UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José Do Rio Preto

FREDERICO BRUNO MENDES BATISTA MORENO

Estudos Estruturais de uma Lectina presente em sementes de *Lotus tetragonolobus*

São José do Rio Preto – SP
Janeiro de 2008

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Campus de São José Do Rio Preto

FREDERICO BRUNO MENDES BATISTA MORENO

Estudos Estruturais de uma Lectina presente em sementes de *Lotus tetragonolobus*

Tese apresentada ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista, Campus de são José do Rio Preto, para obtenção do título de Doutor em Biofísica Molecular (Área de concentração: Cristalografia de Proteínas).

Orientador: Prof. Doutor Walter Filgueira de Azevedo Júnior.

São José do Rio Preto – SP
Janeiro de 2008

Moreno, Frederico Bruno Mendes Batista
Estudos estruturais de uma lectina presente em sementes de *Lotus Tetragonolobus* / Frederico Bruno Mendes Batista Moreno. – São José do Rio Preto : [s.n.], 2008.
119 f.: il. ; 30 cm

Orientador: Walter Filgueira de Azevedo Junior
Co-orientador: Valmir Fadel
Tese (doutorado) – Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas

1. Biofísica molecular. 2.Proteínas – Estrutura. 3. Cristalografia. 4.Lectinas. I. Azevedo Junior, Walter Filgueira. II Fadel, Walmir. III. Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas. IV. Título

CDU – 577.112

Ficha catalográfica elaborada pela Biblioteca do Ibilce
Campus de São José do Rio Preto - UNESP

O presente trabalho foi realizado no Laboratório de Sistemas Biomoleculares (BMSYS), Departamento de Física, Instituto de Biociências, Letras e Ciências Exatas – IBILCE de São José do Rio Preto – SP, Universidade Estadual Paulista – UNESP, sob orientação do Prof. Dr. Walter Filgueira de Azevedo Júnior, co-orientação do Prof. Dr. Valmir Fadel, com auxílio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

Dedico ao meu filho *Bruno*, à minha esposa *Márcia* e aos meus pais *José Mendes* e *Francisca Moreno*, pela atenção, tolerância e incentivo.

AGRADECIMENTOS

Agradeço a todos que de alguma forma contribuíram para a conclusão deste trabalho.

Agradeço especialmente à equipe de Pós-graduação do BMSYS:
Marisa, Márcio, Henrique, Denis, Denise, Diego, Helen, Ana Helena, Karisa, Nelson, José Renato e Marcos Michel.

Aos alunos de iniciação científica:
Brianna, Daina, Janaina, Joane, Angélica, Lívia, Mariane, Michele, Nathália, Hugo, Nelson, Samantha, Bona e Thaís.

Aos amigos que acompanharam de perto a conclusão deste trabalho:
Taianá, Magno e Gustavo. Comportaram-se como verdadeiros amigos e foram essenciais para as conclusões que levantei no decorrer das investigações.

Aos meus familiares, em especial aos meus pais e a minha irmã Gabriela Moreno e ao meu irmão Mendes. Ao meu sogro, José Serra e a minha sogra, Fernanda Serra.

AGRADECIMENTOS ESPECIAIS

Ao Professor Dr. *Walter Filgueira*,
Pelo incentivo e ensinamentos. A sua crença pelos trabalhos foi essencial para que eu pudesse desenvolver esta tese. Além de tudo, considero o Prof. Walter um grande amigo.

Ao Professor Dr. *Valmir Fadel*,
Por todas as formas de ensinamento.

À Professora Dra. *Fernanda Canduri*,
Pela educação, incentivo e compreensão. Obtive muito apoio da Dra. Canduri durante o primeiro ano que passei em Rio Preto longe da família. Sua presença foi essencial.

Ao Professor Dr. *Ramon Abrego*,
Pela colaboração, respeito e humildade. Foi a primeira vez que conheci alguém tão competente e ao mesmo tempo extremamente discreto.

Ao Professor Dr. *Geraldo Nery*
Que com o pouco tempo que o conheço, senti que seu bom humor torna o ambiente muito mais descontraído. E pelas inúmeras lições de vida que vão me servir de agora em diante.

SUMÁRIO

LISTA DE ABREVIAÇÕES

RIP – Proteína de inativação do ribossomo

ConBr – Lectina de *Canavalia brasiliensis*

PNA – Aglutinina do amendoim (Peanut)

Xyl - Xilose

Man – Manose

Gal – Galactose

GlicNAc – *N*-acetyl-glicosamina

GalNAc – *N*-acetyl-glucosamina

Fuc – Fucose

ConA – Lectina de *Canavalia ensiformis* (Concanavalina A)

LTA – Aglutinina de *Lotus tetragonolobus*

EcorL – Lectina de *Erythrina coralodendron*

DB58 – Lectina de *Dolichus biflorus*

TIM Barril – Estrutura quaternária, barril beta

PPL – Lectina de *Parkia platycephala*

N – Glicosilação do tipo “N”

O – Glicosilação do tipo “O”

GSL – Lectina de *Griffonia simplicifolia*

UEA – Aglutinina de *Ulex europaeus*

Mn – Manganês

Ca – Cálcio

GS1 – Tetrâmero da lectina de *Griffonia simplicifolia*

Crotina – Lectina de *Crotun tiglium*

Abrina – Lectina de *Abrus precatorios*

Robina – Lectina de *Robinia pseudoacacia*

LNLS – Laboratório Nacional de Luz Síncrotron

ÍNDICE DE FIGURAS e LEGENDAS

RESUMO

Esta tese tem como foco o estudo estrutural de uma lectina presente em sementes da espécie vegetal *Lotus tetragonolobus* (LTA). Inicialmente a LTA, previamente purificada, foi cristalizada. Os cristais foram obtidos a uma temperatura constante de 20ºC, durante 30 dias, utilizando a técnica de cristalização por difusão de vapor. Dois conjuntos de dados foram coletados a 2.00 e 2.35 Å de resolução, através da fonte de Raios X do Laboratório Nacional de Luz Síncrotron (Campinas – Brasil). Os cristais são monoclínicos e apresentam simetria do tipo $P2_1$, com parâmetros de cela de a=68.89, b=65.83 e c=102.53 Å. A substituição molecular foi feita utilizando-se o monômero da lectina presente na espécie *Arachis Hypogae* (Peanut). O homotetrâmero, produto da substituição, foi utilizado como ponto de partida para o refinamento cristalográfico. Diversos ciclos de refinamento por satisfação das restrições parciais foram feitos até a conversão total para valores satisfatórios de R_{free} e R_{factor}. A análise da estrutura revelou que a LTA possui uma forma tetramérica não identificada dentre outras lectinas de leguminosas. Sua estrutura é composta por dois dímeros, um deles semelhante ao dímero GS4 presente na lectina de *Griffonia simplicifolia* e outro que é único, caracterizado por ser um dímero do tipo LTA-dímero. Diversos fatores são responsáveis pela forma de tetramerização diferenciada da LTA, dentre os quais podemos destacar a influência da glicosilação no resíduo ASN4. Para investigarmos a estrutura da LTA em meio aquoso foi utilizada a técnica de espalhamento de Raios X a baixo ângulo (SAXS), que mostrou que a LTA possui a mesma forma tetramérica em solução da que foi observada no retículo cristalino. Nossas

investigações mostram que a disposição espacial dos sítios de ligação a *L*-fucose no tetrâmero da LTA são fundamentais para que tal proteína possa interagir através de ligações cruzadas em 2 dimensões com oligossacarídeos divalentes.

ABSTRACT

The lectin, LTA, an agglutinin found in *Lotus tetragonolobus* seeds have been crystalized. Crystals grew at a temperature of 20º C during a month and were obtained using the vapor diffusion method. Two data sets were collected at 2.00 and 2.35 Å resolution using a sincrotron radiation source at Laboratório Nacional de Luz Síncrotron (Campinas – Brasil). The LTA crystals are monoclinic belonging to the *P21* space group with a=68.89, b=65.83 and c=102.53 Å. The Molecular replacement was performed using the *Arachis Hypogae* lectin monomer (Peanut) that yielded an homotetramer which was used to initiate the crystallographic refinement. Several steps of restrained refinement were performed to obtain the best values for R_{free} and R_{factor}. Structural analysis of the LTA showed that its tetramer adopts a new structural oligomerization in contrast of that observed for others legume lectins. Its structure contains two GS4-like dimers disposed in an unusual dimer-interface, named as LTA-dimer. Several properties are involved in the new mode of tetramerization adopted for LTA, which includes the glycosilation at ASN4. The LTA tetramer investigation at aqueous solution was perfomed by Small Angle X-Ray Scattering (SAXS), showing that LTA behaves as a tetramer in solution which corroborates with the crystalline structure. Our investigations suggest that the *L*-fucose binding sites of LTA are disposed in a conformation that permits to perform two dimensional type-2 cross-linking interaction with divalent *L*-fucosyl-oligosaccharides.

5. INTRODUÇÃO

5.1 Histórico

O primeiro relato sobre lectinas remonta do século XIX. Stillmark (1888) publicou sua dissertação relatando diversos fenômenos bioquímicos observados a partir de experimentos utilizando extratos da planta *Ricinus communis*. Inicialmente ele observou que extratos de sementes da mamona (*Ricinus communis*) e de quatro outros vegetais pertencentes à família Euphorbiaceae eram capazes de aglutinar hemácias de diversos animais (STILLMARK, 1888). Esse fenômeno, pouco conhecido na época, ocorreu devido à presença da molécula "Ricina" nestes extratos vegetais descritos acima. Trata-se da descoberta da capacidade que essa molécula tem de se ligar através de ligações cruzadas com receptores de membranas das células vermelhas, formando, aglomerados de hemácias. Esse fenômeno, denominado de hemaglutinação, despertou interesse na época, pois a formação destes coágulos eram perfeitamente visíveis, estáveis, bem avermelhados e surgiram em um curto período de tempo.

Atualmente a Ricina é conhecida por ser uma molécula composta por três domínios. Dois deles assemelham-se por serem homólogos e estruturalmente semelhantes e ambos possuem a capacidade de se ligar a *D*-galactose de uma forma reversível, ou seja, não provocam nenhum tipo de alteração conformacional na estrutura dos carboidratos relacionados com a *D*-galactose. Por outro lado, o terceiro domínio, mais conhecido como domínio enzimático da Ricina (RIP – Ribossome Inactivate Protein), consiste em uma cadeia do tipo a/ß, ligada covalentemente aos outros dois domínios já descritos, que possui atividade distinta de reconhecimento aos açúcares, ou seja, desempenha uma atividade enzimática a qual está relacionada com a

inativação de síntese protéica nos ribossomos (RUTENBER *et al*., 1991)(Figura 01).

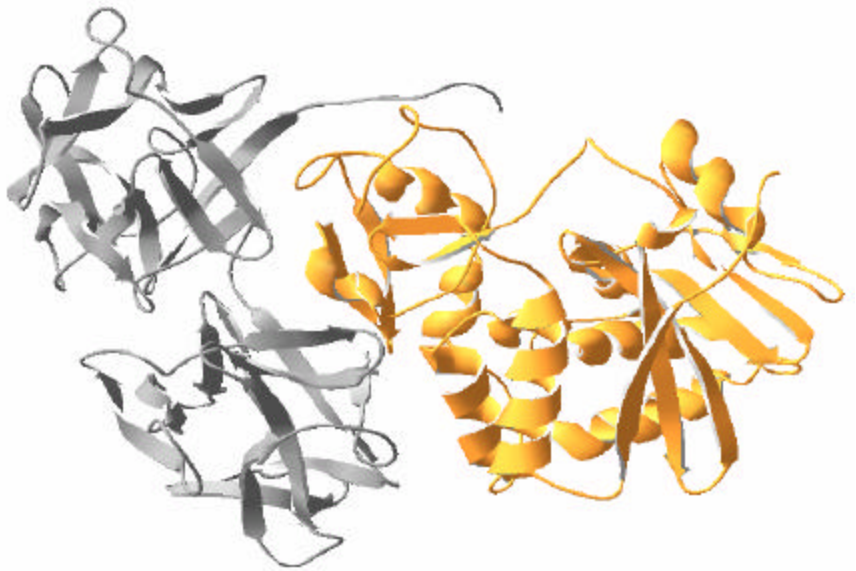


Figura 01. Ilustração da Ricina, toxina presente em sementes da espécie *Ricinus communis*. Em cinza, dois domínios homólogos de reconhecimento a *D*-galactose e em laranja domínio do tipo a/ß com atividade distinta de reconhecimento a carboidratos. Código PDB: 1ABR

Embora Stillmark (1888), em um primeiro momento, não tenha definido precisamente o fenômeno da aglutinação, deu os primeiros passos para novas pesquisas relacionadas com a investigação de extratos de diversos tecidos de vegetais. Em um segundo momento, novas pesquisas revelaram que extratos de outras espécies vegetais apresentaram a mesma capacidade de aglutinação observada no experimento com a Ricina. Tratava-se de proteínas tóxicas provenientes de extratos de *Croton tiglium* (Crotina), *Abrus precatorius* (Abrina) e de *Robinia pseudoacacia* (Robina) que foram identificadas, isoladas e apresentavam a mesma capacidade de aglutinação de células vermelhas do sangue. Inicialmente, acreditava-se que tais proteínas tinham como principal característica bioquímica a propriedade de serem

tóxicas. O que ainda não se tinha observado é que estas proteínas, além de serem capazes de inativar a síntese protéica no ribossomo de uma determinada célula, eram capazes também de se ligar reversivelmente a algum tipo de receptor presente em células, sugerindo que a aglutinação era mediada por ligações fracas como ligações de hidrogênio e forças de Van der Waals.

No início do século XX, a idéia de que as aglutininas vegetais até então descobertas eram proteínas que desempenhavam apenas atividade tóxica passou a ter uma concepção diferenciada. De acordo com as atividades biológicas desempenhadas por novas aglutininas isoladas de sementes de *Pisum sativum*, *Phaseolus vulgaris*, *Lens culinaris* e *Vicia sativa*, observou-se que estas não apresentavam atividade tóxica, entretanto ainda continuava a ser constatado que as mesmas aglutinavam eritrócitos de uma forma semelhante ao que fora visto nas Ricinas por Stillmark no século XIX. Consequentemente, um novo entendimento surgiu: embora a toxicidade de aglutininas vegetais esteja mais relacionada com fenômenos considerados como exceções, a capacidade de se ligar a receptores específicos presentes em membranas celulares é considerado uma regra geral. Portanto, lectinas sempre se ligam a carboidratos, porém, algumas podem apresentar atividades biológicas além da ação de reconhecimento de açucares.

Diversos outros experimentos utilizando outras lectinas revelaram que, realmente, esta classe de proteínas necessita se ligar a receptores presentes em membranas de células para realizarem o fenômeno da aglutinação. Entretanto, ainda não se tinha conhecimento detalhado acerca dos tipos de receptores. A dúvida se estendia também para o fato de cada lectina em questão possuir um tipo de receptor específico (VAN DAMME *et al*., 1998).

Os primeiros experimentos de aglutinação foram feitos com células sanguíneas do sistema ABO humano. Constatou-se que algumas aglutininas tinham preferência para reconhecer determinados tipos de hemácias em relação a outras, e, extraordinariamente algumas também eram capazes de aglutinar mais de um tipo de células ABO. A partir dessas observações, algumas destas proteínas foram categorizadas de acordo com a afinidade por grupos sanguíneos específicos. Consequentemente, com a descoberta de que as aglutininas são proteínas que podem ser distinguidas em relação à afinidade por células sanguíneas do sistema ABO, surgiu uma nova denominação para as aglutininas, ou seja, a partir destes resultados preliminares, as aglutininas passaram a ser chamadas de lectinas, cujo termo vem do latim “legere” e tem como significado “selecionar” (VAN DAMME *et al*., 1998).

Após os experimentos de Stillmark (1888), Summer & Howell (1936) demonstraram que, além da propriedade de aglutinar células, como eritrócitos e alguns tipos de leveduras, a ConcanavalinaA (ConA) - lectina presente em extratos de sementes de *Canavalia ensiformis* - também era capaz de precipitar moléculas de glicogênio e que esta aglutinação era inibida na presença de sacarose. De acordo com essa constatação, foi observado que a ConA interagia com o glicogênio, mas que essa interação, reversível, era revertida na presença de um açúcar menos complexo. Em outras palavras, foi possível entender que a aglutinação se dava a partir de uma reação reversível entre a ConA e carboidratos presentes na superfície dos eritrócitos. Tais resultados foram a primeira indicação experimental de que a ligação a carboidratos era a principal mediação no processo de aglutinação e que a

inibição provocada pela sacarose mostrou que tal interação era reversível (SUMMER & HOWELL, 1936).

O processo de identificação do carboidrato presente em membranas de células foi fundamental para o uso de lectinas na tipagem sanguínea, como também para a determinação de constituintes glicosilados nos diversos tecidos. Essa especificidade sanguínea observada pelas lectinas foi fundamental para se estabelecer as bases químicas da especificidade dos grupos sanguíneos do sistema ABO humano. Em 1952, Watkins & Morgan mostraram que a atividade hemaglutinante específica por eritrócitos humanos do tipo A, causada por extratos de *Phaseolus limenis*, foi inibida por *N*-acetil-galactosamina. Além disso, eles observaram que extratos de *Lotus tetragonolobus*, específicos por eritrócitos do grupo O, tiveram esta capacidade hemaglutinante inibida por α-*L*-fucose (WATKINS & MORGAN, 1952).

Atualmente as lectinas são consideradas um grande grupo de proteínas que possuem afinidade por carboidratos e que estão amplamente distribuídas na natureza, sendo encontradas em plantas, animais e organismos primitivos (LIS & SHARON, 1998). São utilizadas como ferramentas de investigação da arquitetura e dinâmica molecular de carboidratos presentes na superfície celular durante a divisão, diferenciação e processos malignos, assim como para o isolamento e caracterização de glicoconjugados (SHARON & LIS, 1993, 1995). A atividade biológica das lectinas está estreitamente relacionada com a afinidade por carboidratos. O mecanismo de reconhecimento de açúcares pelas lectinas é desenvolvido independentemente em diversos sistemas biológicos. Muitas lectinas desempenham diversas atividades biológicas como comunicação celular, defesa, fertilização, desenvolvimento,

infecção parasitária, metástases tumorais e inflamação (GABIUS & GABIUS, 1997).

Mesmo com a elucidação de inúmeros questionamentos a respeito da interação carboidratos-lectinas, ainda permanecem mistérios em biologia celular a respeito do entendimento de como as células sinalizam-se entre si ou com outros tecidos e de como ocorrem as respostas a estímulos de diferentes meios celulares com extrema precisão. Tais fenômenos estão relacionados com a decifração de glicocódigos encontrados nas mais diversas organelas, e isso garante que o estudo das lectinas passou a ter grande repercussão devido a estas proteínas constituírem desse sistema químico-fisiológico ainda pouco conhecido.

5.2 Definições, Classificação e Considerações Gerais

5.2.1 Principais definições e considerações gerais

Estudos preliminares apontaram que as lectinas poderiam ser definidas como proteínas ou glicoproteínas de origem não imune, que possuíam capacidade de se ligar a carboidratos, aglutinar células e precipitar glicoconjugados (GOLDSTEIN *et al*., 1980). Tal pressuposição originou-se devido à capacidade das lectinas de aglutinar células, podendo essa atividade ser interrompida/inibida na presença de açúcares simples (monossacarídeos). Entretanto, essa definição era pouco abrangente. Diversas outras proteínas, que não eram capazes de aglutinar células, foram posteriormente reconhecidas como lectinas.

Novas evidências a respeito do fenômeno da aglutinação mostraram que apenas algumas lectinas eram capazes de aglutinar células. Para ocorrer a aglutinação de células é necessário que exista pelo menos mais de um sítio de ligação a açúcar por cada estrutura quaternária, em uma determinada lectina. Em outras palavras, caso exista apenas um sítio de ligação ao carboidrato haverá interação com moléculas de açúcar através desse único sítio, mas, o fato de ocorrer apenas uma interação impede a aglutinação, pois não ocorrem ligações cruzadas e contínuas entre uma determinada lectina e receptores glicosilados de membranas celulares, ou seja, não há formação de uma malha estrutural mediada por glicanos e lectinas.

Por outro lado, já era de conhecimento de alguns pesquisadores que o reconhecimento a carboidratos era um fenômeno químico reversível, e que, não provocava nenhum tipo de mudança conformacional nos açúcares. Tais descobertas fizeram com que alguns pesquisadores da época classificassem

erroneamente as lectinas como proteínas distintas de enzimas, visto que, em relação a sua função de reconhecimento a carboidratos, não desempenhavam nenhum tipo de atividade enzimática que pudesse causar qualquer modificação estrutural na conformação nativa dos carboidratos.

Entretanto, Kocourek e Horejsi (1983) propuseram que algumas toxinas, também capazes de aglutinar células, deveriam ser consideradas como lectinas. No entanto, além de muitas dessas proteínas possuírem a capacidade de aglutinar células, desempenhavam ao mesmo tempo uma outra função, podendo ser tóxica ou enzimática. Porém, o que ainda não estava bem elucidado era o fato de que ambas as atividades eram promovidas por domínios protéicos distintos, pertencentes a uma mesma molécula. Ou seja, enquanto parte da proteína era capaz de fazer o reconhecimento do açúcar, ocasionando a aglutinação, a outra possuía a capacidade de desempenhar atividades distintas da ação de reconhecimento à açúcar (KOCOUREK & HOREJSI, 1983).

A constatação de que as lectinas tinham como atividade padrão a de interagir de uma forma reversível com carboidratos, além de poderem desempenhar outras funções distintas, levou alguns pesquisadores a definirem que lectinas eram proteínas que possuíam pelo menos um sítio de ligação a carboidratos e que se ligavam reversivelmente a mono ou oligosacarídeos (PEUMANS & VAN DAMME, 1995b).

De acordo com as proposições de Goldstein (1980), Kocourek e Horejsi (1983) e Peumans e Van Damme (1995), uma determinada proteína de só poderá ser caracterizada como lectina após serem atendidas três características fundamentais (RUDIGER & GABIUS, 2001):

i) São proteínas ou glicoproteínas que se ligam a algum tipo de carboidrato; ou seja, podem ser proteínas que exercem mais de uma função, mas que uma destas seja de reconhecimento à açúcar.

ii) Não são de origem imune;

iii) Não modificam bioquimicamente os carboidratos aos quais se ligam, ou seja, interagem de uma forma reversível.

5.2.2 Classificação das lectinas

5.2.2.1 Quanto aos aspectos estruturais

As três características fundamentais descritas anteriormente são suficientes para que uma ampla diversidade de proteínas sejam caracterizadas como pertencentes à classe de lectinas, incluindo diversos tipos de toxinas e proteínas de membrana. Embora essas características sejam úteis para se classificar tais proteínas de uma forma generalizada, algumas subdivisões foram criadas baseadas em aspectos estruturais. Van Damme (1998) e colaboradores classificaram as lectinas em quatro grupos principais:

a) Merolectinas: trata-se de proteínas compostas de somente um sítio de ligação a carboidrato. São proteínas monoméricas, e devido ao fato de possuírem um único sítio de ligação ao açúcar por estrutura, não apresentam atividade hemaglutinante ou capacidade de precipitar glicoconjugados. São consideradas, estruturalmente, as mais simples. Um exemplo comum de uma merolectina é a heveína, lectina presente no látex da seringueira, *Hevea brasiliensis* (Figura 02).

Figura 02. Representação esquemática de uma merolectina. Proteína monomérica com um único sítio ativo (VAN DAMME *et al*, 1998). A figura ilustra a estrutura da heveína, lectina da espécie *Hevea brasiliensis.* Estrutura determinada através de Ressonância Magnética Nuclear (ANDERSEN *et al*, 1993). Código PDB: 1MMC

b) Hololectinas: possuem dois ou mais domínios idênticos ou homólogos de reconhecimento a carboidratos. Neste caso são geralmente proteínas oligoméricas e a principal característica deste grupo é que possuem a capacidade de aglutinar células e precipitar gliconjugados. Apresentam-se, normalmente como dímeros ou tetrâmeros e, geralmente, têm a capacidade de precipitar glicoconjugados e de aglutinar células. Uma grande quantidade de lectinas presentes em plantas é classificada como hololectinas, as quais diferem entre si pelas diferenças estruturais apresentadas em suas formas diméricas e/ou tetraméricas. A maior parte das lectinas de leguminosas são classificadas por serem hololectinas, dentre as quais podemos citar as lectinas de *Canavalia brasiliensis* (ConBr), *Artocarpus integrifolia* (Jacalina) e *Griffonia simplicifolia* (Figura 03).

