33
Val99 N	O6	3.01
Tyr100 O	O6	3.36
Tyr 100 N	O6	3.01
Asp207 OD2	O6	2.97

to the configuration of CRD in the lectin of *C. maritima* (ConM), associated with the substitution of Pro for Ser in position 202, bringing Tyr12 closer to the disaccharides forming H-bonds. This was later adduced to explain why ConM can interact with endothelial cell glycans and stimulate NO production by the activation of endothelial nitric oxide synthase (11). Surprisingly, the difference in position between ConM and ConA was enough to increase NO production at least 3-fold, despite 96% homology (49).

The position of the indolyl group in CPL and other ConA-like lectins differs by 0.9 Å, enough to prevent the formation of H-bonds with Tyr12. In fact, the hydrophobic interactions of Val/Leu significantly increase the association constant of the ligand in the CRD. The residue can model the ligand inside the site by hydrophobic interaction, possibly modifying the affinity of Val99-lectins and Leu99-lectins for carbohydrate complexes. This is a small but seemingly



**Fig. 9 Structural comparison between CPL (black) and DVL (PDB code: 3AX4) (grey).** (a) difference in the position of X-Man in CPL and DVL due to the substitution of Leu for Val in position 99. (b) Hydrogen bond formation between the hydroxyl group of Tyr12 and the N1 component of X-Man oriented by Leu99 in DVL (grey dashes). The presence of Val99 does not favour H-bond formation due to the reorientation of X-Man in the CRD. The ligand is positioned at a distance of 3.98 Å from the hydroxyl group of Tyr12 (black dashes).

**Table V. Distances between OH Tyr12 and N1 of the X-Man in carbohydrate recognition domain.**

Diocleinae lectin	Amino acid 99	Distance (Å)	PDB code
<i>Camptosema pedicellatum</i>	Valine	3.98	3U4X
<i>Dioclea violacea</i>	Leucine	3.08	3AX4
<i>Dioclea wilsonii</i>	Leucine	3.20	3SH3
<i>Dioclea guianensis</i>	Leucine	3.19	2JDZ

important variation within these otherwise quite homologous molecules.

The diversity of biological activities induced by Diocleinae lectins have been shown to be related to several factors, alone or in combination, including biological assembly and dimer-tetramer equilibrium (38), changes in CRD configuration induced by punctual amino acid modifications (49), CRD volume (46) and, as shown in the present study, substitutions in hydrophobic subsite amino acid residues.

The crystal structure of CPL features a conservative mutation in the hydrophobic subsite, a constituent of the CRD. This modification shows the relevance of hydrophobic interactions in the establishment of interactions between carbohydrates and hydrophobic substituents and how these interactions determine the ligand orientation at the carbohydrate binding site, enhancing the lectin-carbohydrate interaction specificity. The substitution and the analysis of the interactions with X-Man revealed that the hydrophobic effect caused by the presence in the hydrophobic subsite of Val99 instead of Leu99 made H-bond formation difficult due to the reorientation of the indolyl group in

the CRD. These structural aspects are responsible for part of the observed variability in biological activity in otherwise relatively homologous proteins.

## Acknowledgements

We also thank the Laboratório Nacional de Luz Síncrotron (LNLS), Campinas—Brazil. K.S.N., C.S.N., A.H.S., P.D., B.A.M.R. and B.S.C. are senior investigators of CNPq. We thank Jesper Sampaio for English language editing of the manuscript.

## Funding

Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP, partial), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, partial) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, partial).

## Conflict of interest

None declared.