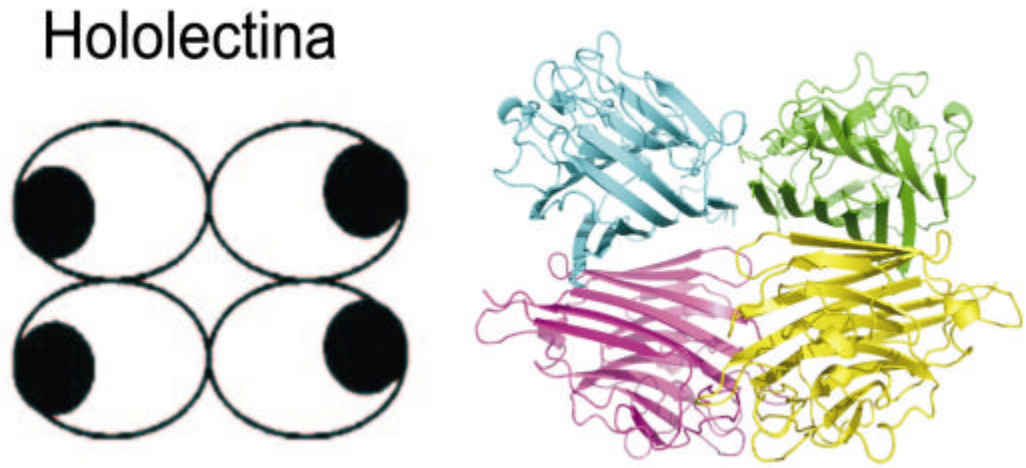


Figura 03. Representação de uma hololectina com quatro sítios ativos iguais/homólogos (VAN DAMME *et al*, 1998). A figura ilustra a estrutura da PNA, lectina da espécie *Arachis Hypogaea* na sua forma tetramérica com quatro monômeros idênticos de ligação a carboidratos (RAVISHANKAR *et al*., 2001). Código PDB: 1CQ9

c) Quimerolectinas: constituem em lectinas que possuem pelo menos um único sítio de ligação a carboidratos, e que junto a este, possui um outro domínio que exerce uma função distinta de reconhecimento a açúcar. Neste grupo geralmente são encontrados proteínas com dupla função, mas que são complementares entre si, como por exemplo, proteínas relacionadas com invasividade celular, as quais utilizam o domínio de reconhecimento a açúcar como ponto para ancoragem e, subsequentemente, agem com o outro domínio para efetivar a entrada ou até mesmo a destruição de membranas. Dependendo da quantidade de sítios de ligações a carboidrato, as quimerolectinas podem se comportar como merolectinas ou hololectinas. No caso das RIPs do tipo 2, estas apresentam mais de um domínio de ligação a carboidrato comportando-se, também, como hololectinas, como veremos a seguir na figura 04.

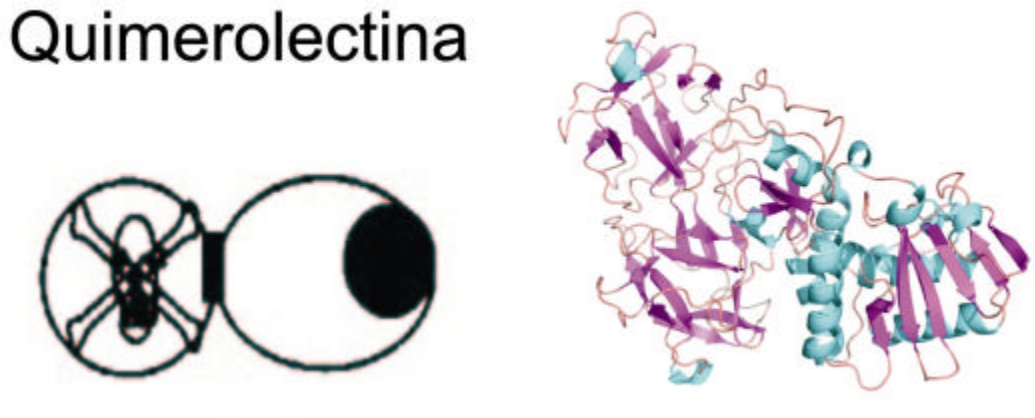


Figura 04. Representação de uma quimeroletina com um sítio de ligação a carboidrato e um outro domínio que possui uma função não lectínica (VAN DAMME *et al*, 1998). A figura 4 ilustra a estrutura da ricina, lectina da espécie *Ricinus communis,* com dois domínios homólogos de ligação a carboidrato e um outro domínio com função enzimática (RUTEMBER *et al*, 1991). Código PDB: 1ABR

d) Superlectinas: as superlectinas são classificadas como proteínas que possuem pelo menos dois sítios de ligação a carboidratos, assim como as hololectinas, porém com especificidade a açúcares distintos. São menos encontradas do que as hololectinas, porém novas estruturas para esse grupo de lectinas vêm sendo resolvidas, como, por exemplo, a estrutura da banana-lectin que possui dois sítios distintos de reconhecimento a laminaribiose e Xyl-beta1,3-Man-alpha-O-Methyl em domínios distintos (figura 05) (MEAGHER *et al*., 2005).

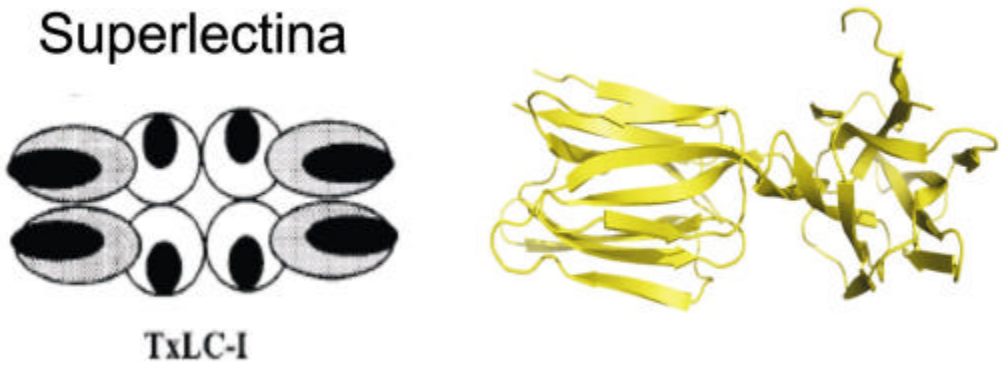


Figura 05. Ilustra o modelo esquemático de uma superlectina com dois domínios diferentes com afinidade por carboidratos distintos (VAN DAMME *et al*, 1998). A estrutura em amarelo representa a lectina da espécie *Musa acuminata* que possui dois domínios que reconhecem carboidratos distintos (MEAGHER *et al*., 2005). Código PDB: 1X1V

Inicialmente essa classificação foi feita, em quatro grupos, abrangendo significativamente inúmeras proteínas vegetais. Entretanto algumas lectinas de plantas, hoje descobertas e estruturalmente resolvidas, não se enquadram especificamente em nenhuma destas classificações. Como exemplo podemos citar a lectina, recém descoberta, presente na semente de *Parkia platycephala*, homóloga a família das hidrolases. Tal lectina possui um domínio com capacidade de reagir enzimaticamente com o polissacarídeo composto de monômeros de *N*-acetyl-D-glucosamina (quitina) e, ao mesmo tempo, dentro deste mesmo domínio, possui um outro sítio de reconhecimento a carboidrato (figura 05) (CAVADA *et al*., 2006).

Figura 06. Lectina presente na espécie *Parkia platicephala*. Proteína do tipo Tim-Barril. Desempenha duas funções distintas. Possui um sítio de ligação a carboidrato (função lectínica) e um outro sítio enzimático de quebra de polímeros de quitina (função enzimática) (CAVADA *et al.*, 2006). Código PDB: 2GSJ

5.2.3 Classificação em famílias evolutivamente relacionadas. Ênfase em lectinas de leguminosas.

O avanço da biologia estrutural em conjunto com modernas técnicas de DNA recombinante deu um salto significativo na elucidação das mais diversas moléculas presentes em estruturas dos sistemas biológicos. Atualmente, cerca de mais de 300 estruturas de lectinas de vegetais encontram-se depositadas no *Protein Data Bank* (BERMAN *et al*., 2000). Destas, além de uma grande maioria possuir um detalhamento estrutural a nível atômico (acima de 2,0 Å de resolução), muitas já possuem seus sítios de reconhecimento a carboidratos bem estabelecidos.

De acordo com o exposto no item anterior, do total de lectinas estudadas, várias apresentam semelhanças quanto a aspectos estruturais. Outras características como seqüência primária, reconhecimento a açúcares e até

mesmo quanto a função desempenhada já são bem conhecidas e estudadas. Interessa-nos observar que tais semelhanças ou discrepâncias são puramente reflexos do processo evolutivo que ocorre em relação a esta classe de proteínas. Isso ocorre de uma forma em que é possível classificar lectinas não somente quanto aos aspectos estruturais vistos anteriormente, mas como também em famílias que estejam evolutivamente relacionadas, ou seja, que não apenas possuam estruturas quaternárias semelhantes, mas que compartilham outras características que possam estar diretamente relacionadas com a função desempenhada por cada uma delas.

Dentre as famílias evolutivamente relatadas de lectinas de plantas podemos destacar:

a) **Lectinas de monocotiledôneas do tipo manose;**
b) **Lectinas específicas a quitina e homólogas a heveína;**
c) **Lectinas homólogas a Jacalina**
d) **Lectinas homólogas ao tipo RIP-2 (Ricina)**
e) **Lectinas de leguminosas;**
f) **Lectinas da família das Amarantaceae**
g) **Lectinas de floema de Curcubitaceae**

Essa classificação, feita em 1998, não é exaustiva, pois novas classificações surgem no decorrer das descobertas. Inclusive, outras lectinas estão classificadas em grupos diversos como: Apiaceae, Araucariaceae, Celastraceae, Euphorbiaceae, Gramineae, Labiatae, Solanaceae, etc (VAN DAMME *et al*.,1998).

5.2.3.1 Considerações gerais e exemplos de estruturas resolvidas de plantas

a) Lectinas de monocotiledônes do tipo manose

Essas lectinas referem-se às proteínas que são estritamente reconhecedoras de manose, e que são encontradas em monocotiledôneas. Foram classificadas, até então, seis diferentes famílias de plantas: Alliaceae, Amaryllidaceae, Araceae, Bromeliaceae, Liliaceae e Orchidaceae. Exemplares com estruturas resolvidas estão representados na figura 07.

	Lectina manose específica presente no bulbo da espécie *Scilla campanulata.* PDBs: 1B2P e 1DLP (WOOD *et al*., 1999) (WRIGHT *et al*., 2000). Encontra-se resolvida na forma nativa e em complexos com os açucares: a) *D*-manose b) Man(1-6)man c) Man1-6(man)1-3man
	Lectina manose específica presente na espécie *Narcissus pseudonarcisus* PDB: 1NPL (SAUERBOM *et al*., 1999). Encontra-se complexada com o açucar: a) Man(1-3)man.
	Lectina manose específica presente em na espécie *Allium sativum.* PDBs: 1BWU e 1KJ1 (RAMACHANDRAIAH *et al*., 2002) (CHANDRA *et al*., 1999). Encontra-se complexada com o açucar: *a) D*-manose.

	Lectina manose específica presente na espécie *Gastrodia elata.* PDBs: 1XD5 e 1XD6 (LIU *et al*., 2005). Encontra-se resolvida apenas na forma nativa.
	Lectina manose específica presente na espécie *Galantus nivalis.* PDBs: 1MSA, 1NIV e 1JPC (WRIGHT & HESTER, 1996) (HESTER & WRIGHT, 1996) (HESTER *et al*., 1995). Encontra-se resolvida na forma nativa e complexada com os açucares: a) *D*-manose-metil b) Man(1-3)man-metil c) Man1-6(man)1-3man-metil

Figura 07. Representação em forma de *ribbon* das lectinas de monocotiledônias específicas a manose. As estruturas foram resolvi das e complexadas apenas com açúcares constituídos por manose. São estruturas ricas em folhas betas interconectadas por loops.

b) Lectinas especiíficas a quitina e homólogas a heveína (Domínio heveínico)

São constituídas por lectinas que se ligam à quitina, que, por definição, compreendem todas as proteínas de plantas que possuem pelo menos um domínio semelhante a Heveína. A Heveína é isolada do látex da seringueira (*Hevea brasiliensis*) (RAIKHEL *et al.,* 1993*;* VAN DAMME *et al.,* 1998) com uma seqüência de 43 resíduos de aminoácidos que reconhece quitina e é o protótipo de proteínas dessa classe. Lectinas que se ligam à

quitina e não possuem o domínio de Heveína são classificadas em outras famílias de lectinas. A aglutinina de gérmen de trigo (WGA) é a principal representante dessa classe, é composta de duas subunidades de 18 kDa consistindo de quatro domínios de Heveína estruturalmente semelhantes. (RAIKHEL *et al.,* 1993). Separados por uma curta α-hélice de cinco resíduos, cada domínio de Heveína é destituído de estruturas secundárias regulares e, portanto consiste principalmente de estrutura em espiral (“coil’’) e curva (“turns”). Estes domínios de Heveína da WGA contêm quatro pontes dissulfeto que estabilizam a estrutura funcional da proteína. Os quatro sítios de ligação a carboidrato para GlcNAc (ou Neu5Ac) são localizados na interface entre suas subunidades (WRIGHT, 1977). Os exemplares com estruturas resolvidas estão representados na figura 08.

Lectina similar a heveína presente na espécie *Amaranthus caudatus.* PDB: 1MMC. (MARTINS *et al*., 1996). Encontra-se resolvida apenas na forma nativa.

Heveína, lectina presente na espécie *Hevea brasiliensis.* PDB: 1HEV. (ANDERSEN *et al*., 1993).Encontra-se resolvida na forma nativa e em complexos com os açucares:

a) Chitobiose
b) Chitotriose

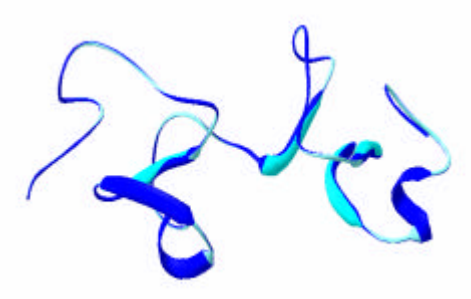
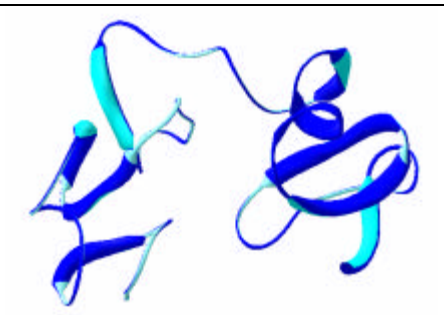
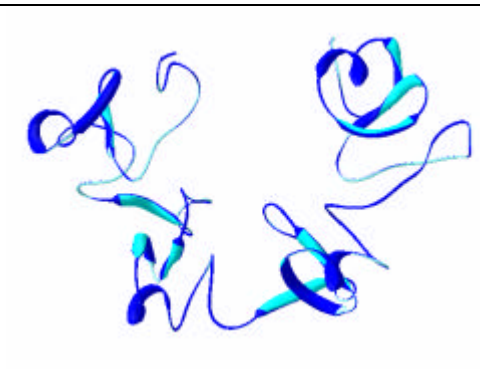
	Lectina presente na espécie *Phytolacca americana.* PDBs: 1ULN, 1UHA, 1ULK e (FUJII *et al*., 2004) (HAYASHIDA *et al*., 2003). Encontra-se resolvida na forma nativa e em complexos com o açucar: a) Chitotriose
	Lectina presente na espécie *Urtica dioica.* PDBs: 1EHD, 1EHH, 1EIS, 1ENM, 1EN2 e 1IQB. (HARATA & MURAKI, 2000) (SAUL *et al*., 2000) (HARATA *et al*., 2001). Encontra-se resolvida na forma nativa e em complexos com os açucares: a) Tri-*N*-acetylchitotriose b) *N*-acetylchitotriose c) *N*-acetylchitotetraose
	WGA. PDBs: 1WGC, 2CWG, 9WGA, 2WGC, (WRIGHT & JAEGER, 1993) (WRIGHT, 1990) (HARATA *et al*., 1995). Encontra-se resolvida na forma nativa e em complexos com os açucares: a) Sialyllactose b) Sialoglycopeptide c) *N*-acetyl-*D*-neuraminic acid d) Chitobiose e) ßGlcNAc(1-4)GlcNAc f) ßGlcNAc(1-6)Gal g) ßGlcNAc(1-6) ßGal(1-4)Glc

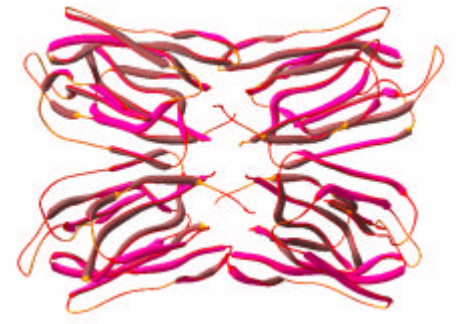
Figura 08. Representação em ribbon de lectinas com domínios homólogos ao da Heveína (*Hevea brasiliensis*). Estruturas ricas em regiões de loops, estabilizados principalmente por pontes de dissulfeto.

c) Lectinas homólogas a jacalina (lectinas do tipo Prisma-ß)

A família de lectinas relacionadas à jacalina compreende todas as lectinas de plantas que estão relacionadas estruturalmente à aglutininas presente nas sementes de *Artocarpus integrifolia* (Jaca). Basicamente, a família da jacalina compreende dois subgrupos específicos: i) que reconhecem

galactose e ii) que reconhecem manose. Esse primeiro grupo possui uma alta afinidade pelo antígeno-T [Gal β(1,3)GalNAc] (Sastry *et al*., 1986) (SARKAR *et al*., 1981).

Em relação as lectinas específicas por manose relacionadas a jacalina, estas foram identificadas em muitas espécies em grupos taxonomicamente diferentes. O primeiro membro desta subfamília de lectinas relacionadas a jacalina foi isolada de *Calystegia sepium* (erva daninha: visgo) (VAN DAMME *et al.,*1996e), posteriormente outras lectinas semelhantes foram descobertas nas sementes de *Helianthus tuberosus* (VAN DAMME *et al.,*1999), *Musa acuminata* (PEUMANS *et al.,*2000) e de arroz (ZHANG *et al.,* 2000). Sabe-se, até então, que as lectinas relacionadas a jacalina, que se ligam a manose, possuem duas ou quatro subunidades não glicosiladas de aproximadamente 150 resíduos de aminoácidos. Exemplares com estruturas resolvidas estão representados na figura 09.



Lectina presente na espécie *Artocarpus integrifolia, Artocarpus hirsuta, Maclura pomífera, Morus nigra.* Todas homólogas a jacalina. PDBs: 1J4S, 1J4T, 1J4U, (RAO *et al*., 2004)(PRATAP *et al*., 2002) (JEYAPRAKASH *et al*., 2004) (LEE *et al*., 1998)(RABIJIN *et al*., 2005) (SASTRY *et al*., 1986). Encontra-se resolvida na forma nativa e em complexos com os açucares:

a) *N*-acetylalanine, Metil-manoside, Man1-6man1-3, Man1-6(Man1-3)Man1-6(aMan1-3)Man), Metil-Galactose.
b) Antígeno-T, Porfirina, *D*-galactose, *D*-methyl-N-acetylgalactosamine
c) ßGal(13)aGalNAc-O-Me, ßGalNAc(13)aGal-O-Me, Mellibiose

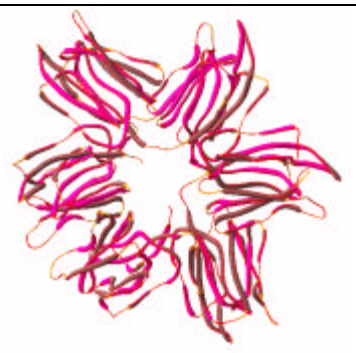
	Lectina presente na espécie *Musa acuminata.* PDBs: 2BMY, 2BMZ, 2BNO e 1X1V (MEAGHER *et al*., 2005) (MCLUSKEY *et al*., 2005) (SINGH et al., 2005). Encontra-se resolvida na forma nativa e em complexos com os açucares: a) ßXyl(1-3)aMan-O-Me b) Laminaribiose c) Metil-manoside
	Lectina presente na espécie *Calystegia sepium.* PDB: 1OUW (BOURNE *et al*., 2004). Encontra-se resolvida apenas na forma nativa.
	Lectina presente na espécie *Helianthus tuberosus.* PDB: 1C3K (BOURNE *et al*., 1999)Encontra-se resolvida na forma nativa e em complexos com os açucares: a) Man1-3man b) Man1-2man
	Lectina presente na espécie *parkia platicephala.* PDBs: 1GZR, 1GZS (GALLEGO DEL SOL *et al*., 2005)Encontra-se resolvida na forma nativa e em complexos com o açucar: a) Manose

Figura 09. Representação em *ribbon* de lectinas com domínios homólogos ao da jacalina. Estruturas ricas em regiões de folhas betas interconectadas por loops. São pobres em pontes de dissulfeto e podem ter afinidade por oligossacarídeos ricos em *D*-galactose e *D*-manose.

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

As RIPs tipo 2 constituem uma típica família de quimerolectinas constituída de uma proteína constituída por dois domínios interconectados pela extremidade *N*-terminal da cadeia A (domínio enzimático) com o C-terminal da cadeia B (domínio lectínico) (BARBIERI *et al.,* 1993). O termo “proteína que inativa ribossomos” (RIP) refere-se à capacidade destas de inativarem cataliticamente os ribossomos, que ocorre através da remoção enzimática de um resíduo de adenina especifico em um loop altamente conservado (A4324 do rRNA em rato) na grande subunidade do RNA ribossomal 28S.

As RIPs foram identificadas em espécies de várias famílias de plantas, como por exemplo, as de Euphorbiaceae (*Ricinus comunis, Cróton sp.),* Leguminosae (*Abrus precatorius),* Viscaceae (*Viscum album e Phoradendron californicum),* Passifloraceae (*Adenia digitata e A.volkensii),* Ranunculaceae (*Eranthis hyemalis),* Lauraceae (*Cinnamonum camphora),* Sambucaceae (*Sambucus sp*.), Curcubitaceae (*Mormodica charantia),* e Iridaceae (*Íris sp*.) (VAN DAMME *et al.,*1998).

A atividade e especificidade de ligação a carboidrato das RIPs são estabelecidas somente por sítios de ligação a carboidratos situados na cadeia B. Estudos de especificidade indicaram que a maioria das RIPs são efetivamente inibidas por Galactose ou GalNAc ou por ambos. Todavia os inibidores mais potentes que quaisquer outros monossacarídeos são os dissacarídeos ou oligossacarídeos ricos em galactose. A lectina de *Abrus pulchelus* possui especificidade por galactose, mas apresenta uma maior afinidade para seqüências de lactose [β-D-Gal-(1,4)-D-Glc], *N*-acetillactosamina [β-D-Gal-(1,4)-D-GlcNAc] e lacto-*N*-biose [β-D-Gal-(1,3)-D-GlcNAc]. Depois que

se constatou que a ricina e outras RIPs interagiam fortemente com oligossacarídeos carregados por galactosídeos terminais, a ocorrência de sítios estendidos de ligação a carboidratos tornou-se bastante provável de existir em outras lectinas. Os exemplares com estruturas resolvidas estão representados na figura 10.

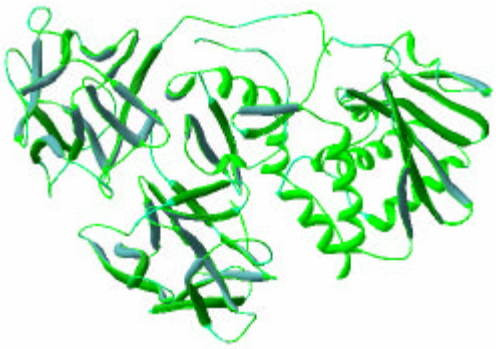
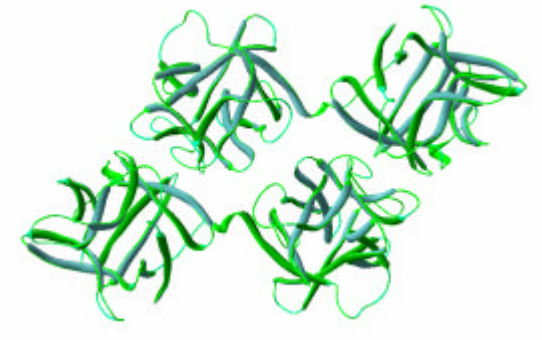
	Proteína homólogas a Ricina. Presente nas espécies *Abrus precatorius*, *Ricinus communis*, *Viscum album, Sambucus ebulus,* PDBs: 1ABR, 2AMZ, 2AAI, 1RZO, 1HWO, (BAGARIA *et al*., 2006) (TAHIROV *et al*., 1995) (RUTENBER *et al*., 1991)(PASCAL *et al*., 2001). Encontra-se resolvida na forma nativa e em complexos com os açucares: a) Gal-1β3-GalNAc-aO-benzyl b) Lactose c) Galactose d) Pteroic acid e) 2-amino-4-isopropyl-pteridine-6-carboxyl acid
	Lectina presente na espécies *Amaranthus caudatus* PDBs: 1JLX, 1JLY (TRANSUE *et al*., 1997).Encontra-se resolvida na forma nativa e em complexos com o açícar: a) Gal-1β,3-GalNAc-aO-benzyl

Figura 10. Representação em ribbon de lectinas com domínios homólogos ao Ricina. São, geralmente, quimerolectinas com estruturas ricas em regiões de folhas betas e alfa hélices. Podem ser dotadas de domínios distintos, que se ligam reversivelmente a carboidratos e que também possuem atividade enzimática.

e) Lectinas de Leguminosas

Dentre as classificações em família evolutivamente relacionadas, as lectinas mais bem estudadas pertencem ao grupo das leguminosas. Inúmeras

estruturas deste grupo foram obtidas nos mais diversos grupos de pesquisa, que conseguiram ao longo de três décadas, estabelecer padrões estruturais que definem com precisão quais características são comuns para todos exemplares até então investigados. Hoje, estas proteínas são utilizadas como excelentes modelos, para comparação e estudos estruturais de diversas outras proteínas presentes nos mais diversos organismos. Normalmente são proteínas/glicoproteínas que possuem cerca de 220 a 250 resíduos de aminoácidos e são extremamente semelhantes quanto aos aspectos estruturais. Apresentam entre si o mesmo tipo de enovelamento em relação a seus monômeros. Esses monômeros são compostos quase que exclusivamente por fitas-beta antiparalelas (cerca de 60%) arranjadas em duas ou mais folhas beta interconectadas por loops, mais conhecido como enovelamento do tipo "β-sanduíche". Cada monômero geralmente possui dois ou mais sítios de ligação a íons como também podem ser dotados de um ou mais sítios de ligação a carboidratos.

O sítio de ligação a íons é amplamente conservado nestas proteínas. Vale destacar que a atividade de reconhecimento a açúcar depende da presença do íon cálcio e de um outro íon, geralmente relatado como sendo o manganês. Esses metais se localizam a uma distância de aproximadamente 4,5 Å e ambos são coordenados por cadeias laterais de quatro aminoácidos e duas moléculas de água. O reconhecimento das lectinas de leguminosas por íons é reversível e a remoção desses metais resulta em importantes mudanças conformacionais. Com a falta dos íons cálcio e manganês, o peptídeo, Ala207-Asp208 (neste caso tomando como exemplo a ConA), que para participar efetivamente do reconhecimento ao seu carboidrato específico necessita estar na conformação

trans, passa a assumir uma conformação do tipo *cis*, ou seja, as lectinas de leguminosas podem assumir duas conformações quanto a presença dos íons cálcio e manganês. Na presença destes assumem a conformação “locked” e na ausência “unlocked”. Portanto, a conformação mais favorável do sítio de ligação a carboidratos das lectinas de leguminosas ocorre exclusivamente na presença dos íons cálcio e manganês. A figura 11 ilustra a estrutura da ConA nas formas “locked” e “unlocked” (BOUCKAERT *et al*., 1995).

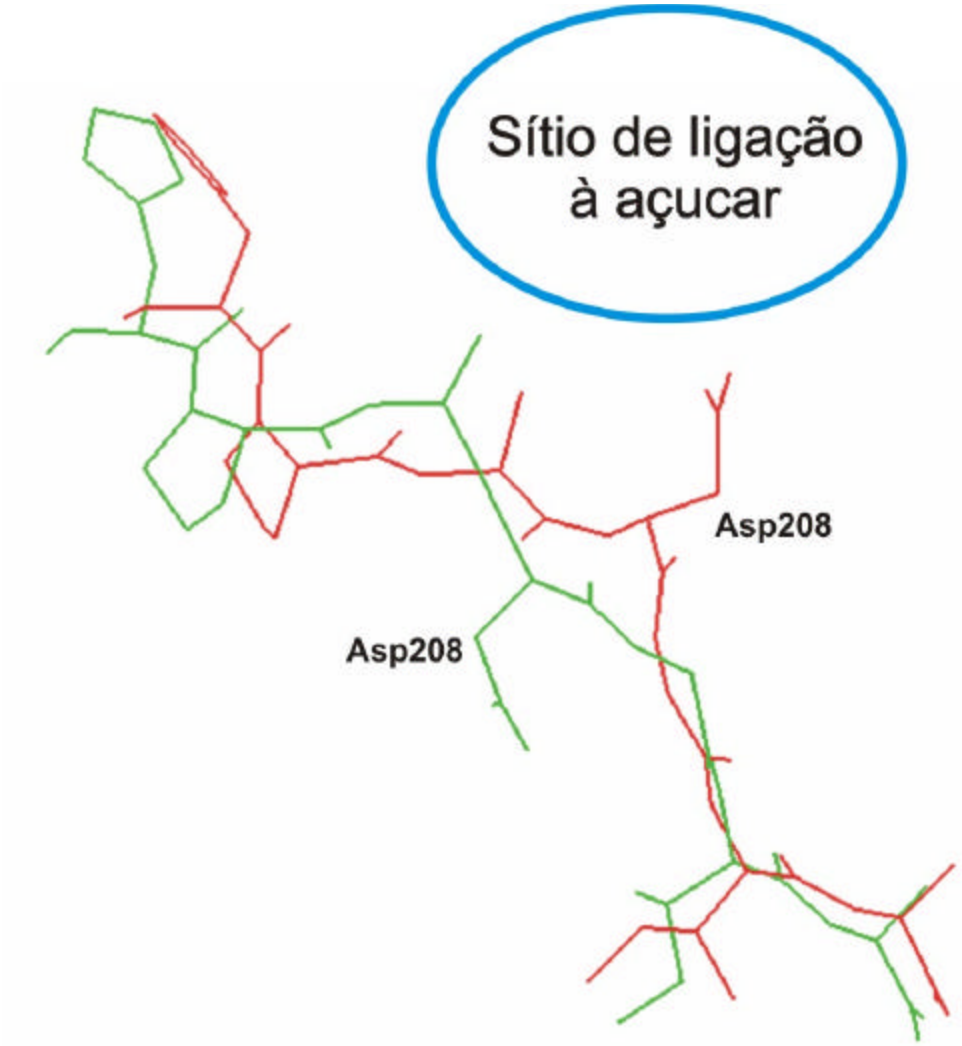


Figura 11. Superposição das estruturas resolvidas da lectina de *Canavalia ensiformis* (ConA) na presença e ausência dos íons cálcio e manganês. Em vermelho, conformação “locked”, com o aspártico 208 voltado para o sítio de ligação à açúcar. Em verde, estrutura resolvida na ausência dos íons cálcio e manganês. Conformação “unlocked”, com o aspártico 208 posicionado para o lado oposto ao sítio de ligação ao açúcar (BOUCKAERT *et al*., 1995).

Entretanto muitas lectinas de leguminosas possuem glicanos do tipo "N" em suas estruturas nativas, podendo ser glicosiladas em um ou mais de um ponto. O interessante é que as lectinas de leguminosas apresentam-se, estruturalmente, formando dímeros ou tetrâmeros que podem ser glicosilados ou não, existindo, com isto, a possibilidade de combinação de várias formas tridimensionais, como ilustrado na figura 12 (VAN DAMME *et al.*, 1998).

Muitas lectinas são expressas nas formas glicosiladas e não glicosiladas. É o caso da lectina de *Vatairea macrocarpa* que apresenta cadeias ativas com um único ponto de glicosilação e com dois pontos de glicosilação, além de fragmentos, também ativos com frações glicosilados e não glicosiladas (CALVETE *et al*., 1998). Algumas variações também são encontradas nas lectinas de sementes de *Phaseolus vulgaris*, *Griffonia simplicifolia* e *Vicia villosa* (VAN DAMME *et al*., 1998). Alguns pesquisadores citam que a glicosilação pode ser uma estratégia celular para que novas estruturas tridimensionais surjam devido a impedimento espacial causado pela presença do glicano (SHAANAN *et al*., 1991). Outros relatam que a glicosilação em lectinas pode estar envolvida em outras funções celulares, que como conseqüência levam ao enovelamento diferenciado (TURTON *et al*., 2004).

Investigações mais recentes afirmam que a glicosilação em lectinas de leguminosas não exerce nenhum tipo de efeito no arranjo oligomérico das mesmas. É o caso, por exemplo, da glicolectina presente em sementes de *Erythrina cristagalli,* que de acordo com sua estrutura nativa e recombinante, não apresentou nenhuma novidade quanto a sua forma de tetramerização, confirmando que a presença do *N*-glicano, ligado covalentemente na Asn-17, não influencia o modo de oligomerização da mesma (TURTON *et al*., 2004).

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

5.2.3.1.1 Classificação das lectinas de leguminosas quanto a afinidade por monossacarídeos

As lectinas de leguminosas possuem capacidade de interagir com açúcares simples e complexos (DAM & BREWER, 2002). Essa interação, conforme sucintamente comentado anteriormente, depende, exclusivamente,

de um sítio de ligação primário a carboidratos presente na superfície molecular das mesmas. Esse sítio, comumente chamado de primário, interage diretamente com um único açúcar simples, daí sua denominação.

Diversos grupos de lectinas foram descritos quanto à especificidade por monossacarídeos, entre eles as lectinas do tipo manose/glicose, *D*-galactose, *N*-acetylglicosamina, *N*-acetylgalactosamina, *L*-fucose, ácido *N*-acetyl-neuramínico, etc. Entretanto, a maioria das lectinas, que compartilham entre si a característica de reconhecerem primariamente o mesmo tipo de monossacarídeo, diferem entre si em relação a afinidades por oligossacarídeos. Em outras palavras, isso quer dizer que a interação com monossacarídeos é bem menos complexa do que por polissacarídeos.

O sítio de ligação a carboidratos das lectinas de leguminosas consiste em uma região na superfície da molécula composto por quatro loops que proporcionam a formação de uma cavidade na superfície da mesma. Nestes loops encontram-se três resíduos altamente conservados, Asp, Asn e Gly/Arg, os quais são responsáveis pela formação de quatro ligações de hidrogênio com as hidroxilas dos carbonos três e quatro presentes nos monossacarídeos como manose ou glicose. A estabilização dos monossacarídeos ocorre também através de interações hidrofóbicas com os resíduos de aminoácido como Phe, Tyr, Trp ou Leu. O interessante é que as variações evolutivas observadas nas lectinas de leguminosas ocorrem especificamente no sítio de ligação a carboidratos e na maneira como essas proteínas formam oligômeros. Em outras palavras, se existirem variações no sítio primário de ligação haverá provavelmente mudança na especificidade da lectinas pelo seu respectivo monossacarídeo. Por outro lado, caso essas

variações sejam observadas na forma de oligomerização haverá mudanças na forma em como essas proteínas interagem com oligossacarídeos complexos presentes em membranas celulares. Ou seja, lectinas contendo a mesma especificidade por monossacarídeos podem desempenhar atividades biológicas distintas e vice-versa.

Como exemplo podemos citar as lectinas ConA e PNA, que possuem idênticos sítios de ligação aos metais cálcio e manganês, porém apresentam cerca de 50% de diferenças em relação aos resíduos responsáveis pela estabilização reversível dos carboidratos. Tais considerações apontam para que a conservação do sítio de ligação a metais seja extremamente essencial para que haja uma correta estabilização do glicano, visto que esse sítio é sempre conservado. Entretanto, o sítio de ligação ao açúcar, propriamente dito, sofre pequenas variações de acordo com as espécies e estas, por sua vez, são responsáveis pelas diferenças nas atividades biológicas apresentas por diversas lectinas em experimentos biológicos (SANZ-APARICIO *et al*., 1997 e BANERJEE *et al*., 1996).

5.2.3.2 Estrutura quaternária e as variações dímero-dímero de lectinas de leguminosas

As lectinas de leguminosas são similares entre si quando comparadas em relação às suas estruturas primárias e terciárias, porém diferem extraordinariamente quanto à estrutura quaternária (BRINDA *et al*., 2005). Essas moléculas raramente ocorrem na forma monomérica, sendo mais comum encontrá-las nas formas dimérica e tetramérica, onde os tetrâmeros são formados a partir de arranjos espaciais entre dímeros. Cada interação dimérica distinta gera uma forma tetramérica específica.

Conforme comentado anteriormente, as variações observadas quanto as diferenças de atividade biológica apresentadas pelas lectinas de leguminosas podem estar relacionadas às mudanças nas formas oligoméricas em que se encontram. Cada estrutura quaternária de uma lectina de leguminosa interage de uma forma específica com um determinado tipo de oligossacarídeo, ou seja, muitos fenômenos biológicos são desencadeados a partir de interações cruzadas com receptores celulares glicosilados, o que nem sempre ocorre da mesma maneira para lectinas distintas.

As interações entre lectinas de leguminosas e receptores celulares glicosilados podem ocorrer de diversas formas. Basicamente três tipos de interações (ligações cruzadas) foram identificados:

i) cross-linking em uma dimensão, que ocorre normalmente em lectinas diméricas,

ii) cross-linking em duas dimensões, que pode ser visto em homolectinas tetraméricas e planares, e

iii) cross-linking em três dimensões que são geralmente formadas por lectinas tetraméricas não planares (Figura 13).

Embora existam poucos relatos a respeito de informações estruturais de "cross-linking" entre lectinas de leguminosas e receptores glicosilados, estudos a partir de Microscopia Eletrônica identificaram diferenças observadas nas ligações cruzadas formadas pelas lectinas de *Lotus tetragonolobus* (LTA) e *Canavalia ensiformis* (ConA) na presença de diversos oligossacarídeos (CHENG *et al*., 1998). Essas lectinas são capazes de formar ligações cruzadas em duas e três dimensões com glicanos divalentes, respectivamente, porém ainda não se sabe como esse fenômeno ocorre no meio vivo celular. Acredita-se que essas interações cruzadas assumem formas complexas, longas e periódicas, que podem ser até mesmo consideradas como estruturas super-quaternárias.

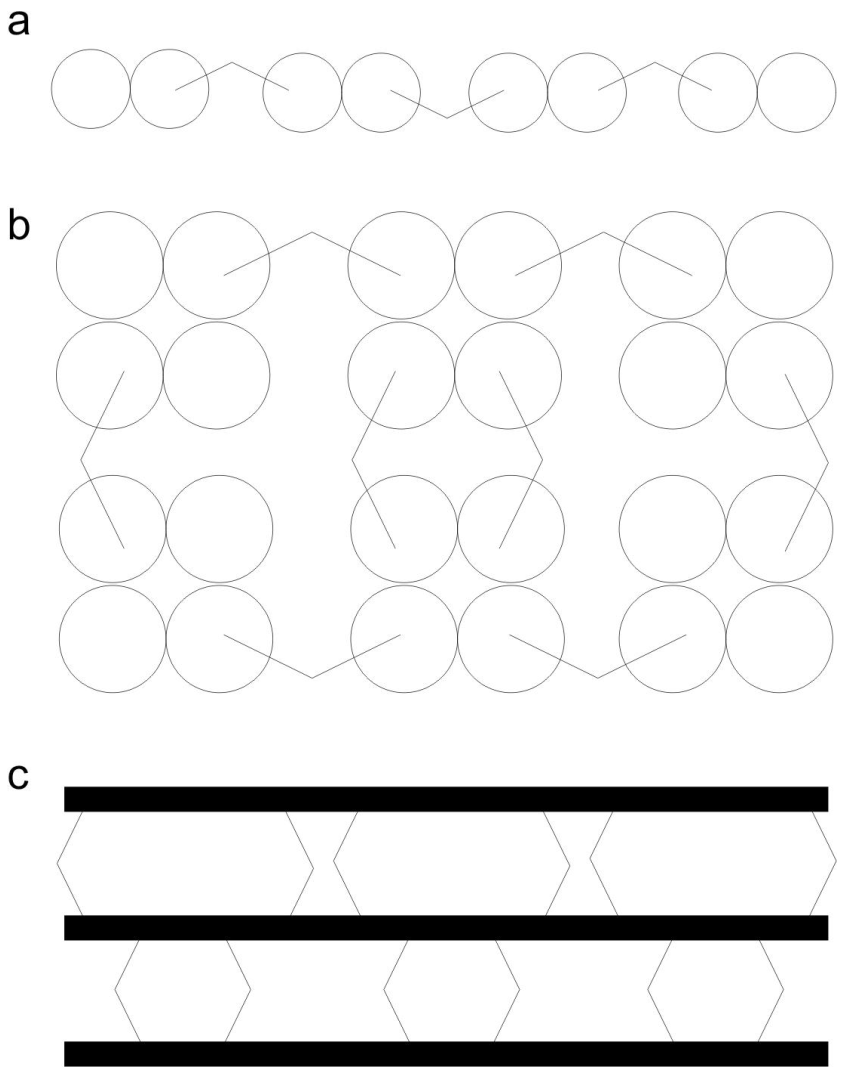


Figura 13. Representação esquemática de cross-linking entre lectinas nas três dimensões. a) Cross-linking em uma dimensão. Lectinas diméricas. Cada circunferência representa um monômero com um único sítio de ligação ligado a um carboidrato divalente. b) Lectinas tetraméricas com sítios de ligação dispostos no mesmo plano. Formação de cross-linking em duas dimensões. c) Lectinas tetraméricas com sítios de ligação dispostos em dois ou mais planos. Formação de cross-linking em três dimensões.

Atualmente o que se tem visto em relação ao uso de lectinas com glicoconjugados é que as interações tornam-se mais complexas na medida em que o tamanho do açúcar aumenta. Isso se deve ao fato de que a interação entre lectinas e glicoconjugados ocorre de uma forma não tão simples. Alguns carboidratos divalentes são capazes de se ligar a duas moléculas de lectinas, portanto se uma lectina assume uma forma oligomérica como um dímero ou tetrâmero haverá a formação de ligações cruzadas, que por sua vez surgem de acordo com o tamanho de cada receptor glicídico e a forma estrutural em que encontra a lectina (Figura 14).

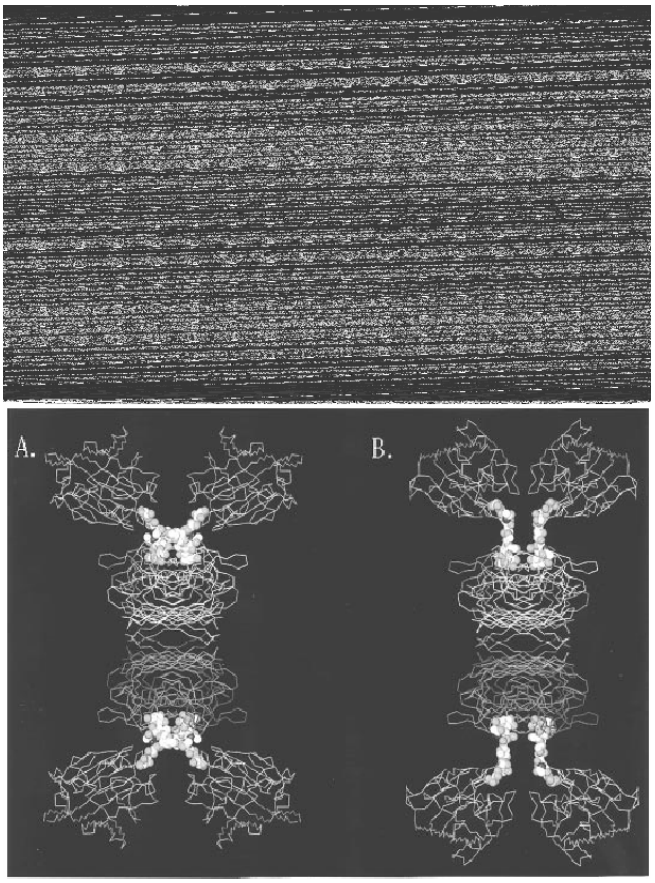


Figura 14. Representação em duas vistas do cross-linking formado entre sacarídeos divalentes e a lectina de soja (Laurence *et al*., 1997).

Por outro lado, conforme comentado anteriormente, as lectinas de leguminosas podem ocorrer nas formas monoméricas, diméricas e tetramêricas. Atualmente nove tipos de estruturas quaternárias foram relatadas para as lectinas de leguminosas, que compreendem o monômero identificado na espécie *Phaseolus vulgaris*, quatro estruturas diméricas e quatro estruturas tetraméricas. O interessante é que os monômeros presentes nessas estruturas diméricas são extremamente semelhantes ao monômero da lectina de *Phaseolus vulgaris* e os tetrâmeros são combinações dessas estruturas diméricas. Sugere-se, portanto, que as lectinas são moléculas que desempenham funções correlacionadas, produto de um processo evolutivo em

que mutações em suas estruturas primárias refletiram na maneira em como essas proteínas se arranjam espacialmente.

Estudos a partir de técnicas computacionais são utilizados na tentativa de identificar através da estrutura primária quais mutações estão envolvidas no processo de oligomerização destas proteínas, dado que são extremamente semelhantes. No entanto, tais esforços são apenas suficientes para se determinar o grau de oligomerização das mesmas, ou seja, pode-se inferir, a partir de algumas sequências primárias, que determinadas lectinas são diméricas ou tetraméricas, porém a forma em como ocorre o arranjo espacial desses dímeros ou tetrâmeros é visto apenas por técnicas experimentais (BRINDA *et al*., 2004).

As estruturas quaternárias observadas na figura 15 foram evidenciadas através da cristalografia de proteínas. Como anteriormente mencionado, todas estas são muito semelhantes em relação a suas estruturas primárias, porém apresentam formas de oligomerização distintas, o que como conseqüência reflete na atividade biológica desempenhada pelas mesmas. Entretanto, essas diferenças, produto de um processo evolutivo, além de estarem relacionadas com variações na estrutura primária conforme comentado anteriormente, podem também sofrer influência de acordo com o "padrão de glicosilação" apresentado pelas mesmas.

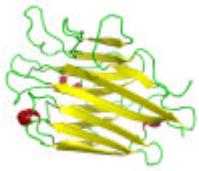
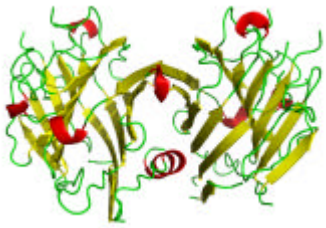
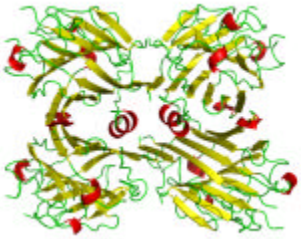
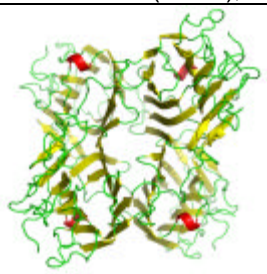
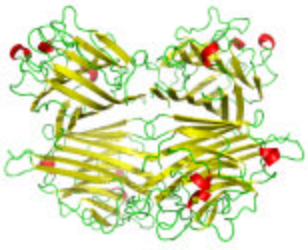
Em relação à estrutura primária, alguns resíduos de aminoácidos, que se posicionam em regiões de interface entre as cadeias, podem sofrer mutações, ocasionando leves distorções espaciais entre dois dímeros. No entanto, para o caso da presença de glicanos, sabe-se que estes são introduzidos covalentemente em resíduos de Serina/Treonina (*O*-glicanos) ou

de Asparagina (*N*-glicanos) através de mudanças pós-traducionais. Normalmente, o que se tem observado é que a maioria das lectinas de leguminosas possui um padrão específico de glicosilação, contudo, a maneira como ocorrem essas mudanças pode também influenciar na forma em como os monômeros destas proteínas se interagem para formar estruturas quaternárias distintas.