## References

- Sharon, N. and Lis, H. (2004) History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* **4**, 53–62
- Gabius, H.J. and Gabius, S. (1997) *Glycoscience. Status and Perspectives.*, Chapman & Hall, Weinheim, Germany
- Peumans, W.J. and Van Damme, E.J.M. (1995) Lectins as plant defense proteins. *Plant Physiol.* **109**, 347–352
- Loris, R., Van Walle, I., De Greve, H., Beeckmans, S., Deboeck, F., Wyns, L., and Bouckaert, J. (2004) Structural basis of oligomannose recognition by the *Pterocarpus angolensis* seed lectin. *J. Mol. Biol.* **335**, 1227–1240

5. Van Damme, E.J.M., Peumans, W.J., Barre, A., and Rougé, P. (1998) Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. *Crit. Rev. Plant Sci.* **17**, 575–692
6. Calvete, J.J., Thole, H.H., Raida, M., Urbanke, C., Romero, A., Grangeiro, T.B., Ramos, M.V., Rocha, I.M.A., Guimarães, F.N., and Cavada, B.S. (1999) Molecular characterization and crystallization of Diocleinae lectins. *Biochim. Biophys. Acta* **1430**, 367–375
7. Nagano, C.S., Calvete, J.J., Baretino, D., Pérez, A., Cavada, B.S., and Sanz, L. (2008) 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
8. Barral-Netto, M., Santos, S.B., Barral, A., Moreira, L.I.M., Santos, C.F., Moreira, R.A., Oliveira, J.T.A., and Cavada, B.S. (1992) Human lymphocyte stimulation by legume lectins from the Diocleae tribe. *Immunol. Invest.* **21**, 297–303
9. Assreuy, A.M., Fontenele, S.R., Freitas, P.A., Fernandes, D.C., Rodrigues, N.V., Bezerra, E.H., Moura, T.R., Nascimento, K.S., and Cavada, B.S. (2009) Vasodilator effects of Diocleinae lectins from the *Canavalia* genus. *N-S Arch. Pharmacol.* **380**, 509–521
10. Andrade, J.L., Arruda, S., Barbosa, T., Paim, L., Ramos, M.V., Cavada, B.S., and Barral-Netto, M. (1999) Lectin-induced nitric oxide production. *Cell. Immunol.* **194**, 98–102
11. Gadelha, C.A.A., Moreno, F.B.M.B., Santi-Gadelha, T., Cajazeiras, J.B., Rocha, B.A.M., Assreuy, A.M.S., Mota, M.R.L., Pinto, N.V., Meireles, A.V.P., Borges, J.C., Freitas, B.T., Canduri, F., Souza, E.P., Delatorre, P., Criddle, D.N., Azevedo, W.F. Jr, and Cavada, B.S. (2005) Native crystal structure of a nitric oxide releasing lectin from the seeds of *Canavalia maritima*. *J. Struct. Biol.* **152**, 185–194
12. Gomes, J.C., Rossi, R.R., Cavada, B.S., Moreira, R.A., and Oliveira, J.T.A. (1994) Histamine release induced by glucose (mannose)-specific lectins isolated from Brazilian beans. Comparison with concanavalin A. *Agent Actions* **41**, 132–135
13. Barbosa, T., Arruda, S., Cavada, B.S., Grangeiro, T.B., Freitas, L.A.R., and Barral-Netto, M. (2001) In vivo lymphocyte activation and apoptosis by lectins of the diocleinae subtribe. *Mem. Inst. Oswaldo Cruz.* **96**, 673–678
14. Barauna, S.C., Kaster, M.P., Nascimento, K.S., Rossi, F.M., Teixeira, E.H., Cavada, B.S., Rodrigues, A.L.S., and Leal, R.B. (2006) Antidepressant-like effect of lectin from *Canavalia brasiliensis* (ConBr) administered centrally in mice. *Pharmacol. Biochem. Behav.* **85**, 160–169
15. Pinto, V.P.T., Teixeira, E.H., Teixeira, A.H., Carneiro, V.A., Cristino-Filho, G., Dus, D., Debray, H., Sampaio, A.H., and Cavada, B.S. (2010) Lectins isolated from Brazilian beans as markers of membrane glycoconjugates of human colon cancer cells. *J. Cancer Res. Exp. Oncol.* **2**, 54–59