Glicanos do tipo "*N*" são comumente encontrados nas superfícies das lectinas de leguminosas. São moléculas estruturadas a partir de ligações glicosídicas de monosacarídeos como GlicNac, Fucose, Manose, Xylose, etc. O padrão de glicosilação é dependende da posição das asparaginas, porém a composição destes glicanos ainda é considerado um mistério, pois diversas lectinas, mesmo muito semelhantes, podem apresentar glicanos que possuem desde 200Da até 1500Da. O interessante é que tetrâmeros distintos podem ocorrer de acordo com o padrão de glicosilação das lectinas, como por exemplo, a EcorL, que apresenta um tetrâmero distinto dos demais observados devido a presença de um heptasacarideo do tipo N. (SHAANAN *et al*., 1991).

Conclui-se, portanto que, o modo de tetramerização das lectinas de leguminosas, está intimamente relacionado a um processo evolutivo referente a mutações pontuais nas estruturas primárias das mesmas. No entanto não se pode descartar que diferenças nos padrões de glicosilação destas lectinas contribuem também para tal processo. Pesquisadores afirmam que o padrão de glicosilação pode, realmente, influenciar estereoquimicamente na maneira em como essas proteínas oligomerizam, porém caso isto seja uma afirmação verdadeira, torna-se difícil compreender se a célula utilizou sua capacidade de glicosilar para realmente formar novas estruturas quaternárias, ou se estas

estruturas quaternárias surgem por consequência da presença destes glicanos que foram incorporados para realizarem uma outra função.

	
a) Monômero da lectina de sementes de *Phaseolus vulgaris* (Arcelina – 5);	b) Dímero canônico tipo-2 (lectinas da sub tribo Diocleinae);
	
c) Dímero da lectina de sementes de *Erythrina coralodendron* (EcorL);	d) Dímero da lectina de sementes de *Griffonia simplicifolia* (GS4-dímero);
	
e) Dímero da lectina presente em sementes de *Dolichus biflorus* (DB58);	f) Tetrâmero da Lectina de *Dolichus biflorus* lectin (DB58);
	
g) Tetrâmero da lectina de *Canavalia ensiformis* (ConA);	h) Tetrâmero da lectina presente na espécie *Arachis hypogaea* (Peanut);

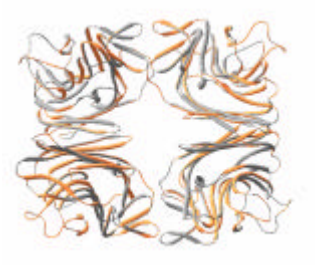
 **i) Tetrâmero da lectina de *Griffonia simplicifolia* (GS1);**	

Figura 15. Exemplificação dos nove tipos de estrutura quaternária encontradas em lectinas de leguminosas.

5.3 Glicobiologia

A sequência de genes e proteínas depositadas cresce de uma forma significativamente superior a de estruturas de carboidratos. Isso não ocorre devido ao fato de haver mais empenho na elucidação daquelas do que destas, mas devido ao fato de que a determinação de estruturas de açúcares complexos tem apresentado muitos desafios. Uma das dificuldades está no fato destes açucares formarem complexos com outras moléculas e possuírem variações estruturais resultantes de suas vias biossintéticas. Essa carência reflete exclusivamente no fato de que a falta de informações estruturais/biológicas de estruturas de carboidratos é uma pedra no caminho para quando se quer conseguir entender fenômenos biológicos como interação proteína-carboidratos, ou até mesmo proteínas-glicoproteínas e proteínas-glicolipídeos (HASHIMOTO & KEGG, 2006).

Os carboidratos são moléculas com alto potencial para formar estruturas complexas, pois um único dissacarídeo pode formar bem mais estruturas distintas do que um dipeptídeo. Isso ocorre devido à diversidade e graus de liberdade das ligações glicosídicas. Por exemplo, enquanto dois aminoácidos podem formar um dipeptídeo a partir de uma única possível ligação peptídica, duas hexoses podem se conectar através de 04 dos seus carbonos hidroxilados. Essa vasta combinação é um dos fatores em que explica o porquê que as interações moleculares baseadas no reconhecimento entre estruturas complexas e carboidratos é uma das maiores formas de comunicação intra- e inter-celular concorrendo para a propagação de diversos processo biológicos relevantes.

Diversos processos bioquímicos que ocorrem nas comunicações celulares mediadas por carboidratos ainda hoje permanecem sem explicações convincentes. No entanto, esforços no sentido de elucidar tais fenômenos surgiram com o uso de lectinas vegetais, que são proteínas que possuem a capacidade de se ligar reversivelmente a carboidratos, ou seja, decifram glicocódigos presentes em diversos tipos de estruturas celulares. A investigação dessas proteínas e seu papel no reconhecimento de células, assim como a aplicação dessas proteínas para o estudo de carboidratos em solução e na superfície de células, está proporcionando contribuições importantes para o avanço da glicobiologia. O desenvolvimento desse último campo tem um impacto profundo na pesquisa de lectinas, de tal forma que essas duas áreas avançam lado a lado (DWEK, 1996).

6.0 CAPÍTULO 1

6.1 A Lectina de sementes de *Lotus tetragonolobus*

A espécie *Lotus tetragonolobus* pertence à Família das Leguminosae-Papilionoideae e tribo Loteae. O gênero Lotus possui mais de 100 tipos plantas catalogadas, dentre as quais são bastante estudadas as espécies *Lotus corniculatos* e *Lotus japonicus*. Esta última já possui seis cromossomos sequenciados e disponíveis na base de dados “http://www.kazusa.or.jp/lotus/”. A espécie *Lotus. tetragonolobus* ou *Tetragonolobus purpurea* é considerada por ser uma planta medicinal. É comumente encontrada no continente europeu e também pode ser vulgarmente chamada por “*Asparagus Peã”* ou “*Winged Pea”* (Figura 16).

Figura 16. Fotografia da espécie *Lotus tetragonolobus* (Papilionoideae; Loteae; Lotus). "http://hortiplex.gardenweb.com/plants/p1/gw1024138.html"

tetragonolobus remota do ano de 1963: "*Research on phytoagglutinin from the seeds of L. and some other plants*". Trata-se de um trabalho publicado na Polônia por Sagan e colaboradores (SAGAN *et al*., 1963). Posteriormente, Yariv e colaboradores, (1967), descreveram, pela primeira vez, através do artigo "*Isolation of an L-fucose binding protein from Lotus tetragonolobus seed*", informações sobre o isolamento de aglutininas presente nas sementes da mesma (YARIV *et al*., 1967) e que a espécie *L. tetragonolobus* expressava uma proteína que tinha capacidade de aglutinar eritrócitos. Entretanto, um ano mais tarde, em 1968, foi descoberto que existiam pelo menos três isoformas de aglutinina presentes nas suas sementes (KALB, 1968).

O uso da lectina de *Lotus tetragonolobus* (LTA), como ferramenta bioquímica, se deu na década de 70. Estudos da sua especificidade sanguínea no sistema ABO humano (PEREIRA e KABAT, 1974), de caracterização de antígenos da saliva humana através da precipitação (NAPIER *et al*., 1974) e no isolamento de leucócitos polimorfonucleares humanos (VAN & TUNG, 1977) foram algumas das aplicações biológicas que proporcionaram um superficial entendimento de que a LTA era específica para alguns tipos de receptores presentes em determinadas membranas celulares. Na realidade já se tinha conhecimento de que a especificidade da LTA ocorria por *L*-fucose, mas ainda não estava elucidado que essa interação se dava de uma maneira bem mais complexa, ou seja, através de oligossacarídeos constituídos por *L*-fucose.

Posteriormente, nos anos 80, inúmeras aplicações biológicas foram desenvolvidas com a LTA. Sua capacidade de reconhecimento a carboidratos

ricos em *L*-fucose era uma característica que a distinguia de outras lectinas, estudadas na época, como a Concanavalina A (específica por manose/glicose). Dentre os principais trabalhos a respeito da utilização da LTA, com base em sua especificidade por oligossacarídeos *L*-fucosídicos, podemos destacar os estudos a partir de células malignas com o propósito de se identificar antígenos anômalos presentes nas membranas de células cancerígenas (RAEDLER *et al*., 1983) (KOCH *et al*., 1983) (HOLTHOFER *et al*., 1983) (WALKER, 1984) (ULRICH *et al*., 1985) (DAVIDSSON *et al*., 1987) (GRIFFITHS & STEPHERSON, 1988) (RAJU & LEE, 1988) (FINNE *et al*., 1989) (IMURA *et al*., 2004) (TURNER *et al*., 1995). Essas investigações foram direcionadas devido ao fato de que oligossacarídeos fucosilados são amplamente encontrados em células durante o processo de diferenciação ou estágios de metástase, comuns em diversos tipos de carcinomas, como neuroblastomas ou adenocarcinomas (THOMAS & SUROLIA, 2000).

No entanto, o entendimento da LTA como apenas uma lectina que se liga a fucose é insuficiente para que se possa explicar os inúmeros fenômenos biológicos que a mesma desempenha. O reconhecimento da LTA por *L*-fucose já está bem elucidado apenas em termos de testes simples por inibição, porém a maneira que como ocorrem estas interações com oligossacarídeos complexos ainda dá os primeiros passos. Alguns estudos com Microscopia Eletrônica e Ressonância Plasmônica de Superfície mostraram que a LTA têm capacidade de reconhecer antígenos fucosilados do tipo Lewis^{x} e oligossacarídeos divalentes como difucosyllacto-N-neohexa-ose – FucOcta, porém nenhuma informação estrutural a respeito dessas interações têm sido relatadas (STAUDACHE *et al*., 1999) (CHENG *et al*., 1998).

Dentre todos os trabalhos produzidos até então, o que mais explora convincentemente a forma de reconhecimento de oligossacarídeos complexos pela LTA foi relatado por Cheng e colaboradores (CHENG *et al*., 1998). A LTA foi utilizada para precipitar Fuc-Octa (difucosyllacto-N-neohexaose – FucOcta) (Figura 17). Esse precipitado, quando analisado por Difração de Raios X e Microscopia Eletrônica revelou que havia um padrão de simetria entre o açúcar e a lectina, ou seja, durante o processo de aglutinação ocorreu a formação de ligações cruzadas a uma razão de 2:1 (duas moléculas do açúcar divalente para uma de LTA). Essas ligações, de natureza reversível, ocorrem entre o açúcar e a estrutura tetramérica da LTA, que possui 04 sítios de ligação à carboidratos homólogos, em um *cross-linking* do tipo-2 (bidimensional) (CHENG *et al*., 1998).

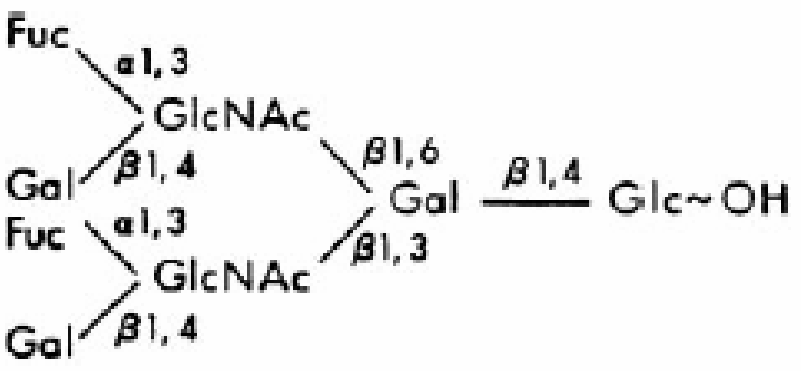


Figura 17. Representação molecular do difucosyllacto-N-neohexaose – FucOcta. Oligossacarídeo biantenário e divalente. Interage com a LTA na razão de 2:1 (CHENG *et al*., 1998).

Ultimamente, a LTA continua sendo uma molécula extremamente utilizada para diversas aplicações biológicas (MANSOUR *et al*., 2005) (MONTISCI *et al*., 2001), mas até então não se sabe como a LTA encontra-se na sua forma tetramérica, ou seja, como seus dímeros de dímeros estão arranjados

espacialmente e como o seu sítio de ligação a *L*-fucose está estruturado. Essas informações são cruciais para se entender como se dá o processo de reconhecimento de oligossacarídeos pela LTA, característica fundamental para que se entenda o mecanismo de todas as atividades biológicas desempenhadas pela mesma. Além do que, outras informações estruturais mais complexas ainda não são conhecidas como por exemplo a maneira como ocorre o fenômeno de *cross-linking* em duas dimensões entre a LTA e oligossarídeos biantenários apresentado anteriormente por Cheng e colaboradores (CHENG *et al*., 1998).'

7.0 OBJETIVOS

Objetivos gerais

Esta tese tem como objetivo principal o estudo estrutural de uma lectina presente em sementes da espécie *Lotus tetragonolobus*, específica a *L*-fucose, de modo a aumentar o entendimento da relação estrutura-função biológica que a mesma desempenha.

Objetivos específicos

a) Cristalizar na forma nativa a LTA através da técnica de difusão de vapor;

b) Coletar dados de difração de raios X de monocristais da LTA;

c) Resolver a estrutura tridimensional por substituição molecular;

d) Fazer análises estruturais da estrutura da lectina refinada e correlacionar com atividades biológicas conhecidas.

8.0 RESULTADOS E DISCUSSÃO

Artigo 01

Moreno FB, Martil DE, Cavada BS, de Azevedo WF Jr. Crystallization and preliminary X-ray diffraction analysis of an anti-H(O) lectin from *Lotus tetragonolobus* seeds. Acta Crystallogr Sect F Struct Biol Cryst Communnications. 2006 Jul 1;62(Pt 7):680-3. Epub 2006 Jun 26.

Artigo 02

Moreno FB, de Oliveira TM, Martil DE, Viçoti MM, Bezerra GA, Abrego JR, Cavada BS, Filgueira de Azevedo W Jr. Identification of a new quaternary association for legume lectins. ***Journal of Structural Biology***. 2008 Feb;161(2):133-43. Epub 2007 Oct 15.

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

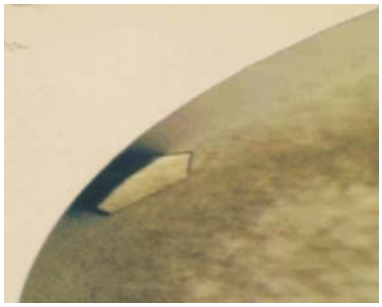
Crystallization and preliminary X-ray diffraction analysis of an anti-H(O) lectin from *Lotus tetragonolobus* seeds

Frederico Bruno Mendes Batista Moreno,[a] Daiana Evelin Martil,[a] Benildo Sousa Cavada[b] and Walter Filgueira de Azevedo Jr[c]*

[a]Programa de Pós-graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, [b]BioMol-Lab UFC, Caixa Postal 6043, 60.455-970, Fortaleza-CE, Brazil, and [c]Faculdade de Biociências–PUCRS, Av. Ipiranga 6681, Porto Alegre-RS, CEP 90619-900, Brazil

Correspondence e-mail: walter.junior@pucrs.br

Received 20 April 2006
Accepted 5 June 2006



The seed lectin from *Lotus tetragonolobus* (LTA) has been crystallized. The best crystals grew over several days and were obtained using the vapour-diffusion method at a constant temperature of 293 K. A complete structural data set was collected at 2.00 Å resolution using a synchrotron-radiation source. LTA crystals were found to be monoclinic, belonging to space group $P2_1$, with unit-cell parameters a = 68.89, b = 65.83, c = 102.53 Å, $\alpha \neq \gamma \neq 90$, $\beta \neq 92^\circ$. Molecular replacement yielded a solution with a correlation coefficient and R factor of 34.4 and 51.6%, respectively. Preliminary analysis of the molecular-replacement solution indicates a new quaternary association in the LTA structure. Crystallographic refinement is under way.

1. Introduction

The legume lectins are a large group of homologous carbohydrate-binding proteins of non-immune origin that are found mainly in the seeds of most leguminous plants. These lectins act by deciphering specific glycocodes encoded in the structure of glycans. The interaction between lectins and carbohydrates plays a biological role in cellular processes such as cell communication, host defence, fertilization, development, parasitic infection, tumour metastasis and plant defence against herbivores and pathogens (Vijayan & Chandra, 1999; Gallego del Sol *et al.*, 2005). Lectins are ubiquitous in animals, plants and microorganisms. Over 250 three-dimensional structures of lectins from diverse sources are available and legume lectins represent a significant proportion of these proteins. Lectins share a common structural fold but they differ in their carbohydrate specificities. An outstanding feature of the group of legume lectins is that although all the monomers have similar tertiary structures, they show different modes of quaternary association (Brinda *et al.*, 2004). This structural feature deserves more in-depth investigation in order to identify possible structural aspects that may dictate the determinants of the quaternary association of legume lectins.

Bianchet and coworkers characterized fucose-binding lectins as proteins that bind fucose and share a specific sequence motif whose function is associated with immune recognition in vertebrates and invertebrates (Bianchet *et al.*, 2002). Fucose-binding lectins are widespread among microorganisms, animals and plants, including *Pseudomonas aeruginosa* lectin (PA-IIL), *Anguilla anguilla* agglutinin (AAA), *Morone saxatilis* agglutinin (MsaFBP32), *Dicentrarchus labrax* agglutinin, *Chromobacterium violaceum* lectin (CV-IIL), *Ralstonia solanacearum* lectin (RS-IIL), *Ulex europaeus* agglutinin (UEA-I) and *Lotus tetragonolobus* agglutinin (LTA) (Bianchet *et al.*, 2002; Mitchell *et al.*, 2005; Vandonselaar & Delbaere, 1994; Konami *et al.*, 1990; Zinger-Yosovich *et al.*, 2006; Odom & Vasta, 2006; Cammarata *et al.*, 2001).

Fucose-binding lectins have been widely studied because the molecular-recognition properties of glycoproteins and glycolipids containing L-fucose are often characterized in terms of their interactions with fucolectins. Moreover, there have been several reports that L-fucosyl oligosaccharides are found in most common human cancers, particularly in adenocarcinomas and neuroblastomas. Several pieces of evidence suggest that fucosyl oligosaccharides are important cell-surface recognition determinants (Hakomori, 1984; Santer *et al.*, 1983; Cheng *et al.*, 1998).

LTA and UEA-I are two homologous fucolectins from the Leguminosae family. Although they possess the same primary specificity for fucose, they display considerable diversity in their carbohydrate binding to fucosylated oligosaccharides. UEA-I recognizes the H-type 2 determinant [α-L-Fuc(1→2)β-D-Gal(1→4)β-D-GlcNAc] but not the Le^a and Le^x determinants. However, LTA recognizes Le^x and other different divalent L-fucosyl carbohydrates (Cheng *et al.*, 1998). Despite their similarity and apparently conserved sequences (41% amino-acid identity), the above lectins possess different biological activities. Furthermore, a previous study of LTA and other legume lectins with unknown structures have shown that the LTA structure should have a quaternary association type that differs from those of known legume lectins (Brinda *et al.*, 2004). Thus, structural studies of LTA will help us to understand its binding to carbohydrates, which is directly responsible for its biological activity, and to elucidate its quaternary association.

LTA is a member of the legume family (Leguminosae, Papilionoideae, Loteae) of lectins, which have been widely used to explore the properties of membranes from both normal and transformed cells (Shirahama *et al.*, 1993; Mansour *et al.*, 2005). The primary structure of LTA has been previously determined and is a glycoprotein containing 240 amino-acid residues with a molecular weight of 26 273.17 Da. Furthermore, the affinity of LTA for L-fucosyl oligosaccharides has been investigated by nuclear magnetic resonance and electron microscopy (Cheng *et al.*, 1998; Haselhorst *et al.*, 2001). In order to establish the crystal structure of this new member of the fucose-binding lectins, we carried out a crystallization study and preliminary X-ray diffraction analysis of LTA, a fucolectin from *Lotus tetragonolobus* seeds, with the aim of solving its native quaternary structure (Pereira & Kabat, 1974*a*,*b*; Konami *et al.*, 1990).

2. Material and methods

2.1. LTA crystallization, data collection and processing

LTA was purchased from Sigma–Aldrich (USA). The purified lectin was dissolved completely at a concentration of 5.0 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.6 containing 1 m*M* $CaCl_2$ and $MnCl_2$. For hanging-drop crystallization trials, the sample was submitted to various crystallization conditions using Crystal Screens I and II and SaltRx, Index and PEG/Ion screens (Hampton Research, Riverside, CA, USA). The drops were composed of equal volumes (1 µl) of protein solution and reservoir solution and were equilibrated against 500 µl reservoir solution. Crystals were grown in Linbro plates at 293 K by the vapour-diffusion method (Jancarik & Kim, 1991).

X-ray diffraction data were collected at a wavelength of 1.47 Å using a synchrotron-radiation source (MX1 station, Laboratório Nacional de Luz Síncrotron, Campinas, Brazil) and a CCD detector (MAR Research). Data were collected from the LTA crystal at 100 K and to avoid freezing crystals were soaked in a cryoprotectant solution containing 75% mother liquor and 25% glycerol. Using an oscillation range of 1.5° and an exposure time of 90 s per frame, 150 images were collected to a maximum resolution of 2.0 Å. Data were processed, indexed and integrated using *MOSFLM* and scaled using *SCALA* (Collaborative Computational Project, Number 4, 1994; Leslie, 1992).

2.2. Molecular replacement

The molecular-replacement method was used to determine the crystal structure of LTA using *MOLREP* (Collaborative Computational Project, Number 4, 1994). Rotation and translation functions were obtained using the peanut lectin monomer coordinates (PDB code 1cr7; Ravishankar *et al.*, 2001) as a search model. Three space groups were tested (*P*2, $P2_1$ and *C*2) and the best solution for each model was chosen based on the magnitude of the correlation coefficient and the *R* factor.

3. Results and discussion

3.1. Optimization of LTA crystals

LTA has previously been purified and its primary sequence determined. The lectin was shown to be a glycolectin and its haemagglutinating activity was found to be inhibited by the presence of L-fucose (Konami *et al.*, 1990).

Small irregular crystals of LTA appeared in 0.1 *M* trisodium citrate pH 5.6 containing 20% 2-propanol and 20% PEG 4000 after a week (Hampton Research Crystal Screen II, condition No. 40). These crystals were not suitable for X-ray diffraction experiments. However, several optimization steps were performed, changing the pH and precipitant concentration. The LTA crystals grew over several days in 0.1 *M* trisodium citrate pH 5.6 containing 8% 2-propanol and 16% PEG 4000 to maximum dimensions of 0.1 × 0.2 × 0.4 mm (Fig. 1).

3.2. Data collection and processing

LTA crystals diffracted to a maximum resolution of 2.0 Å using a synchrotron-radiation source (LNLS, Campinas, Brazil). The complete data set (150 frames) was indexed, integrated and scaled in the resolution range 40.42–2.0 Å. LTA crystals were monoclinic,

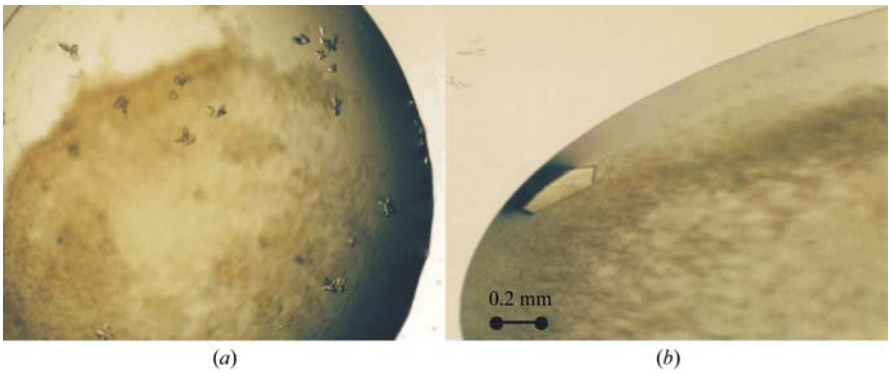


Figure 1
(*a*) Crystals from 0.1 *M* trisodium citrate pH 5.6 containing 20% 2-propanol and 20% PEG 4000. (*b*) Optimized LTA crystal suitable for X-ray diffraction experiments.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell (2.11–2.00 Å).

Wavelength (Å)	1.47
Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 68.89, b = 65.83, c = 102.53, $\alpha \neq \gamma \neq 90$, $\beta \neq 92$
Resolution range (Å)	40.42–2.00 (2.11–2.00)
Unique reflections	61847 (8835)
Completeness (%)	99.4 (97.9)
$\langle I/\sigma(I)\rangle$	10.8 (2.6)
R_{sym} (%)	9.5 (45.9)
Asymmetric unit content	One tetramer (960 residues)

belonging to the $P2_1$ space group, with unit-cell parameters a = 68.89, b = 65.83, c = 102.53 Å, $\alpha = \gamma \neq 90$, $\beta = 92°$. The volume of the unit cell was 464 772.31 Å^3, with a V_M value of 2.2 Å^3 Da^{-1} (Matthews, 1968). THe LTA crystal contained one tetramer per asymmetric unit. A summary of data-collection statistics is given in Table 1.

3.3. Molecular replacement

Sequence-alignment analysis was performed using the *BLASTP* software at the National Center of Biotechnology Information to compare the LTA primary sequence with those deposited in the PDB (Altschul *et al.*, 1990).

Molecular replacement was performed using the *MOLREP* software (Vagin & Teplyakov, 1997). The best result was obtained using the peanut lectin monomer. We found four monomers per asymmetric unit in the LTA crystal structure. The rotation and translation functions yielded a planar homotetramer with a correlation coefficient and R factor of 34.4 and 51.6%, respectively (Fig. 2). The model produced was then submitted to initial crystallographic refinement using *REFMAC5* (Collaborative Computational Project, Number 4, 1994). The first initial step of refinement was performed using rigid-body followed by restrained refinement (maximum-likelihood method), resulting in a model with an R factor of 29.4% and an R_{free} of 35.4%. Complete refinement of the LTA structure is in progress.

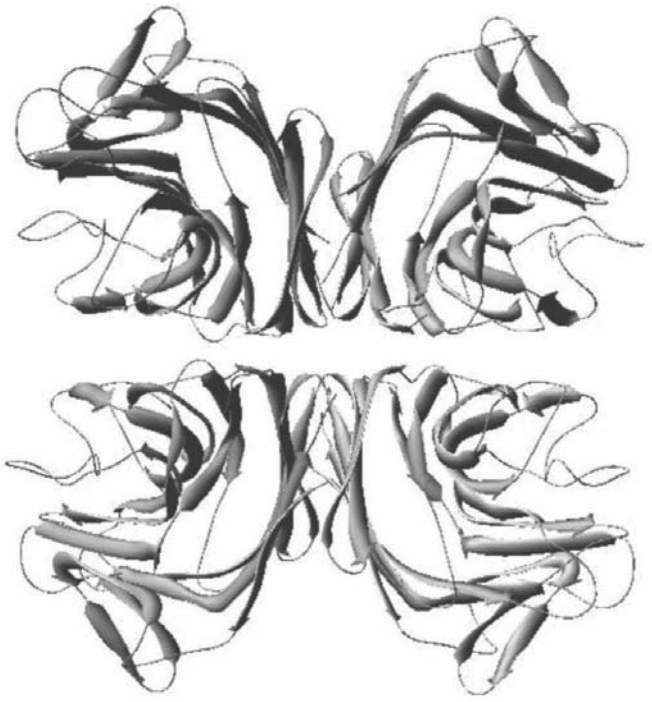


Figure 2
Molecular-replacement homotetramer solution for the LTA crystal using the peanut lectin monomer as a model.

The crystallographic packing of the LTA structure is shown in Fig. 2. Preliminary comparison of the LTA tetramer structure with the quaternary association of other legume lectins indicates that is has a different quaternary association. As described previously, the peanut lectin consists of a homotetramer composed of two different dimers. Its tetramer is made up of a canonical dimer and an unusual dimer interface, leading to an open quaternary structure. The canonical dimer is found in several legume lectins, as described for ConA, which is a homotetramer composed of two canonical dimers linked *via* a specific dimer–dimer interface. Normally, legume lectin tetramers are essentially dimers of dimers and the different modes of tetramerization are a consequence of the diverse combination of dimeric interfaces seen in these quaternary structures. Our present work shows that the preliminary initial model of LTA is a tetramer composed of two back-to-back interconnected dimers (non-canonical), yielding a planar homotetramer. Although GS4, a lectin from *Griffonia simplicifolia*, possesses the same type of dimer–dimer interface as LTA, its tetramers are significantly different. The molecular-replacement solution leads to acceptable crystal packing, but further refinement results will clarify the particular structural properties of LTA.

This work was supported by grants from FAPESP (SMOLBNet 01/07532-0, 02/04383-7, 04/00217-0), CNPq, CAPES and Instituto do Milênio (CNPq-MCT). WFA and BSC are researchers of the Brazilian Council for Scientific and Technological Development. We thank Dr A. Leyva for English language editing of the manuscript.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Bianchet, M. A., Odom, E. W., Vasta, G. R. & Amzel, L. M. (2002). *Nature Struct. Biol.* **9**, 628–634.
Brinda, K. V., Mitra, N., Surolia, A. & Vishveshwara, S. (2004). *Protein Sci.* **13**, 1735–1749.
Cammarata, M., Vazzana, M., Chinnici, C. & Parrinello, N. (2001). *Biochim. Biophys. Acta*, **1528**, 196–202.
Cheng, W., Bullitt, E., Bhattacharyya, L., Brewer, C. F. & Makowski, L. (1998). *J. Biol. Chem.* **273**, 35016–35022.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Gallego del Sol, F., Nagano, C., Cavada, B. S. & Calvete, J. J. (2005). *J. Mol. Biol.* **353**, 574–583.
Haselhorst, T., Weimar, T. & Peters, T. (2001). *J. Am. Chem. Soc.* **123**, 10705–10714.
Hakomori, S. (1984). *Annu. Rev. Immunol.* **2**, 103–126.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Konami, Y., Yamamoto, K. & Osawa, T. (1990). *FEBS Lett.* **268**, 281–286.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Mansour, M. H., Abdul-Salam, F. & Al-Shemary, T. (2005). *Immunobiology*, **210**, 335–348.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mitchell, E. P., Sabin, C., Snajdrova, L., Pokorna, M., Perret, S., Gautier, C., Hofr, C., Gilboa-Garber, N., Koca, J., Wimmerova, M. & Imberty, A. (2005). *Proteins*, **58**, 735–746.
Odom, E. W. & Vasta, G. R. (2006). *J. Biol. Chem.* **281**, 1698–1713.
Pereira, M. E. & Kabat, E. A. (1974*a*). *Ann. NY Acad. Sci.* **234**, 301–305.
Pereira, M. E. & Kabat, E. A. (1974*b*). *Biochemistry*, **13**, 3184–3192.
Ravishankar, R., Thomas, C. J., Suguna, K., Surolia, A. & Vijayan, M. (2001). *Proteins*, **43**, 260–270.
Santer, U. V., Glick, M. C., van Halbeek, H. & Vliegenthart, J. F. (1983). *Carbohydr Res.* **120**, 197–213.

Shirahama, T., Ikoma, M., Muramatsu, T., Kayajima, T., Ohi, Y., Tsushima, T., Akebi, N., Ohmori, H., Hirao, Y. & Okajima, E. (1993). *Cancer*, **15**, 1329–1334.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vandonselaar, M. & Delbaere, L. T. (1994). *J. Mol. Biol.* **243**, 345–346.
Vijayan, M. & Chandra, N. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.
Zinger-Yosovich, K., Sudakevitz, D., Imberty, A., Garber, N. C. & Gilboa-Garber, N. (2006). *Microbiology*, **152**, 457–463.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 161 (2008) 133–143
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

Identification of a new quaternary association for legume lectins

Frederico Bruno Mendes Batista Moreno [a], Taianá Maia de Oliveira [b], Daiana Evelin Martil [a], Magno Massolino Viçoti [a], Gustavo Arruda Bezerra [b], José Ramon Beltran Abrego [a], Benildo Sousa Cavada [b,*], Walter Filgueira de Azevedo Jr. [c,*]

[a] *Curso de Pós-graduação em Biofísica Molecular, Departamento de Física, Universidade Estadual Paulista Júlio de Mesquita Filho, R. Cristovão Colombo, 2265, Nazareth, São José do Rio Preto, São Paulo, Brazil*
[b] *Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Biomol-LAB, Campus do Pici S/N, Fortaleza, Ceará, Brazil*
[c] *Faculdade de Biociências-PUCRS, Av. Ipiranga, 6681, Porto alegre, Rio Grande do Sul, Brazil*

Received 27 June 2007; received in revised form 27 September 2007; accepted 4 October 2007
Available online 15 October 2007

Abstract

Lotus tetragonolobus lectin (LTA) is a fucose-specific legume lectin. Although several studies report a diverse combination of biological activities for LTA, little is known about the mechanisms involved in L-fucosyl oligosaccharide recognition. The crystal structure of LTA at 2.0 Å resolution reveals a different legume lectin tetramer. Its structure consists of a homotetramer composed of two back-to-back GS4-like dimers arranged in a new mode, resulting in a novel tetramer. The LTA *N*-linked carbohydrate at Asn4 and the unusual LTA dimer–dimer interaction are related to its particular mode of tetramerization. In addition, we used small angle X-ray scattering to investigate the quaternary structure of LTA in solution and to compare it to the crystalline structure. Although the crystal structure of LTA has revealed a conserved metal-binding site, its L-fucose-binding site presents some punctual differences. Our investigation of the new tetramer of LTA and its fucose-binding site is essential for further studies related to cross-linking between LTA and complex divalent L-fucosyl carbohydrates.

Keywords: *Lotus tetragonolobus* lectin (LTA); Crystal structure; L-Fucosyl

1. Introduction

Plant lectins are a group of proteins of non-immune origin that recognize and bind to carbohydrates without modifying them. It comprises a heterogeneous class of (glyco)proteins that play a range of crucial roles in many cell–cell recognition events such as host defense, fertilization, development, parasitic infection, tumor metastasis and plant defense against parasites. Although they encompass different members that are similar in their primary and tertiary structures, several differences have been reported with regard to their mode of quaternary association (Brinda et al., 2005). In general, legume lectins are essentially dimers of dimers, and the different modes of tetramerization are a consequence of a diverse combination of dimeric interfaces seen in these quaternary structures (Moreno et al., 2006; Delatorre et al., 2007). Additionally, the mode of tetramerization of legume lectins is closely related to their affinity for complex oligosaccharides. Differences in the carbohydrate cross-linking properties of some lectins in cellular recognition and signal transduction processes are reflected in biological properties (Dam et al., 2005). Thus, a detailed understanding of biological interaction mechanisms between lectins and complex carbohydrates is a key feature for establishing the chemical basis of important logical recognition functions performed by plant lectins.

Legume lectins are known to exist mainly as homodimers and homotetramers. They show considerable differences in their quaternary associations and modes of

* Corresponding authors. Fax: +55 51 3320 3629.
E-mail addresses: bscavada@ufc.br (B.S. Cavada), walter.junior@pucrs.br (W. Filgueira de Azevedo).

1047-8477/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2007.10.002

monomer organizations in the dimeric/tetrameric assemblage. Most structural studies of Leguminosae lectins have involved members of Papilionoideae subfamily, where the majority is composed of three sets of β-sheets interconnected by loops, commonly designated as "jelly roll" or "β-sandwich motif". In general, the term legume lectin refers exclusively to lectins from the Leguminosae family, but some of them are not considered to belong to the legume lectin class. For example, some chitinases and the ricin-like lectins (*Riccinus communis* lectin and *Abrus precatorius* lectin) are classified as chitin-binding and type-2 RIP lectins, respectively (Van Damme et al., 1998) .

On the basis of their overall quaternary structure, legume lectins from the Papilionoideae subfamily have been classified into nine types, which are composed of seven different dimeric interfaces. The seven different dimeric types include the canonical dimer type-II (Concanavalin A, from *Canavalia ensiformis*) and its non-canonical interface, the *Dolichos biflorus* seed lectin, EcorL (*Erythrina corallodendron* lectin), GS-I and GS-IV (*Griffonia simplicifolia* lectins 1 and 2, respectively) and the unusual interface of PNA (Chandra et al., 2001; Loris et al., 1998; Brinda et al., 2005) .

Phytoagglutinins from *Lotus tetragonolobus* seeds were first investigated in 1963 (Sagan et al., 1963). Further studies revealed the existence of three fucolectins present in the seeds of *L. tetragonolobus* and provided biochemical and sugar specificity characterization (Yariv et al., 1967; Kalb, 1968). To date, the primary structure of only one isolectin from *L. tetragonolobus* seeds (LTA) has been determined, showing that LTA was composed of about 240 amino acids with a molecular mass of 26273.17 Da (Konami et al., 1990). Furthermore, LTA was shown to be specific for L-fucose and also for mono- and biantennary oligosaccharides containing L-fucosyl residues. Results from NMR and Surface Plasmon Resonance indicate that LTA recognizes diverse L-fucosyl oligosaccharides including the Le^X antigenic determinant and its divalent form (difucosyllacto-*N*-neohexaose—Fuc-octa). Interestingly, fucosylated sugars are widely distributed in nature, from plant to man, and play a variety of roles in biology, particularly in recognition processes. Several studies have reported that L-fucosyl oligosaccharides are often expressed on cells during differentiation or metastatic stages; also they have been correlated with metastatic potential in some types of tumors, particularly in neuroblastomas and adenocarcinomas (Kalb, 1968; Staudacher et al., 1999).

In addition, due to its particular biological property of recognizing L-fucosyl oligosaccharides present in glycoprotein and glycolipid receptors on cells, LTA has been used as a tool in various biochemical studies: (i) in the characterization of antigens and insect cells (Thomas and Surolia, 2000; Napier et al., 1974; Butters and Hughes, 1978) (ii) in examinations of surface saccharide composition in different cells (Mayliepfenninger and Jamieson, 1979; Gurd, 1979) and mainly (iii) in lectin histochemistry of malignant cells for the purpose of investigating anomalous membrane oligosaccharides and differentiating carcinomas (Raedler et al., 1983; Koch et al., 1983; Miettinen et al., 1983; Walker, 1984; Ulrich et al., 1985; Davidsson et al., 1987; Griffiths and Stephenson, 1988; Raju and Lee, 1988; Finne et al., 1989; Imura et al., 2004; Turner et al., 1995).

A striking feature is that although LTA has been extensively used in various experiments as an L-fucose-binding lectin, its affinity for complex L-fucosyl oligosaccharides is not well understood. Previous results revealed that the affinity for various L-fucosyl oligosaccharides is predominantly related to the LTA fucose-binding pocket. Consequently, some of these results also showed that binding affinity decreased with increasing complexity of these saccharides (Haselhorst et al., 2001), which suggests that no extended binding site was present for LTA, in contrary to what was described earlier for other lectins (Moothoo and Naismith, 1998). Moreover, Electron Microscopy and X-ray diffraction studies suggested that LTA, when cross-linked with Fuc-octa, is arranged in a tetramer form (Cheng et al., 1998), but there has been no other experimental elucidation of its mode of oligomerization. Brinda and co-workers investigated LTA by computational methods with the aim of finding a relationship between its primary sequence and the nature of its quaternary association, but their data showed only that the LTA oligomer was likely different from the other known legume lectins (Brinda et al., 2004).

In a previous article, we reported on the crystallization and preliminary X-ray analysis of LTA in the native form (Moreno et al., 2006). In the present study, we have investigated the quaternary structure and oligomerization state of LTA using a combination of Protein Crystallography and small angle X-ray scattering (SAXS) techniques. Our investigation shows that this homotetrameric lectin adopts a novel quaternary structure resulting from the assembly of two GS4-related dimers through a newly described interface. In addition, our data provide new insights into the structural basis for the recognition on L-fucosyl saccharides, lectin molecular evolution and diverse biological applications of LTA.

2. Results and discussion

2.1. Overall crystallographic structure of LTA

LTA consists of a homotetramer composed of two different dimer–dimer interfaces as shown in Fig. 1. Each monomer consists of a standard legume-like lectin ("jelly roll motif") complexed with calcium and manganese, as previously established for several legume lectins (Van Damme et al., 1998). Each monomer is composed of 234 residues and presents a glycosylation located at Asn4. The statistics of refinement data are presented in Table 1 and the pdb entry is 2EIG. Although legume lectin tetramers exist mainly as homotetramers with an internal

Table 1
Statistics of data collection, refinement and quality of the structure

	Overall resolution data set	Highest resolution data set
Data collection		
Total number of observations	225,127	30,889
Total number of unique observations	61,847	8835
R_{merge}	0.095	0.459
Highest resolution limit (Å)	2.00	2.00
Lowest resolution limit (Å)	40.42	2.11
Completeness (%)	99.4	97.9
Multiplicity	3.6	3.5
$I/\sigma(I)$	10.8	2.6
Wavelength (Å)	1.431	
Space group	$P2_1$	
Cell parameters (Å)	$a = 68.89$, $b = 65.83$, $c = 102.53$	
Refinement		
Resolution range (Å)	2.00–34.46	
R_{factor} (%)	18.81	
R_{free} (%)	25.08	
Number of non-hydrogen atoms in protein structure	7876	
Number calcium ions	04	
Number of manganese ions	04	
Number of water molecules	597	
RMS deviations from ideal values		
Bond lengths (Å)	0.021	
Bond angles (degrees)	2.201	
Temperature factors		
Average B value for main chains (Å^2)	21.00	
Average B values for water molecules (Å^2)	27.94	
Ramachandran plot		
Residues in most favored regions	700 (86.6%)	
Residues in additional allowed regions	90 (11.1%)	
Residues in generously allowed regions	10 (1.2%)	
Residues in disallowed regions	8 (1.0%)	

hydrophobic cavity as described for the lectins from the Diocleinae subtribe (ConA-like lectins), *D. biflorus*, *G. Simplicifolia* and peanut seeds, the LTA tetramer showed a different internal cavity resultant from its particular mode of tetramerization.

Several steps of molecular refinement were performed to obtain the LTA tetramer at 2.0 Å resolution. The $2F_o - F_c$ final density map contoured at 1σ showed that a few residues were not well fitted in the electron density map when using the LTA primary sequence previously established. Consequently, with the aim of solving the complete LTA structure, we used our electron density map to decipher the missing residues.

The LTA structure contains a covalently bound glycan at Asn4 which corroborates preliminary results of Konami and co-workers (Konami et al., 1990). The first carbohydrate position (GlcNac) was fitted correctly in three of the four monomers. Residual and poor densities were found surrounding the first Asn4 *N*-linked glycan residue, but there were no sufficient data to correctly model the remaining sugar residues.

2.2. Ion-binding sites

The carbohydrate-binding activity of legume lectins depends on the presence of calcium and a transition metal ion, usually manganese. Each LTA monomer contains two binding sites for calcium and manganese as described for several other legume lectins. These ions are coordinated directly by Glu118, His135, Asp120, His122, Asn124 and Asp127 and indirectly by hydrogen bonds between residues Ser145, Ile143, Gly100, Asp80, Trp126 and four water molecules. The LTA ion-binding site is conserved compared to other legume lectins, with exception of residue His122, which interacts directly with the calcium ion through the oxygen from its backbone. This difference does not seem to be relevant enough to consider LTA ion-binding site as a new folding arrangement. Moreover, the structure of LTA in the presence of Ca^{+2} and Mn^{+2} displayed the *cis*-Asp80 in a "locked" conformational state as described for the *cis*-Asp208 in concanavalin A. On the other hand, the side chain of His122 is located at the opposite side of the calcium ion, in the vicinity of the LTA carbohydrate-binding site, and its position suggests that this residue plays a significant role in the recognition of L-fucosyl residues by LTA.

Since there was no other particular structural difference observed in the LTA ion-binding site, it is reasonable to conclude that only evolutionary changes in the carbohydrate-binding site of LTA are related to its specificity for L-fucosyl.

2.3. Dimer–dimer interfaces

Most legume lectins, with a few exceptions, are known to exist mainly as homodimers and homotetramers, where these tetramers are dimers of dimers. To date, seven types of dimer associations have been described for legume lectins, including the canonical dimer and non-canonical interface (lectins from Diocleinae subtribe), EcorL (lectins from *E. corallodendron* and *Psophocarpus tetragonolobus*), GS-I and GS-IV (lectin IV from *G. simplicifolia*), DB58 (lectin from *D. biflorus*) and the unusual interface of PNA. Up to now, four models of quaternary associations for these dimers have been reported into tetrameric structures: the ConA-type (two canonical dimers packed against each other), the DBL-type (combination of canonical and DB58 types), the PNA-type (combination of canonical, GS-IV and an unusual type of interface, seen only in PNA) and GSL (with the GS-IV type and an unusual type of interface, seen only in *G. simplicifolia* type lectins) (Fig. 3). Furthermore, sequences from the five types of dimeric interfaces described above were investigated by bioinformatics procedures, and the results showed that for each type, a specific signature motif determines and differentiates each dimerization state (Brinda et al., 2004). On the other hand, legume lectins may exist as a monomer, dimer or tetramer under different conditions of pH and solvent; however, analysis of known dimer interfaces, using

computational sequence investigation, is not satisfactory to typify with a high level of certainty the real quaternary oligomerization state of legume lectins at different chemical environments.

Our work reports a crystallographic structure of LTA, a homotetramer not previously described for any legume lectins. Fig. 3 illustrates legume lectin tetramers, including the new tetramer of LTA. Additionally, LTA is composed of two dimer interfaces, a GS4-type, similar to that from *G. simplicifolia* (1GSL) and an unknown interface that is displayed in Fig. 1c. The *G. simplicifolia* lectin (1GSL) is essentially a tetramer with a GS4-type dimer and an unusual dimer interface. The GS4 dimer is formed by the association of the six-stranded β-sheets of each monomer, with the β-strands of one sheet being perpendicular to those of the other (Delbaere et al., 1993).

The GS4 dimers of LTA, PNA and GSL share the same spatial configuration, but while the two first show nearly equal angular arrangement of the monomers, they contrast to GSL that shows a considerable difference in the angle of association. Sequence alignment of the residues located in the GS-IV back-to-back interface revealed high identity between LTA and PNA; identity which was obtained in a much less extent when comparing both of them to GSL, as depicted in Fig. 2. This difference is likely the reason for why this angular discrepancy appears.

Other two aspects may ascribe for aforementioned discrepancy. The first one, the glycosylation at Asn4, which is located nearby the contact interface between monomers A and B (Fig. 1c), may influence the molecular arrangement, accounting for the difference seen in GS4-like dimer of LTA in contrast of that observed in GS4 of GSL. Second, crystal packing may also contribute to the spatial arrangement of LTA dimers, influencing the displacement seen in the GS4-like dimer of LTA, since this displacement

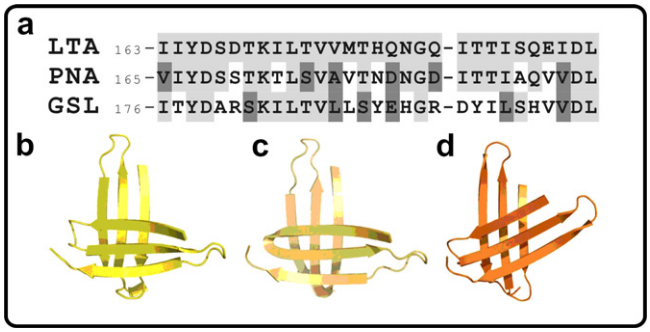


Fig. 2. (a) Sequence alignment of the residues located in the GS-IV back-to-back interface LTA, PNA and GSL. The identical residues are highlighted in light grey and the conserved residues are highlighted in dark grey. Notice the great similarity between LTA and PNA in this region. Three-dimensional representation of the interface region from the GS-IV back-to-back interface for (b) LTA, (c) PNA and (d) GSL. Observe that LTA and PNA have a nearly equal angular arrangement of the monomers, contrasting to GSL whose angle of association is distinct from the two other lectins.

could not be confirmed by the SAXS technique on account of its lower resolution range.