16. Cavada, B.S., Barbosa, T., Arruda, S., Grangeiro, T.B., and Barral Netto, M. (2001) Revisiting proteus: do minor changes in lectin structure matter in biological activity? Lessons from and potential biotechnological uses of the Diocleinae subtribe lectins. *Curr. Protein Pept. Sci.* **2**, 123–135
17. Wah, D.A., Romero, A., Gallego del Sol, F., Cavada, B.S., Ramos, M.V., Grangeiro, T.B., Sampaio, A.H., and Calvete, J.J. (2001) 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
18. Brinda, K.V., Mitra, N., Suroliya, A., and Vishveshwara, S. (2004) Determinants of quaternary association in legume lectins. *Protein Sci.* **13**, 1735–1749
19. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685
20. Moreira, R.A. and Perrone, J.C. (1977) Purification and partial characterization of a lectin from *Phaseolus vulgaris*. *Plant Physiol.* **59**, 783–787
21. Ferrige, A.G., Seddon, M.J., Green, B.N., Jarvis, S.A., and Skilling, J. (1992) Disentangling electrospray spectra with maximum entropy. *Rapid Commun. Mass Spectrom.* **6**, 707–711
22. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., and Mann, M. (2006) In gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860
23. Gouet, P., Robert, X., and Courcelle, E. (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structure of proteins. *Nucleic Acids Res.* **31**, 3320–3323
24. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExpASY Server in *The Proteomics Protocols Handbook* (Walker, J.M., ed.), pp. 571–607, Humana Press, Totowa, NJ
25. Jancarik, J. and Kim, S.H. (1991) Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* **24**, 409–441
26. Leslie, A.G.W. (1997) *MOSFLM User's Guide: MOSFLM Version 550*. MRC Laboratory of Molecular Biology, Cambridge, England
27. Evans, P.R. (1997) Scala. joint CCP4 and ESF-EACBM. *Newsletter* **33**, 22–24
28. Vargin, A. and Teplyakov, A. (1997) MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30**, 1022–1025
29. Collaborative Computational Project, Number 4. (1994) The CCP4 suite: programs for crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**, 760–763
30. Delano, W.L. (2002) *The Pymol Molecular Graphics System*. DeLano Scientific, San Carlos, CA
31. Cavada, B.S., Grangeiro, T.B., Ramos, M.V., Cordeiro, E.F., Oliveira, J.T.A., and Moreira, R.A. (1996) Isolation and partial characterization of a lectin from *Dioclea rostrata* Benth seeds. *R. Bras. Fisiol. Veg.* **8**, 31–36
32. Oliveira, J.T.A., Cavada, B.S., and Moreira, R.A. (1991) Isolation and partial characterization of a lectin from *Cratylia floribunda* seeds. *Revist. Brasil. Bot.* **14**, 61–66
33. Ceccatto, V.M., Cavada, B.S., Nunes, E.P., Nogueira, N.A.P., Grangeiro, M.B., Moreno, F.B.M.B., Teixeira, E.H., Sampaio, A.H., Alves, M.A.O., Ramos, M.V., Calvete, J.J., and Grangeiro, T.B. (2002) Purification and partial characterization of a lectin from *Canavalia grandiflora* Benth. seeds. *Protein Peptide Lett.* **9**, 67–73
34. Ramos, M.V., Moreira, R.A., Cavada, B.S., Oliveira, J.T.A., and Rouge, P. (1996) Interaction of lectins from the sub tribe Diocleinae with specific ligands. *R. Bras. Fisiol. Veg.* **8**, 193–199