As reported earlier, the LTA crystal structure revealed an unknown dimer–dimer interface never described before for legume lectins. While the LTA GS4-dimer type is an association between β-strands of adjacent monomers, the unknown dimer interface of LTA is stabilized by hydrogen bonds and hydrophobic contacts present in inter-contacts between loops of monomers A/C and B/D. Two contact interfaces are described here for the new LTA-dimer interface, and the residues involved are shown in Fig. 4.

Brinda and co-authors (2004) used lectins with unknown three-dimensional structures until then to test the ability of their method to predict the nature of quaternary association in legume lectins given the sequence. LTA was one

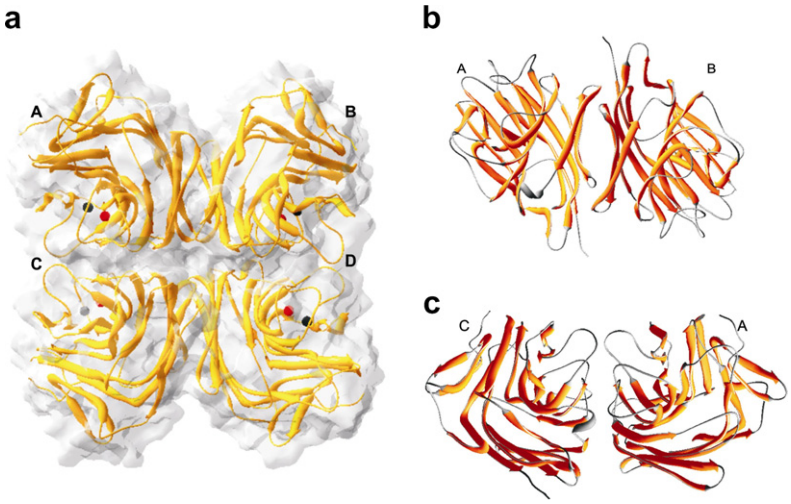


Fig. 1. Overall crystal structure of LTA. (a) The LTA homotetramer. Four identical domains are interconnected by two dimer–dimer interfaces. (b) The GS4-like dimer seen in LTA. Both monomers are interconnected by back-to-back interfaces as described for the lectin from *G. simplicifolia*. (c) The new LTA-dimer observed in LTA tetramer configuration. Both monomers are interconnected by hydrophilic interactions at external loops and a hydrophobic core. figures were generated by *Deep View*.

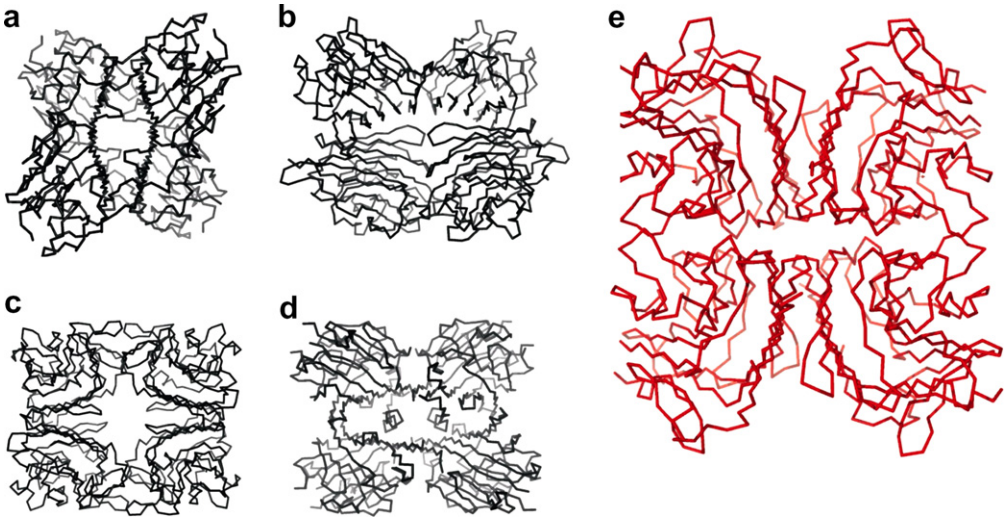


Fig. 3. Representation of the five types of tetramer associations among legume lectins. (a) Ribbon representation of a ConA-like lectin which is composed of two canonical dimers interconnected by a non-canonical interface. (b) The peanut lectin showing a tetramer composed of a canonical dimer, a back-to-back interface and an unusual type of dimer–dimer interface seen only in this case. (c) The lectin from *G. simplicifolia* which represents two back-to-back dimer interfaces and an unusual dimer interface. (d) The *D. biflorus* lectin which is composed of a canonical dimer and two DB58 dimer–dimer interfaces. (e) The LTA tetramer composed of two back-to-back interfaces interconnected by two particular LTA-dimer interfaces. The dimer–dimer interactions yielded a homoplanar tetramer. Figures were generated by *Pymol.*

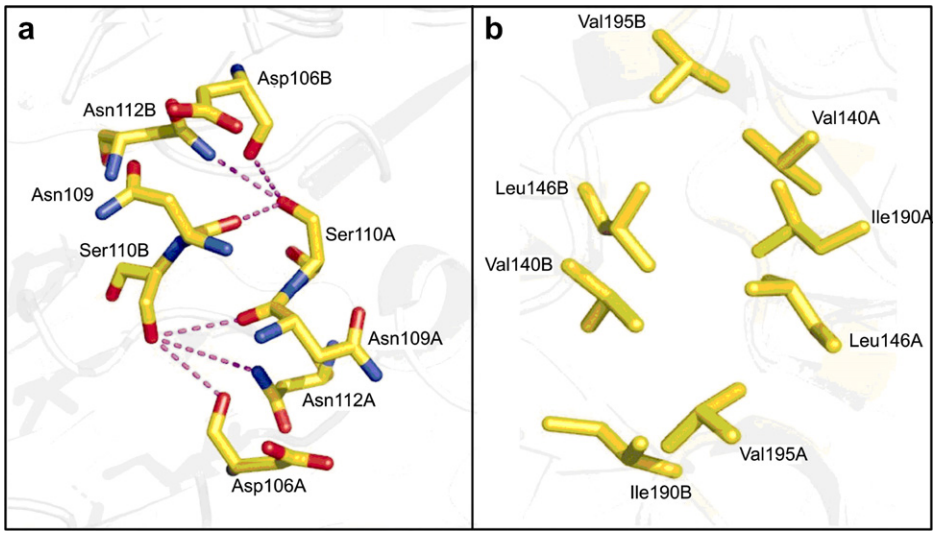


Fig. 4. The LTA-dimer interface. (a) Hydrogen bonds between loops from Asp106 to Asn112 at the surface of LTA molecule. (b) The hydrophobic residues arranged at the center of the LTA-dimer interface. Both figures are representations of the contacts at the interface between monomers A and B. Figures were generated by *Pymol.*

of the lectins used in order to do that. They concluded that all seven types of interface types were likely to be destabilized in the case of LTA. Our results prove them to be partially correct, since there is indeed a novel interface type in LTA, but there's also an already known type which is nearly identical to the one present in PNA.

The crystallographic refinement revealed, at initial stages, that the loops from residues ILE104 to GLN113 and Gly86 to ASP92 were not well defined in the electronic density map. Thus, most of the time required for converging R_{free} and R_{factor} was a consequence of amino acid modeling at the loops from this new dimer interface. We named this "unknown" dimer interface the LTA-dimer, and we suggest here that these variations observed in LTA loop pattern in the LTA-dimer may be strictly related to the association of GS4-like dimers in the tetramer.

2.4. Small angle X-ray scattering and the homotetramer structure of LTA

As described previously, legume lectins are basically dimers of dimers and the modes of quaternary association

is a consequence of the interactions observed between these dimers. The crystal structure of LTA belongs to the $P2_1$ space group containing four monomers per asymmetric unit, resulting in a regular crystallographic homotetramer as arranged in Fig. 1. The initial crystallographic results showing a new tetrameric form observed for LTA needed further support. First, the differences seen in overall structure could be a consequence of crystal packing, and second, the crystal structure of diverse proteins may assume a structural form that is not observed in a biological aqueous environment (Filgueira de Azevedo et al., 2003).

The LTA tetramer is a combination of two GS4-type dimers which may assume the structural conformation seen in *G. simplicifolia* lectin depending on the space group. However, the presence of four monomers per asymmetric unit yielded the LTA homotetramer as arranged in Fig. 1a, and moreover, the symmetry operations for $P2_1$ space group are not sufficient to configure an LTA tetramer similar to that observed for *G. simplicifolia* lectin in the crystal packing configuration. In addition, the glycan position at ASN4 is a new structural feature of LTA that is not yet associated with quaternary folding. While *G. simplicifolia* lectin possesses a non-glycosylated ASN5 placed at the center of molecule, the ASN4 of LTA is placed at the external face of the tetramer.

Studies of the non-glycosylated legume lectins ConA and PNA proposed that the variability of quaternary association in these cases is related to differences in the amino acid sequence of these proteins, since glycans were absent in both structures (Hardman and Ainswort, 1972; Banerjee et al., 1996). Moreover, the crystal structure of the glycolectin *Erythrina cristagalli* (*N*-linked at Asn17) and its recombinant form showed that both lectins associate into identical dimers confirming that the presence of *N*-linked glycan does not influence the mode of dimerization of ECL and suggesting that intrinsic factors to the primary structure of the lectin dictate its quaternary structure (Turton et al., 2004). Thus, we come to interesting questions: were there no glycosylation at Asn4 of LTA, would this new assembly be formed? And what forces would allow LTA to assume this new spatial configuration?

Preliminary analysis of LTA and GSL showed that both are glycosylated at Asn4 and Asn18 (not at Asn5 in GSL), respectively, and that their quaternary structures present the *N*-linked glicans exposed to the surface of molecules (Fig. 5). On the other hand, symmetry operations could allow for LTA GS4-like dimers to assume a conformation similar of that of GSL; but for that, its *N*-linked glycan at Asn4 must be oriented toward the center of the tetramer, which is not a common feature observed in legume lectins. Furthermore, the position of the Asn4 glycan on the surface of the LTA tetramer is also a consequence of steric restraints, since there is no suitable chemical location to place all glycans together at the center of the molecule, and also because a favorable environment for *N*-linked carbohydrates is frequently found with solvent exposure.

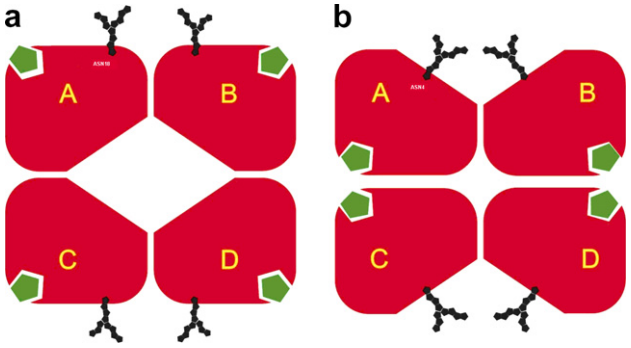


Fig. 5. Schematic representation of dimer association between lectins from *G. simplicifolia* and *L. tetragonolobus*. The red boxes represent a legume lectin monomer. The black dots represent the glycosylation pattern of each lectin and the green boxes represent the carbohydrate-binding site of each monomer. (a) Representation of the two back-to-back dimer interface of GSL. Both dimers are interconnected with an internal hydrophobic cavity of approximately 25 Å. GSL is glycosylated at Asn18 and its *N*-linked oligosaccharides are arranged at the surface of molecule. The carbohydrate-binding sites are at equivalent distances in the equatorial position of each monomer. (b) Representation of LTA tetramer which is a molecule composed of two back-to-back dimer interfaces as arranged in (a). The LTA tetramer is different from that observed for GSL. The glycan at Asn4 instead of Asn18 is a particular characteristic of LTA. As a consequence of its natural tetramerization, the LTA tetramer possesses the carbohydrate-binding sites at positions close to each other. The ability of LTA to recognize L-fucose and the tetramer folding are key features that determine the model of the cross-linking complexes formed by LTA.

Thus, we conclude that differences seen between the LTA tetramer's crystal packing and other legume lectins are strictly related to variations seen in primary structure. Furthermore, although we cannot conclude whether the presence of surface-exposed *N*-linked glicans is a cause for LTA new assembly, we consider it an additional evidence for the existence of LTA's novel arrangement.

Finally, to confirm our investigations on LTA quaternary structure, we have used the small angle X-ray scattering (SAXS) technique to determine definitively how LTA behaves in aqueous solution. Analyses of maximum dimension (D_{max}), external surface area (S) and volume of hydration (V_h) from LTA crystal coordinates yielded values similar to that obtained with SAXS. The values of radii of gyration $R_g = 33.2$ Å, $D_{max} = 100$ Å, $S = 30 \times 10^3$ Å^2 and $V_h = 138 \times 10^3$ Å^3, obtained from SAXS curves, showed that LTA dimensions in aqueous solution corroborate those calculated for the crystalline structure (Fig. 7). Moreover, regarding qualitative information about LTA surface contour, the $\rho(r)$ plot displayed in Fig. 6a revealed that the pattern assumed by this function corresponds to a particle with the appearance of a lamellar shape with a maximum dimension of 100 Å (Volkov and Svergun, 2003). On the other hand, we also generated the external envelope of LTA by the *ab initio* shape determination method to compare it to the crystallographic structure and to investigate how LTA monomers are interconnected with each other. Fig. 6 shows the tetramer envelope of LTA which represents a laminar orthorhom-

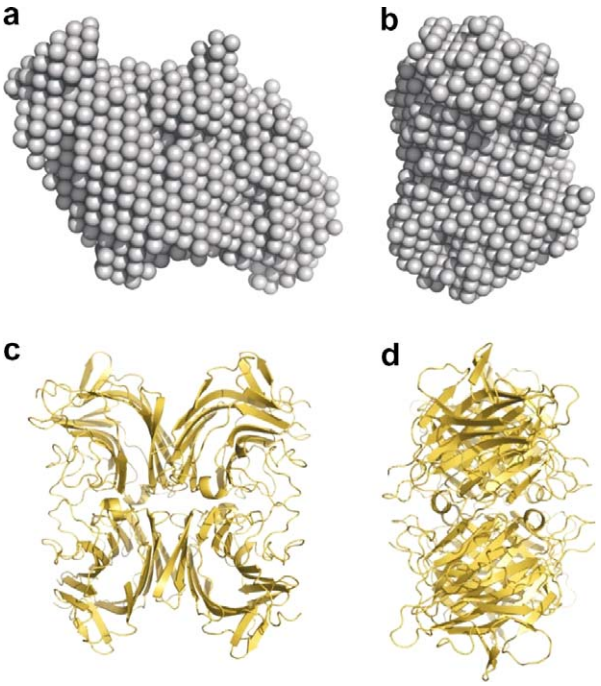


Fig. 6. Small angle X-ray scattering of LTA. (a and b) The dummy atoms. The models were generated by SUPCOMB20 using a difference of 90°. (c and d) The LTA tetramer as a cartoon depiction with a difference of 90°.

bic shape similar to that observed for the LTA crystal structure. Both GS4-like and LTA-dimers can be identified in Fig. 6a.

2.5. LTA carbohydrate-binding site and biological remarks

Crystallographic studies of various legume lectins have revealed the structural basis for molecular recognition of fucosyllactose, fucosylgalactose, Lewis-B and Y-human blood group determinants (*Ulex europaeus* II—1QOT, 1JXN, *E. cristagalli*—1GZ9, *G. simplicifolia*—1LED, 1GSL). Although such complexes were obtained in the presence of carbohydrates containing L-fucosyl residues or derivatives, little is known about the L-fucose monosaccharide primary site in any legume lectins. In addition, structural studies also revealed that the presence of L-fucose in these complexes cited above is not essential for the lectin to bind the whole glycan. For example, *U. europaeus* II lectin was shown to complex with fucosylgalactose and galactose but not with L-fucose, indicating that fucose by itself is not essential to inhibit UEA-II. Hence, Loris and co-workers suggested that although UEA-II possesses a promiscuous carbohydrate-binding site, its primary affinity is designed for both GlcNAc and Gal, but not for L-fucose (Loris et al., 2000). Moreover, LTA precipitating activities using L-fucosyl oligosaccharides has been investigated, and only L-fucose and L-fucosyl oligosaccharides were able to inhibit its hemagglutinating activity. Furthermore, the formation of a unique homogeneous cross-linked complex precipitate between tetrameric

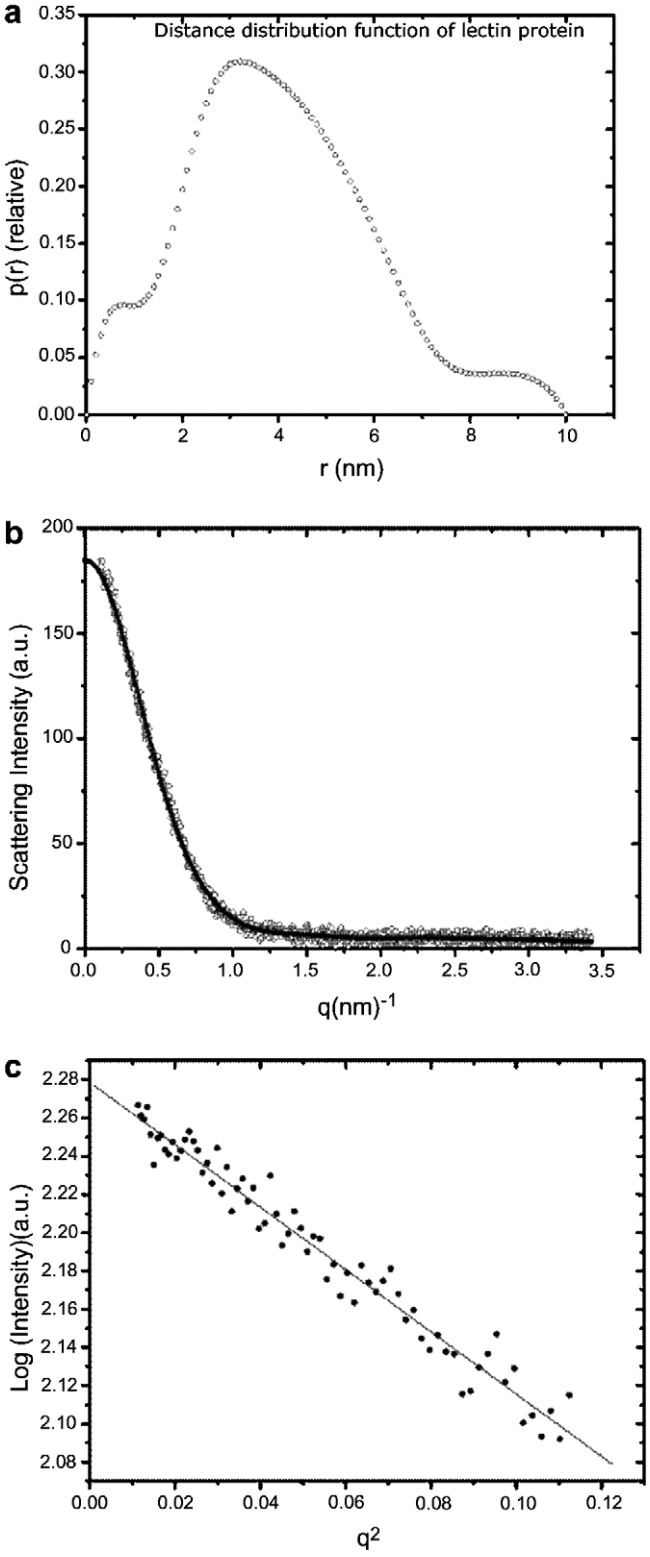


Fig. 7. Small angle X-ray scattering data of LTA. (a) $\rho(r)$ showing the maximum dimension of scattering to be 10 nm. (b) Curve plotting intensity versus scattering angle used to create the low resolution model of LTA. (c) log intensity versus log q^2 plot.

LTA and diverse biantennary L-fucosyl oligosaccharides are inhibited by 0.1 M fucose[38]. Thus, the experimental results described above make it clear that LTA recognizes

L-fucose at its primary site, but only superficial information about its binding mechanism is provided.

The binding site for monosaccharides in legume lectins is highly conserved; however, as consequence of molecular evolution, discrete mutations, which dictate the differences in the affinity of lectins for small carbohydrates, appear. While some amino acid residues are extensively conserved, some of them are found particularly in each type of lectin. The LTA residues ASP, ASN and GLY at positions 80, 124 and 100, respectively, are also conserved in EcorL, DB58, GSL and PNA. In LTA, these residues are structurally positioned in the same spatial configuration as equivalent residues of the above lectins. These amino acids are important in sugar-binding since they play a crucial role in stabilizing and coordinating a specific monosaccharide inside the binding pocket through hydrogen bonds, which are formed between their side chains and the hydroxyl groups of carbon-3 and carbon-4. Thus, the fucose-binding site of LTA is conserved on the face of residues Asp80, Asn124 and Gly100.

The subsequent analysis of the structural site of LTA aimed to investigate the non-conserved residues that are also involved in fucose recognition. The fucose-binding pocket of LTA is shown in Fig. 8. We placed the L-fucose monosaccharide in the conserved orientation according to the highly conserved monosaccharide site of other related legume lectins. Consequently, fucose positioning regarding the conserved monosaccharide binding turned carbon-3 and carbon-4 to equivalent positions of Asp, Asn and Gly based on that previously established for other complexes of similar lectins. This arrangement led us to investigate the role of the remaining two residues located in the area surrounding bound L-fucose.

The monosaccharide L-fucose differs from other related sugars due to the presence of a methyl group bound to carbon-5 instead of a hydroxyl. Moreover, the electron density map of Trp209 is well defined and its side chain position is a few angstroms apart from the CH_3 group of L-fucose, which indicates that this evolutionary punctual mutation may be involved in the stabilization of this hydrophobic contact. Finally, due to the particular conformation of the LTA tetramer, the carbohydrate-binding sites are arranged more closely than observed in other legume lectins (Fig. 5b).

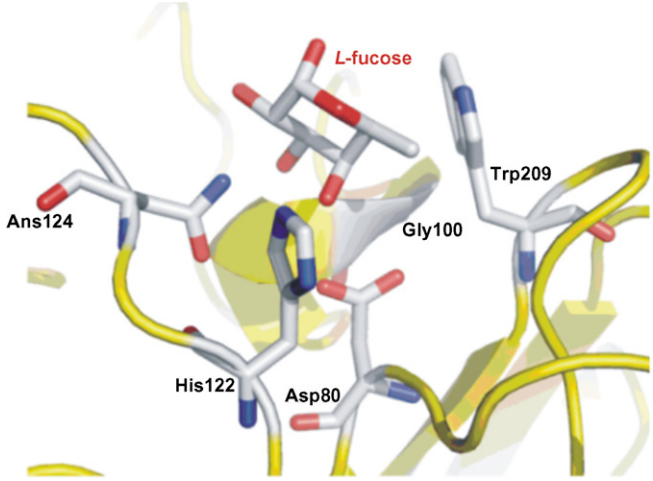


Fig. 8. The carbohydrate-binding site of LTA. L-Fucose is located at the conserved binding pocket of legume lectins. Residues Asp80, Gly100 and Asn124 are also at the conserved positions which interact with hydroxyls of carbon-3 and carbon-4. Trp-209 is close to the CH_3 group of L-fucose carbon-1 and Glu212 is a few angstroms from carbon-5.

2.6. LTA cross-linking properties

LTA showed some specific differences in its carbohydrate-binding site. Besides, the particular tetramerization model is strictly related to its ability to bind different bivalent carbohydrates and to form cross-linking interactions, since the formation of highly ordered lattices depends on the saccharide specificity and the spatial configurations of the sugar-binding sites in the whole tetramer.

ConA dimers have been used to generate all possible orientations of LTA consistent with the observed symmetry of the cross-linking pattern lattice of this lectin in the presence of divalent $Lewis^X$ oligosaccharide. The best results were obtained when the relative orientation of the two LTA dimers in a tetramer were varied relative to a range of about 20° as compared to dimmers in the ConA tetramer, leaving them more parallel in relation one another.

The reason for why the best results were accomplished with this kind of manipulation is related to the novel tetramer structure described here for LTA. The LTA tetramer observed in Fig. 5 shows that the fucose-binding sites are not at equivalent distances between each other, as observed for ConA. While the carbohydrate-binding sites of LTA monomers A/C and B/D are at a distance of 30 Å, the distance observed for the sites between monomers A/B and C/D are at approximately 57 Å. The positions of the four carbohydrate-binding sites of LTA are consequence of its mode of tetramerization, since the monomers are in a planar view in contrast to that observed in ConA. On the other hand, ConA-like lectins possess the four carbohydrate-binding sites of their tetramers at equivalent distances of approximately 60 Å, which corroborates of distances between binding sites A/B and C/D of LTA.

Since Cheng and co-workers used two LTA-dimers similar of that observed for ConA to build a tetramer topology for the whole LTA and had only low resolution data available, they were not able to reach the novel tetramer structure of LTA, which successfully explains the highly organized two-dimensional cross-linked lattice which they proved to be dependent on the geometry and symmetry of the binding sites of the lectin (Cheng et al., 1998).

Hence, it seems that this angular discrepancy between the position of ConA and LTA dimers observed by Electron Microscopy are a key feature for understanding how LTA forms a specific two-dimensional type-2 cross-linked complex between divalent oligosaccharides, since changes in the orientation of its dimmers corroborate our tetramer topoloy described in Fig. 5.

Thus, the particular quaternary structure and the new L-fucose-binding site of LTA are key features in describing the differences seen in biological activities already determined. The ability of LTA to form two-dimensional type-2 organized clusters with divalent L-fucosyl oligosaccharides (Bhattacharyya et al., 1990) may be related to the functions of similar carbohydrates as putative receptors on cell surfaces. Organized clusters mediated by cross-linking complex structures occur with LTA in the presence of diverse biantennary L-fucosyl oligosaccharides (Cheng et al., 1998; Bhattacharyya et al., 1990). These biological phenomena do not seem to be only a consequence of the LTA's affinity for L-fucose, since the evidence presented in this paper of a new LTA tetramer provides new insights into the understanding of how LTA is able to carry out signal transduction processes due to the ability to bind and cross-link specific glycoproteins and glycolipid receptors.

3. Conclusions

In the present paper, we showed that the lectin from *L. tetragonolobus* seeds, LTA, is a homotetramer composed of four legume lectin domains. The new tetramer was found under both crystalline and aqueous conditions, and furthermore, we observed some particular molecular features that appear to account for LTA to assume this new spatial configuration. The spatial position of the *N*-linked oligosaccharide at Asn4 and the LTA-dimer contacts as arranged in Fig. 4 are key features for the understanding of how LTA behaves as a homotetramer with its two GS4-like dimers arranged in an inverted position compared to the tetramer described for the *G. simplicifolia* lectin.

The present findings also provide insights into the molecular ability of LTA to recognize its specific monosaccharide. Such descriptions about its carbohydrate-binding site together with the model of tetramer oligomerization afford a structural basis for understanding the specificity of LTA for some complex bivalent L-fucosyl oligosaccharides, since this type of interaction may account for the formation of homogeneous cross-linking carbohydrate complexes that govern diverse biological signal transduction processes. Moreover, our results present new data for further studies about the evolution of binding sites and molecular oligomerization of legume lectins.

4. Materials and methods

4.1. Crystallization, data collection and processing

Lotus tetragonolobus agglutinin (LTA) was purchased from Sigma–Aldrich (USA) and was crystallized as described elsewhere (Moreno et al., 2006). Small crystals were submitted to X-ray diffraction at LNLS (Laboratório Nacional de Luz Síncrotron, Campinas—Brazil) using a synchrotron radiation source (MX1-station). X-ray diffraction data were collected at a wavelength of 1.47 Å using a CCD detector (Mar research). Two data sets were collected at resolutions of 2.0 and 2.35 Å, respectively. The images were processed, indexed, integrated and scaled using the software *MOSFLM* and *SCALA* (CCP4, 1994; Leslie, 1992).

4.2. Structure determination

The 2.0 Å resolution crystal data set was solved by the molecular replacement method using *MOLREP* software (Vagin and Teplyakov, 1997). We have used several legume lectin coordinates to solve the phase problem, and the best model was chosen based on the magnitude of the CC (correlation coefficient) and R_{factor}. The best model was obtained using the peanut lectin monomer as a search model (PDB code 1CR7—Ravishankar et al., 2001). Initially, rotation and translation functions yielded a tetramer according to the content of the asymmetric unit (four monomers per asymmetric unit) (Moreno et al., 2006).

The position and orientation of the four monomers (Chains A, B, C and D), as a single rigid body entity, were refined by 20 cycles with *REFMAC* (Murshudov et al., 1996) using reflections in a resolution range of 102.5–2.0 Å. Several steps of rebuilding, interspersed with restrained refinement, yielded the current model at 2.0 Å resolution. The appropriate LTA amino acids were positioned by inspection of the $2F_o - F_c$ and $F_o - F_c$ maps, and the stereochemistry was monitored frequently with *PROCHECK* and PARMODEL software (Laskowski et al., 1993; Uchoa et al., 2004). Finally, ions were arranged in the model manually, and the water molecules were positioned after several steps of restrained refinement with *ARP/WARP* and examined manually by *XTALVIEW* (CCP4, 1994; McRee, 1999).

4.3. Small angle X-ray scattering data collection

The X-ray scattering data (SAXS) of LTA were collected at LNLS (Laboratório Nacional de Luz Síncrotron, Campinas—Brazil) using a synchrotron radiation source at the D11A-station beam line. LTA was dissolved in 10 mM Tris–HCl, pH 7.6, containing 18 mM $CaCl_2$ and $MnCl_2$ at three different concentrations of 10.0, 6.0 and 3.0 mg/mL. The samples were then submitted to data collection at room temperature (about 24 °C) from a distance of 831 mm to the detector. The scattering curves were recorded at a wavelength of $\lambda = 1.48$ Å and the scattering intensities were obtained in the range of $0.0144\ \text{Å}^{-1} < q < 0.3423\ \text{Å}^{-1}$, where q is the scattering vector $q = 4\pi\sin\theta/\lambda$ and 2θ is the scattering angle.

The SAXS data were desmering and the distance distribution function $p(r)$ was obtained by indirect Fourier transformation using the program GNOM (Svergun, 1992). The radii of gyration, $R_g (R_g = 2.628\sqrt{-\alpha})$, was obtained from the slope, α, of the linear region of $\log I$ versus q^2 plots using the Guinier's Law (Guinier and Fournet, 1955).

Parameters of volume and external area of LTA were calculated using $V = 2\pi^2 I(0)/Q$ and $S = \pi V \lim[I(q)q^4]/Q$,

respectively, for Q representing the invariant and $I(0)$ the result from Guinier's Law, where $e^{-\frac{Rg^2q^2}{3}} = 1$ (Porod, 1982).

The external envelope of LTA was then generated by *ab initio* shape determination method, and 20 independent *ab initio* models were averaged for the purpose of improving the quality of shape reconstruction using DAMMIM and DAMAVER programs (Svergun, 1999). Finally, the superposition between X-ray crystallographic data (PDB) and dummy SAXS atom coordinates was carried out by SUPCOMB20 (Kozin and Svergun, 2001).

4.4. Structure quality and manipulation

LTA model manipulation, analysis and the confection of figures were executed with the aid of *VMD* and *Deep View* (Guex and Peitsch, 1997; Humphrey et al., 1996). The LTA structure geometry was monitored by *PROCHECK* and Ramachandran plot. The temperature factor analysis was performed using *BAVERAGE* software to obtain graphs and data of *B-factor* versus residues for each chain and water molecules (CCP4, 1994).

Acknowledgments

The authors would like to thank funding from CNPq, CAPES and FAPESP. We also thank Dr. A. Leyva for English language editing of the manuscript. W.F. de Azevedo Jr. and B.S. Cavada are senior investigators of CNPq.

References

Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A., Vijayan, M., 1996. Conformation, protein–carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin-lactose complex. J. Mol. Biol. 259, 281–296.

Bhattacharyya, L., Fant, J., Lonn, H., Brewer, C.F., 1990. Binding and precipitating activities of lotus-tetragonolobus isolectins with L-fucosyl oligosaccharides—formation of unique homogeneous cross-linked lattices observed by electron-microscopy. Biochemistry 29, 7523–7530.

Brinda, K.V., Mitra, N., Surolia, A., Vishveshwara, S., 2004. Determinants of quaternary association in legume lectins. Protein Sci. 13, 1735–1749.

Brinda, K.V., Surolia, A., Vishveshwara, S., 2005. Insights into the quaternary association of proteins through structure graphs: a case study of lectins. Biochem. J. 391, 1–15.

Butters, T.D., Hughes, R.C., 1978. Lectin binding to mosquito aedes-aegyptii and human kb cells—structural comparisons of membrane oligosaccharides. Carbohydr. Res. 61, 159–168.

Collaborative Computational Project, Number 4. 1994. The Ccp4 suite—programs for protein crystallography. Acta Crystallogr. D 50, 760–763.

Chandra, N.R., Prabu, M.M., Suguna, K., Vijayan, M., 2001. Structural similarity and functional diversity in proteins containing the legume lectin fold. Protein Eng. 14, 857–866.

Cheng, W., Bullitt, E., Bhattacharyya, L., Brewer, C.F., Makowski, L., 1998. Electron microscopy and X-ray diffraction studies of *Lotus tetragonolobus* A isolectin cross-linked with a divalent Lewis(x) oligosaccharide, an oncofetal antigen. J. Biol. Chem. 273, 35016–35022.

Dam, T.K. et al., 2005. Thermodynamic, kinetic, and electron microscopy studies of concanavalin A and *Dioclea grandiflora* lectin cross-linked with synthetic divalent carbohydrates. J. Biol. Chem. 280, 8640–8646.

Davidsson, P., Karlsson, B., Svennerholm, L., 1987. Glycoprotein pattern in human-brain tumors studied using lectin binding after sodium dodecyl sulfate–gel electrophoresis and protein blotting. Brain Res. 412, 254–260.

Delatorre, P., Rocha, B.A., Souza, E.P., Oliveira, T.M., Bezerra, G.A., Moreno, F.B., Freitas, B.T., Santi-Gadelha, T., Sampaio, A.H., Azevedo Jr., W.F., Cavada, B.S., 2007. Structure of a lectin from *Canavalia gladiata* seeds: new structural insights for old molecules. BMC Struct. Biol. 7, 52.

Delbaere, L.T.J. et al., 1993. Structures of the lectin-Iv of *Griffonia simplicifolia* and its complex with the Lewis-B human blood-group determinant at 2.0-angstrom resolution. J. Mol. Biol. 230, 950–965.

Filgueira de Azevedo Jr., W., dos Santos, G.C., dos Santos, D.M., Olivieri, J.R., Canduri, F., Silva, R.G., Basso, L.A., Renard, G., da Fonseca, I.O., Mendes, M.A., Palma, M.S., Santos, D.S., 2003. Docking and small angle X-ray scattering studies of purine nucleoside phosphorylase. Biochem. Biophys. Res. Commun. 309 (4), 923–928.

Finne, J., Castori, S., Feizi, T., Burger, M.M., 1989. Lectin-resistant variants and revertants of mouse melanoma-cells—differential expression of a fucosylated cell-surface antigen and altered metastasizing capacity. Int. J. Cancer 43, 300–304.

Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714–2723.

Guinier, A., Fournet, G., 1955. Small-Angle Scattering of X-rays. Jonh Wiley & sons, NY.

Gurd, J.W., 1979. Molecular and biosynthetic heterogeneity of fucosyl glycoproteins associated with rat-brain synaptic junctions. BBA 555, 221–229.

Griffiths, D.W.R., Stephenson, T.J., 1988. Lectin agglutinins—a comparative-study of *Lotus tetragonolobus* and *Ulex europaeus*-1 as markers of vascular endothelium in follicular thyroid-carcinoma. Med. Lab. Sci. 45, 45–51.

Hardman, K.D., Ainswort, C.F., 1972. Structure of concanavalin-a at 2.4-a resolution. Biochemistry 11, 4910–4919.

Haselhorst, T., Weimar, T., Peters, T., 2001. Molecular recognition of sialyl Lewis(x) and related saccharides by two lectins. J. Am. Chem. Soc. 123, 10705–10714.

Humphrey, W., Dalke, A., Schulten, K., 1996. VMD: visual molecular dynamics. J. Mol. Graph. 14, 33.

Imura, J. et al., 2004. Multilocular cystic renal cell carcinoma: a clinicopathological, immuno- and lectin histochemical study of nine cases. Apmis 112, 183–191.

Kalb, A.J., 1968. Separation of 3 L-fucose-binding proteins of *Lotus tetragonolobus*. BBA 168, 532.

Koch, B. et al., 1983. The binding of different lectins on peripheral-blood mononuclear-cells from patients with chronic inflammatory and malignant diseases. Immunobiology 164, 99–109.

Konami, Y., Yamamoto, K., Osawa, T., 1990. The primary structure of the *Lotus tetragonolobus* seed lectin. FEBS Lett. 268, 281–286.

Kozin, M.B., Svergun, D.I., 2001. Automated matching of high- and low-resolution structural models. J. Appl. Crystallogr. 34, 33–41.

Laskowski, R.A., Macarthur, M.W., Moss, D.S., Thornton, J.M., 1993. Procheck—a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.

Leslie, A.G.W., 1992. Jnt CCP4/ESF–EACBM. Newsl. Protein Crystallogr. 26.

Loris, R., Hamelryck, T., Bouckaert, J., Wyns, L., 1998. Legume lectin structure. BBA Protein Struct. Mol. Enzyme 1383, 9–36.

Loris, R. et al., 2000. Structural basis of carbohydrate recognition by lectin II from *Ulex europaeus*, a protein with a promiscuous carbohydrate-binding site. J. Mol. Biol. 301, 987–1002.

Mayliepfenninger, M.F., Jamieson, J.D., 1979. Distribution of cell-surface saccharides on pancreatic cells. 2. Lectin-labeling patterns on mature guinea-pig and rat pancreatic cells. J. Cell Biol. 80, 77–95.

McRee, D.E., 1999. XtalView Xfit—a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. 125, 156–165.

Miettinen, H.H.A. et al., 1983. Cellular-origin and differentiation of renal carcinomas—a fluorescence microscopic study with kidney-specific antibodies, antiintermediate filament antibodies, and lectins. Lab. Invest. 49, 317–326.

Moothoo, D.N., Naismith, J.H., 1998. Concanavalin A distorts the beta-GlcNAc-(1 -> 2)-Man linkage of beta-GlcNAc-(1 -> 2)-alpha-Man-(1 -> 3)-[beta-GlcNAc-(1 -> 2)-alpha-Man-(1 -> 6)]-Man upon binding. Glycobiology 8, 173–181.

Moreno, F.B.M.B., Martil, D.E., Cavada, B.S., deAzevedo, W.F., 2006. Crystallization and preliminary X-ray diffraction analysis of an anti-H(O) lectin from *Lotus tetragonolobus* seeds. Acta Crystallogr. F 62, 680–683.

Murshudov, G., Vagin, A., Dodson, E., 1996. Application of maximum likelihood refinement in the refinement of protein structures. In: Proceedings of Daresbury Study Weekend.

Napier, P.W., Everhart, D.L., Grundbac, F.J., 1974. Characterization of a human saliva antigen precipitated by a lectin from *Lotus tetragonolobus*. Vox Sang. 27, 447–458.

Porod, G., 1982. In: Glatter, O., Kratky, O. (Eds.), Small-Angle X-ray Scattering. Academic Press, London, pp. 17–51.

Raedler, A., Schmiegel, W.H., Raedler, E., Arndt, R., Thiele, H.G., 1983. Lectin-defined cell-surface glycoconjugates of pancreatic-cancer cells and their nonmalignant counterparts. Exp. Cell Biol. 51, 19–28.

Raju, G.C., Lee, Y.S., 1988. Lectin expression in carcinoid-tumors of the gastrointestinal-tract. J. Pathol. 154, 263–268.

Ravishankar, R., Thomas, C.J., Suguna, K., Surolia, A., Vijayan, M., 2001. Crystal structures of the peanut lectin-lactose complex at acidic pH: retention of unusual quaternary structure, empty and carbohydrate bound combining sites, molecular mimicry and crystal packing directed by interactions at the combining site. Proteins 43, 260–270.

Sagan, Z., Olszewska, I., Pawlus, M., 1963. Research on phytoagglutinin from the seeds of *Lotus tetragonolobus* L. and some other plants. Acta Pol. Pharm. 20 (20), 83–88.

Staudacher, E., Altmann, F., Wilson, I.B.H., Marz, L., 1999. Fucose in *N*-glycans: from plant to man. BBA Gen. Subj. 1473, 216–236.

Svergun, D.I., 1992. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. 25, 495–503.

Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. Biophysical J76, 2879–2886.

Thomas, C.J., Surolia, A., 2000. Mode of molecular recognition of L-fucose by fucose-binding legume lectins. Biochem. Biophys. Res. Commun. 268, 262–267.

Turner, G.A., Goodarzi, M.T., Thompson, S., 1995. Glycosylation of alpha-1-proteinase inhibitor and haptoglobin in ovarian-cancer—evidence for 2 different mechanisms. Glycoconj. J. 12, 211–218.

Turton, K., Natesh, R., Thiyagarajan, N., Chaddock, J.A., Acharya, K.R., 2004. Crystal structures of *Erythrina cristagalli* lectin with bound N-linked oligosaccharide and lactose. Glycobiology 14, 923–929.

Uchoa, H.B. et al., 2004. Parmodel: a web server for automated comparative modeling of proteins. Biochem. Biophys. Res. Commun. 325, 1481–1486.

Ulrich, W., Horvat, R., Krisch, K., 1985. Lectin histochemistry of kidney tumours and its pathomorphological relevance. Histopathology 9 (10), 1037–1050.

Vagin, A., Teplyakov, A., 1997. MOLREP: an automated program for molecular replacement. J. Appl. Crystallogr. 30, 1022–1025.

Van Damme, E.J.M., Peumans, W.J., Barre, A., Rouge, 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.

Volkov, V.V., Svergun, D.I., 2003. Uniqueness of ab initio shape determination in small-angle scattering. J. Appl. Crystallogr. 36, 860–864.

Yariv, J., Kalb, A.J., Katchals, E., 1967. Isolation of an L-fucose binding protein from *Lotus tetragonolobus* Seed. Nature 215, 890.

Walker, R.A., 1984. The binding of peroxidase-labelled lectins to human breast epithelium. III-Altered fucose-binding patterns of breast carcinomas and their significance. J. Pathol. 144 (2), 109–117.

REFERÊNCIAS BIBLIOGRÁFICAS

Andersen N.H. et al. Hevein: NMR assignment and assessment of solution-state folding for the agglutinin-toxin motif. **Biochemistry**. Feb 16;32(6):1407-22, 1993.

Bagaria, A., Surendranath, K., Ramagopal, U.A., Ramakumar, S., Karande, A.A. Structure-Function Analysis and Insights into the Reduced Toxicity of Abrus precatorius Agglutinin I in Relation to Abrin. **J.Biol.Chem**. v281 pp.34465-34474 , 2006.

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**. v259 pp.281-296 , 1996.

Barbieri, L., Batelli, G. B., and Stirpe, F.. Ribosome-inactivating proteins from plants. **Biochim. Biophys. Acta** 1154: 237–282, 1993.

Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, Westbrook J. The Protein Data Bank and the challenge of structural genomics. **Nat Struct Biol**. Nov;7 Suppl:957-9, 2000.

Brinda KV, Surolia A, Vishveshwara S. Insights into the quaternary association of proteins through structure graphs: a case study of lectins. **Biochem J**. Oct 1;391(Pt 1):1-15, 2005.

Brinda KV, Mitra N, Surolia A, Vishveshwara S. Determinants of quaternary association in legume lectins. **Protein Sci**. Jul;13(7):1735-49, 2004.

Bouckaert, J., Loris, R., Poortmans, F., Wyns, L. Crystallographic structure of metal-free concanavalin A at 2.5 A resolution. **Proteins** v23 pp.510-524, 1995.

Bourne, Y., Zamboni, V., Barre, A., Peumans, W.J., Van Damme, E.J., Rouge, P. Helianthus tuberosus lectin reveals a widespread scaffold for mannose-binding lectins. **Structure Fold.Des**. v7 pp.1473-1482, 1999.

Bourne, Y., Roig-Zamboni, V., Barre, A., Peumans, W.J., Astoul, C.H., Van Damme, E.J.M., Rouge, P. The crystal structure of the Calystegia sepium agglutinin reveals a novel quaternary arrangement of lectin subunits with a beta-prism fold. **J.Biol.Chem**. v279 pp.527-533 , 2004.

Calvete, J.J. et al. Primary structure and posttranslational processing of Vatairea macrocarpa seed lectin. **J Protein Chem**. Aug;17(6):545-7, 1998.

Cavada BS, Moreno FB, da Rocha BA, de Azevedo WF Jr, Castellon RE, Goersch GV, Nagano CS, de Souza EP, Nascimento KS, Radis-Baptista G, Delatorre P, Leroy Y, Toyama MH, Pinto VP, Sampaio AH, Barettino D, Debray H, Calvete JJ, Sanz L. cDNA cloning and 1.75 A crystal structure determination of PPL2, an endochitinase and N-acetylglucosamine-binding hemagglutinin from Parkia platycephala seeds. **FEBS J**. Sep;273(17):3962-74, 2006.

Chandra, N.R., Ramachandraiah, G., Bachhawat, K., Dam, T.K., Surolia, A., Vijayan, M. Crystal structure of a dimeric mannose-specific agglutinin from garlic: quaternary association and carbohydrate specificity. **J.Mol.Biol**. v285 pp.1157-1168, 1999.

Cheng W, Bullitt E, Bhattacharyya L, Brewer CF, Makowski L. Electron microscopy and x-ray diffraction studies of Lotus tetragonolobus A isolectin cross-linked with a divalent Lewisx oligosaccharide, an oncofetal antigen. **J Biol Chem**. Dec 25;273(52):35016-22, 1998.

Dam T.K. & Brewer C. F. Thermodynamic studies of lectin-carbohydrate interactions by isothermal titration calorimetry. **Chem Rev**. Feb;102(2):387-429, 2002.

Davidsson P, Karlsson B, Svennerholm L. Glycoprotein pattern in human brain tumors studied using lectin binding after sodium dodecyl sulfate-gel electrophoresis and protein blotting. **Brain Res**. Jun 2;412(2):254-60, 1987.

Dwek, R. A. Glycobiology: Toward understanding the function of sugars. **Chemical Reviews,** v.96, n.2, p.683-720, 1996.

Els J. M. Van Damme, Willy J. Peumans, Annick Barre, and Pierre Rougé. Plant Lectins: A Composite of Several Distinct Families of Structurally and Evolutionary Related Proteins with Diverse Biological Roles. **Critical Reviews in Plant Sciences**, 17(6):575–692, 1998.

Finne J, Castori S, Feizi T, Burger MM. Lectin-resistant variants and revertants of mouse melanoma cells: differential expression of a fucosylated cell-surface antigen and altered metastasizing capacity. **Int J Cancer**. Feb 15;43(2):300-4, 1989.

Fujii, T., Hayashida, M., Hamasu, M., Ishiguro, M., Hata, Y. Structures of two lectins from the roots of pokeweed (Phytolacca americana). **Acta Crystallogr.,Sect.D**. v60 pp.665-673, 2004.

Gabius H. –J. & Gabius S., eds Glycoscience Status and Perspectives. **Chapman & Hall**, Weinheim, Germany, 1997.

Gallego Del Sol, F., Nagano, C., Cavada, B.S., Calvete, J.J. The first crystal structure of a mimosoideae lectin reveals a novel quaternary arrangement of a widespread domain. **J.Mol.Biol**. v353 pp.574-583, 2005.

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

Griffiths DW, Stephenson TJ. Lectin agglutinins: a comparative study of Lotus tetragonolobus and Ulex europaeus 1 as markers of vascular endothelium in follicular thyroid carcinoma. **Med Lab Sci**. Jan;45(1):45-51, 1988.

H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne: The Protein Data Bank. **Nucleic Acids Research**, 28 pp. 235-242, 2000.

Harata, K., Schubert, W.D., Muraki, M. Structure of Urtica dioica agglutinin isolectin I: dimer formation mediated by two zinc ions bound at the sugar-binding site. **Acta Crystallogr, Sect.D**. v57 pp.1513-1517, 2001.