35. Carrington, D.M., Auffret, A., and Hanke, D.E. (1985) Polypeptide ligation occurs during post-translational modification of Concanavalin A. *Nature* **313**, 64–67
36. Cunningham, B.A., Hemperly, J.J., Hopp, T.P., and Edelman, M. (1979) Favin versus concanavalin A: circularly permuted amino acid sequences. *Proc. Natl Acad. Sci. USA* **76**, 3218–3222
37. Chrispeels, M.J., Hartl, P.M., Sturm, A., and Faye, L. (1986) Characterization of the endoplasmic reticulum-associated precursor of concanavalin A. *J. Biol. Chem.* **261**, 10021–10024
38. Gallego del Sol, F., Cavada, B.S., and Calvete, J.J. (2007) Crystal structures of *Cratylia floribunda* seed lectin at acidic and basic pHs. Insights into the structural basis of the pH-dependent dimer–tetramer transition. *J. Struct. Biol.* **158**, 1–9
39. Delatorre, P., Rocha, B.A.M., Simões, R.C., Pereira-Júnior, F.N., Silva, H.C., Bezerra, E.H.S., Bezerra, M.J.B., Marinho, E.S., Gadelha, C.A.A., Santi-Gadelha, T., Farias, D.L., Assreuy, A.M.S., Marques-Domingos, G.F.O., Nagano, C.S., and Cavada, B.S. (2011) Mass spectrometry and X-ray diffraction analysis of two crystal types of *Dioclea virgata* lectin: an antinociceptive protein candidate to structure/function analysis. *Appl. Biochem. Biotechnol.* **164**, 741–754
40. Rangel, T.B.A., Assreuy, A.M.S., Pires, A.F., Carvalho, A.U., Benevides, R.G., Simões, R.C., Silva, H.C., Bezerra, M.J.B., Nascimento, A.S.F., Nascimento, K.S., Nagano, C.S., Sampaio, A.H., Delatorre, P., Rocha, B.A.M., Fernandes, P.M.B., and Cavada, B.S. (2011) Crystallization and characterization of an inflammatory lectin purified from the seeds of *Dioclea wilsonii*. *Molecules* **16**, 5087–5103
41. Correia, J.L.A., Nascimento, A.S.F., Cajazeiras, J.B., Gondim, A.C.S., Pereira, R.I., Sousa, B.L., Silva, A.L.C., Garcia, W., Teixeira, E.H., Nascimento, K.S., Rocha, B.A.M., Nagano, C.S., Sampaio, A.H., and Cavada, B.S. (2011) Molecular characterization and tandem mass spectrometry of the lectin extracted from the seeds of *Dioclea sclerocarpa* Ducke. *Molecules* **16**, 9077–9089
42. Sharon, N. and Lis, H. (1990) Legume lectins- A large family of homologous proteins. *FASEB J.* **4**, 3198–3208
43. Kanellopoulos, P.N., Pavlou, K., Perrakis, A., Agianian, B., Vorgias, C.E., Mavrommatis, C., Soufi, M., Tucker, P.A., and Hamodrakas, S.J. (1996) The crystal structure of the complexes of concanavalin A with 40-nitrophenyl-alpha-D-mannopyranoside and 40-nitrophenylalpha-D-glucopyranoside. *J. Struct. Biol.* **116**, 345–355
44. 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. Jr, and Cavada, B.S. (2007) 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. *J. Struct. Biol.* **160**, 168–176
45. 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.M.B., Freire, V.N., Azevedo, W.F. Jr, and Cavada, B.S. (2008) 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. *J. Struct. Biol.* **164**, 177–182
46. Bezerra, E.H.S., Rocha, B.A.M., Nagano, C.S., Bezerra, G.A.A., Moura, T.R., Bezerra, M.J.B., Benevides, R.G., Sampaio, A.H., Delatorre, P., and Cavada, B.S. (2011) Structural analysis of Conbr reveals molecular correlation between the carbohydrate recognition domain and endothelial NO synthase activation. *Biochem. Biophys. Res. Commun.* **408**, 566–570
47. Matthews, B.W. (1968) Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497
48. Hirabayashi, J. and Kasai, K. (1991) Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin. *J. Biol. Chem.* **266**, 23648–23653
49. Delatorre, P., Rocha, B.A.M., Gadelha, C.A.A., Santi-Gadelha, T., Cajazeiras, J.B., Souza, E.P., Nascimento, K.S., Freire, V.N., Sampaio, A.H., Azevedo, W.F. Jr, and Cavada, B.S. (2006) Crystal structure of a lectin from *Canavalia maritima* (ConM) in complex with threose and maltose reveals relevant mutation in ConA-like lectins. *J. Struct. Biol.* **154**, 280–286