Harata K, Muraki M.Crystal structures of Urtica dioica agglutinin and its complex with tri-N-acetylchitotriose. **J Mol Biol**. Mar 31;297(3):673-81, 2000

Harata, K., Nagahora, H., Jigami, Y. X-Ray Structure Of Wheat-Germ-Agglutinin Isolectin-3. **Acta Crystallogr D Biol Crystallogr**. v51 pp.1013-1019, 1995

Hashimoto, K. et al. Kegg as a glycome informatics resource. **Glycobiology**, v.16, n.5, p.63R-70R, 2006.

Hayashida M, Fujii T, Hamasu M, Ishiguro M, Hata Y.Crystallization and preliminary X-ray analysis of lectin C from the roots of pokeweed (Phytolacca americana). **Acta Crystallogr D Biol Crystallogr**. Jul;59(Pt 7):1249-52. Jun 27, 2003.

Hester, G., Wright, C.S. The mannose-specific bulb lectin from Galanthus nivalis (snowdrop) binds mono- and dimannosides at distinct sites. Structure analysis of refined complexes at 2.3 A and 3.0 A resolution. **J.Mol.Biol**. v262 pp.516-531, 1996.

Hester, G., Kaku, H., Goldstein, I.J., Wright, C.S. Structure of mannose-specific snowdrop (Galanthus nivalis) lectin is representative of a new plant lectin family. **Nat.Struct.Biol**. v2 pp.472-479, 1995.

Holthofer H, Miettinen A, Paasivuo R, Lehto VP, Linder E, Alfthan O, Virtanen I. Cellular origin and differentiation of renal carcinomas. A fluorescence

microscopic study with kidney-specific antibodies, antiintermediate filament antibodies, and lectins. **Lab Invest**. Sep;49(3):317-26, 1983.

Imura J, Ichikawa K, Takeda J, Tomita S, Yamamoto H, Nakazono M, Takimoto T, Ueda Y, Fujimori T. Multilocular cystic renal cell carcinoma: a clinicopathological, immuno- and lectin histochemical study of nine cases. **APMIS**. Mar;112(3):183-91, 2004.

Jeyaprakash, A.A., Srivastav, A., Surolia, A., Vijayan, M. Structural basis for the carbohydrate specificities of artocarpin: variation in the length of a loop as a strategy for generating ligand specificity. **J.Mol.Biol**. v338 pp.757-770, 2004.

Kalb AJ. The separation of three L-fucose-binding proteins of Lotus tetragonolobus. **Biochim Biophys Acta**. Dec 3;168(3):532-6, 1968.

Koch B, Regnat W, Schedel I, Hermanek H, Leibold W, Kalden JR. The binding of different lectins on peripheral blood mononuclear cells from patients with chronic inflammatory and malignant diseases. **Immunobiology**. Mar;164(2):99-109, 1983.

Kocourek, J. and Horejsi, V. A note on the recent discussion on definition of the term 'lectin'. **In**: **Lectins: Biology, Biochemistry and Clinical Biochemistry**. Vol. 3, pp. 3–6. Bøg- Hansen, T. C. and Spengler, G. A., Eds., Walter de Gruyter, Berlin, Germany, 1983.

Laurence R. Olsen, Andre´a Dessen,Dipti Gupta, Subramaniam Sabesan, O James C. Sacchettini and C. Fred Brewer. X-ray Crystallographic Studies of Unique Cross-Linked Lattices between Four Isomeric Biantennary Oligosaccharides and Soybean Agglutinin. 15073. **Biochemistry**, 36, 15073-15080, 1997.

Lee, X., Thompson, A., Zhang, Z., Ton-that, H., Biesterfeldt, J., Ogata, C., Xu, L., Johnston, R.A., Young, N.M. Structure of the complex of Maclura

pomifera agglutinin and the T-antigen disaccharide, Galbeta1,3GalNAc. **J.Biol.Chem**. v273 pp.6312-6318, 1998.

Lis H, Sharon N. Lectins: Carbohydrate-Specific Proteins That Mediate Cellular Recognition. **Chem Rev**. Apr 2;98(2):637-674, 1998.

Liu, W., Yang, N., Ding, J., Huang, R.H., Hu, Z., Wang, D.C. Structural Mechanism Governing the Quaternary Organization of Monocot Mannose-binding Lectin Revealed by the Novel Monomeric Structure of an Orchid Lectin **J.Biol.Chem**. v280 pp.14865-14876, 2005.

Mansour MH, Abdul-Salam F, Al-Shemary T. Distinct binding patterns of fucose-specific lectins from Biomphalaria alexandrina and Lotus tetragonolobus to murine lymphocyte subsets. **Immunobiology**.;210(5):335-48, 2005.

Martins, J.C., Maes, D., Loris, R., Pepermans, H.A., Wyns, L., Willem, R., Verheyden, P. H NMR study of the solution structure of Ac-AMP2, a sugar binding antimicrobial protein isolated from Amaranthus caudatus. **J.Mol.Biol**. v258 pp.322-333, 1996.

Meagher JL, Winter HC, Ezell P, Goldstein IJ, Stuckey JA. Crystal structure of banana lectin reveals a novel second sugar binding site. ***Glycobiology***. Oct;15(10):1033-42. Epub 2005 Jun 8, 2005.

Montisci MJ, Dembri A, Giovannuci G, Chacun H, Duchene D, Ponchel G. Gastrointestinal transit and mucoadhesion of colloidal suspensions of Lycopersicon esculentum L. and Lotus tetragonolobus lectin-PLA microsphere conjugates in rats. **Pharm Res**. Jun;18(6):829-37, 2001.

Mcluskey, K., Cameron, S., Hammerschmidt, F., Hunter, W.N. Structure and Reactivity of Hydroxypropylphosphonic Acid Epoxidase in Fosfomycin Biosynthesis by a Cation- and Flavin-Dependent Mechanism. **Proc.Natl.Acad.Sci**. USA v102 pp.14221, 2005.

Napier PW, Everhart DL, Grundbacher FJ. Characterization of a human saliva antigen precipitated by a lectin from Lotus tetragonolobus. **Vox Sang**.;27(5):447-58, 1974.

Pascal, J.M., Day, P.J., Monzingo, A.F., Ernst, S.R., Robertus, J.D., Iglesias, R., Perez, Y., Ferreras, J.M., Citores, L., Girbes, T. 2.8-A crystal structure of a nontoxic type-II ribosome-inactivating protein, ebulin l. **Proteins** v43 pp.319-326, 2001.

Pereira ME, Kabat EA. Blood group specificity of the lectin from Lotus tetragonolobus. **Ann N Y Acad Sci**.; 234(0):301-5, 1974.

Peumans WJ, Zhang W, Barre A, Houles Astoul C, Balint-Kurti PJ, Rovira P, Rouge P, May GD, Van Leuven F, Truffa-Bachi P, Van Damme EJ. Fruit-specific lectins from banana and plantain. **Planta**. Sep;211(4):546-54, 2000.

Peumans, W. J. and Van Damme, E. J. M.. Lectins as plant defense proteins. **Plant Physiol**. **109**: 347–352, 1995b.

Pratap, J.V., Jeyaprakash, A.A., Rani, P.G., Sekar, K., Surolia, A., Vijayan, M. Crystal structures of artocarpin, a Moraceae lectin with mannose specificity, and its complex with methyl-alpha-D-mannose: implications to the generation of carbohydrate specificity. **J.Mol.Biol**. v317 pp.237-247, 2002.

Rabijns, A., Barre, A., Van Damme, E.J.M., Peumans, W.J., De Ranter, C.J., Rouge, P. Structural analysis of the jacalin-related lectin MornigaM from the black mulberry (Morus nigra) in complex with mannose. **Febs J.** v272 pp.3725-3732, 2005.

Raedler A, Schmiegel WH, Raedler E, Arndt R, Thiele HG. Lectin-defined cell surface glycoconjugates of pancreatic cancer cells and their nonmalignant counterparts. **Exp Cell Biol**.;51(1):19-28, 1983.

Raikhel, N. V., Lee, H.-I., and Broekaert, W. F.. Structure and function of chitin-binding proteins. **Annu. Rev. Plant Physiol. Plant Mol. Biol**. 44: 591–615, 1993.

Raju GC, Lee YS. Lectin expression in carcinoid tumours of the gastrointestinal tract. **J Pathol**. Mar;154(3):263, 1988.

Rao, K.N., Suresh, C.G., Katre, U.V., Gaikwad, S.M., Khan, M.I. Two orthorhombic crystal structures of a galactose-specific lectin from Artocarpus hirsuta in complex with methyl-D-galactose. **Acta Crystallogr.,Sect.D.** v60 pp.1404-1412*,* 2004.

Ramachandraiah, G., Chandra, N.R., Surolia, A., Vijayan, M. Re-refinement using reprocessed data to improve the quality of the structure: a case study involving garlic lectin. **Acta Crystallogr., Sect.D.** v58 pp.414-420*,* 2002.

Ravishankar, R., Thomas, C.J., Suguna, K., Surolia, A., Vijayan, M. Crystal structures of the peanut lectin-lactose complex at acidic pH: retention of unusual quaternary structure, empty and carbohydrate bound combining sites, molecular mimicry and crystal packing directed by interactions at the combining site. **Proteins** v43 pp.260-270, 2001.

Rutenber, E., Katzin, B.J., Ernst, S., Collins, E.J., Mlsna, D., Ready, M.P., Robertus, J.D. Crystallographic refinement of ricin to 2.5 A. **Proteins**. v10 pp.240-250, 1991.

Sagan Z, Olszewska I, Pawlus M. Research on phytoagglutinin from the seeds of Lotus tetragonolobus L. and some other plants. **Acta Pol Pharm**; 20:83-8, 1963.

Sanz-Aparicio, J., Hermoso, J., Grangeiro, T.B., Calvete, J.J., Cavada, B.S. The crystal structure of Canavalia brasiliensis lectin suggests a correlation between its quaternary conformation and its distinct biological properties from Concanavalin A. **FEBS Lett**. v405 pp.114-118, 1997.

Sarkar, M., Wu, A. M., and Kabat, E. A.. Immunochemical studies on the carbohydrate specificity of Maclura pomifera lectin. **Arch. Biochem. Biophys**. 209: 204–218, 1981.

Sastry, M. V. K., Banerjee, P., Patanjali, S. R., Swamy, M. J., Swarnalatha, G. V., and Surolia, A.. Analysis of the saccharide binding to Artocarpus integrifolia lectin reveals specific recognition of T-antigen (b-D Gal(1,3)DGalNAc). **J. Biol. Chem**. 261: 11726–11733, 1986.

Sauerborn, M.K., Wright, L.M., Reynolds, C.D., Grossmann, J.G., Rizkallah, P.J. Insights into carbohydrate recognition by Narcissus pseudonarcissus lectin: the crystal structure at 2 A resolution in complex with alpha1-3 mannobiose. **J.Mol.Biol**. v290 pp.185-199*,* 1999.

Saul, F.A., Rovira, P., Boulot, G., Damme, E.J., Peumans, W.J., Truffa-Bachi, P., Bentley, G.A. Crystal structure of Urtica dioica agglutinin, a superantigen presented by MHC molecules of class I and class II. **Structure Fold.Des**. v8 pp.593-603*,* 2000*.*

Shaanan B, Lis H, Sharon N. Structure of a legume lectin with an ordered N-linked carbohydrate in complex with lactose. **Science**. Nov 8;254(5033):862-6, 1991.

Sharon, N. & Lis H.. Carbohydrates in cell recognition. **Sci. Am**. *268*:82-88, 1993.

Sharon, N. & Lis H.. Lectins-proteins with a sweet tooth: Functions in cell recognition. **Essays Biochem** 30:59-75, 1995.

Singh, D.D., Saikrishnan, K., Kumar, P., Surolia, A., Sekar, K., Vijayan, M. Unusual sugar specificity of banana lectin from Musa paradisiaca and its probable evolutionary origin. Crystallographic and modelling studies **Glycobiology** v15 pp.1025-1032*,* 2005.

Staudacher E, Altmann F, Wilson IB, Marz L. Fucose in N-glycans: from plant to man. **Biochim Biophys Acta**. Dec 6;1473(1):216-36, 1999.

Stillmark, H.. Über Ricin ein giftiges Ferment aus den Samen von Ricinus communis L. und einige anderen Euphorbiaceen. Inaugural Dissertation Dorpat, Tartu, 1888.

Summer, J. B. and Howell, S. F.. The identification of the hemagglutinin of the Jack bean with concanavalin A. **J. Bacteriol**. 32: 227–237. 86: 922–926, 1936.

Tahirov, T.H., Lu, T.H., Liaw, Y.C., Chen, Y.L., Lin, J.Y. Crystal structure of abrin-a at 2.14 A. **J.Mol.Biol**. v250 pp.354-367, 1995.

Thomas CJ, Surolia A. Mode of molecular recognition of L-fucose by fucose-binding legume lectins. Biochem Biophys Res Commun. 2000 Feb 16;268(2):262-7. Erratum in: **Biochem Biophys Res Commun.** Apr 2;270(1):329, 2000.

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**. v4 pp.779-783, 1997.

Turner GA, Goodarzi MT, Thompson S. Glycosylation of alpha-1-proteinase inhibitor and haptoglobin in ovarian cancer: evidence for two different mechanisms. **Glycoconj J**. Jun;12(3):211-8, 1995.

Turton K, Natesh R, Thiyagarajan N, Chaddock JA, Acharya KR. Crystal structures of Erythrina cristagalli lectin with bound N-linked oligosaccharide and lactose. **Glycobiology**. Oct;14(10):923-9, 2004.

Ulrich W, Horvat R, Krisch K. Lectin histochemistry of kidney tumours and its pathomorphological relevance. **Histopathology**. Oct;9(10):1037-50, 1985.

Van Damme EJ, Barre A, Mazard AM, Verhaert P, Horman A, Debray H, Rouge P, Peumans WJ. Characterization and molecular cloning of the lectin from Helianthus tuberosus. **Eur J Biochem**. Jan;259(1-2):135-42, 1999.

Van Damme, E. J. M., Barre, A., Verhaert, P., Rougé, P., and Peumans, W. J.. Molecular cloning of the mitogenic mannose/maltosespecific rhizome lectin from Calystegia sepium. **FEBS Lett**. 397: 352–356, 1996e.

VanEpps DE, Tung KS. Fucose-binding Lotus tetragonolobus lectin binds to human polymorphonuclear leukocytes and induces a chemotactic response. **J Immunol**. Sep;119(3):1187-9, 1977.

Walker RA. The binding of peroxidase-labelled lectins to human breast epithelium. III--Altered fucose-binding patterns of breast carcinomas and their significance. **J Pathol**. Oct;144(2):109-17, 1984.

Wright, C.S., Hester, G. The 2.0 A structure of a cross-linked complex between snowdrop lectin and a branched mannopentaose: evidence for two unique binding modes. **Structure** v4 pp.1339-1352, 1996.

Wright, C.S., Jaeger, J. Crystallographic refinement and structure analysis of the complex of wheat germ agglutinin with a bivalent sialoglycopeptide from glycophorin A. **J.Mol.Biol**. v232 pp.620-638, 1993.

Wright, C.S. 2.2 A resolution structure analysis of two refined N-acetylneuraminyl-lactose--wheat germ agglutinin isolectin complexes. **J.Mol.Biol**. v215 pp.635-651, 1990.

Wright, L.M., Reynolds, C.D., Rizkallah, P.J., Allen, A.K., Van Damme, E.J., Donovan, M.J., Peumans, W.J. Structural characterisation of the native fetuin-binding protein Scilla campanulata agglutinin: a novel two-domain lectin. **FEBS Lett**. v468 pp.19-22, 2000.

Wright, C. S.. The crystal structure of wheat germ agglutinin at 2.2 Å resolution. **J. Mol. Biol**. 111: 439–457, 1977

Watkins, W. M. and Morgan, W. T. J.. Neutralization of the anti-H agglutinin in eel by simple sugars. **Nature.** 169: 825–826, 1952.

Wood, S.D., Wright, L.M., Reynolds, C.D., Rizkallah, P.J., Allen, A.K., Peumans, W.J., Van Damme, E.J.M. Structure of the Native (unligated) Mannose-Specific Bulb Lectin from Scilla Campanulata (bluebell) at 1.7A Resolution. **Acta Crystallogr. Sect.D**. vD55 pp.1264-1272, 1999.

Yariv J, Kalb AJ, Katchalski E. Isolation of an L-fucose binding protein from Lotus tetragonolobus seed. **Nature**. Aug 19;215(5103):890-1, 1967.

Zhang W, Peumans WJ, Barre A, Astoul CH, Rovira P, Rouge P, Proost P, Truffa-Bachi P, Jalali AA, Van Damme EJ. Isolation and characterization of a jacalin-related mannose-binding lectin from salt-stressed rice (Oryza sativa) plants. **Planta**. May;210(6):970-8, 2000.

11.0 Anexo 01 – Artigos publicados relacionados

Artigo 01

Moreno FB, Bezerra GA, de Oliveira TM, de Souza EP, da Rocha BA, Benevides RG, Delatorre P, Cavada BS, de Azevedo WF Jr. New crystal forms of Diocleinae lectins in the presence of different dimannosides. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2006 Nov 1;62(Pt 11):1100-3. Epub 2006 Oct 20.

Artigo 02

Gadelha CA, **Moreno FB**, Santi-Gadelha T, Cajazeiras JB, Rocha BA, Rustiguel JK, Freitas BT, Canduri F, Delatorre P, Azevedo WF Jr, Cavada BS. Crystallization and preliminary X-ray diffraction analysis of a lectin from *Canavalia maritima* seeds. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2005 Jan 1;61(Pt 1):87-9. Epub 2004 Dec 2.

Artigo 03

Cavada BS, Castellón RE, Vasconcelos GG, Rocha BA, Bezerra GA, Debray H, Delatorre P, Nagano CS, Toyama M, Pinto VP, **Moreno FB**, Canduri F, Azevedo WF Jr. Crystallization and preliminary X-ray diffraction analysis of a new chitin-binding protein from *Parkia platycephala* seeds. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2005 Sep 1;61(Pt 9):841-3. Epub 2005 Aug 31.

Artigo 04

Cavada BS, Marinho ES, Souza EP, Benevides RG, Delatorre P, Souza LA, Nascimento KS, Sampaio AH, **Moreno FB**, Rustiguel JK, Canduri F, de Azevedo WF Jr, Debray H. Purification, partial characterization and preliminary X-ray diffraction analysis of a mannose-specific lectin from *Cymbosema roseum* seeds. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2006 Mar 1;62 (Pt 3):235-7. Epub 2006 Feb 10.

Artigo 05

Bezerra GA, Oliveira TM, **Moreno FB**, de Souza EP, da Rocha BA, Benevides RG, Delatorre P, de Azevedo WF Jr, Cavada BS. 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. 2007 Nov;160(2):168-76. Epub 2007 Aug 16.

Artigo 06

Cavada BS, **Moreno FB**, da Rocha BA, de Azevedo WF Jr, Castellón RE, Goersch GV, Nagano CS, de Souza EP, Nascimento KS, Radis-Baptista G, Delatorre P, Leroy Y, Toyama MH, Pinto VP, Sampaio AH, Barettino D, Debray H, Calvete JJ, Sanz L. cDNA cloning and 1.75 A crystal structure determination of PPL2, an endochitinase and N-acetylglucosamine-binding hemagglutinin from *Parkia platycephala* seeds. FEBS J. 2006 Sep;273(17):3962-74.

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

New crystal forms of Diocleinae lectins in the presence of different dimannosides

Frederico Bruno Mendes Batista Moreno,[a] Gustavo Arruda Bezerra,[b] Taianá Maia de Oliveira,[b] Emmanuel Prata de Souza,[b] Bruno Anderson Matias da Rocha,[b,c] Raquel Guimarães Benevides,[b] Plínio Delatorre,[c] Benildo Sousa Cavada[b]* and Walter Filgueira de Azevedo Jr[d]*

[a]Departamento de Física–IBILCE–Universidade Estadual Paulista, São José do Rio Preto 15054-000, Brazil, [b]Biomol-Lab, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza-CE, Caixa Postal 6043, CEP 60455-970, Brazil, [c]Departamento de Ciências Biológicas, Universidade Regional do Cariri, Crato, CE 63195-000, Brazil, and [d]Faculdade de Biociências–PUCRS, Av. Ipiranga 6681, TECNOPUC-Predio 92[a], 90619-900 Porto Alegre, RS, Brazil

Correspondence e-mail: bscavada@ufc.br, walter.junior@pucrs.br

Received 21 July 2006
Accepted 22 September 2006

Studying the interactions between lectins and sugars is important in order to explain the differences observed in the biological activities presented by the highly similar proteins of the Diocleinae subtribe. Here, the crystallization and preliminary X-ray data of *Canavalia gladiata* lectin (CGL) and *C. maritima* lectin (CML) complexed with Man(α1-2)Man(α1)OMe, Man(α1-3)Man(α1)-OMe and Man(α1-4)Man(α1)OMe in two crystal forms [the complexes with Man(α1-3)Man(α1)OMe and Man(α1-4)Man(α1)OMe crystallized in space group $P3_2$ and those with Man(α1-2)Man(α1)OMe crystallized in space group $I222$], which differed from those of the native proteins ($P2_12_12$ for CML and $C222$ for CGL), are reported. The crystal complexes of ConA-like lectins with Man(α1-4)Man(α1)OMe are reported here for the first time.

1. Introduction

Lectins are carbohydrate-binding proteins which function as cognate receptors for various cell-surface glycoproteins, resulting in several important cellular-mediated events ranging from mitogenic processes to plant defence mechanisms (Weis & Drickamer, 1996).

Plant lectins, especially those purified from species of the Leguminosae family, represent the best studied group of carbohydrate-binding proteins (Van Damme *et al.*, 1998). Legume lectins have been used for decades as a model system for studying protein–sugar interactions. Fundamental insights obtained from investigating these protein model systems can often be readily applied to lectins outside this family, such as the pharmaceutically important C-type lectins and galectins (Hamelryck *et al.*, 1998).

The legume lectins from the Diocleinae subtribe are highly similar proteins that present significant differences in the potency/efficacy of their biological activities (Delatorre *et al.*, 2006). For instance, even though they only differ by three amino-acid residues, *Canavalia gladiata* lectin (CGL) and *C. maritima* lectin (CML) present different patterns in several activities, such as in aorta-relaxation experiments (to be published). In spite of these activities being very well documented, little is known about the receptors with which they interact and how this happens. The crystal structures of many plant lectins have revealed how they specifically recognize their carbohydrate ligands (Lis & Sharon, 1998) and this may give us a direction for elucidating their mechanism of action.

An outstanding feature of Diocleinae lectins is that although all the monomers have similar tertiary structures, they show different modes of quaternary association. Minor differences in the ratio between the dimeric and tetrameric forms, together with changes in the relative orientations of the carbohydrate-binding sites in the quaternary structures, have been hypothesized to contribute to the differing biological activities possessed by these lectins (Brinda *et al.*, 2004).

Native CGL and CML have been crystallized and their structures have been solved in space groups $C222$ and $P2_12_12$, respectively (Moreno *et al.*, 2004; Gadelha *et al.*, 2005). Both have tetramers composed of 237 amino-acid monomers in the asymmetric unit. Crystals of CML in complex with trehalose and maltose were subsequently obtained under the same conditions as described for the native protein and presented the same space group $P2_12_12$ (Delatorre *et al.*, 2006).

We report here the crystallization and preliminary data of six dimannoside-complexed crystals of CGL and CML in two different space groups, including the first crystal of a ConA-like lectin with the dimannoside Man(α1-4)Man(α1)OMe.

2. Material and methods

2.1. Crystallization

CML and CGL were purified according to the methods of Moreira & Cavada (1984) and Ceccatto *et al.* (2002), respectively. Each dimannoside [Man(α1-2)Man(α1)OMe, Man(α1-3)Man(α1)OMe and Man(α1-4)Man(α1)OMe; purchased from Sigma–Aldrich] was added in a 0.08 molar proportion to 10 mg ml^{-1} of the pure lectin dissolved in 20 m*M* Tris–HCl pH 7.6 containing 5 m*M* $CaCl_2$/$MnCl_2$ and incubated for 24 h at room temperature.

Crystallization trials were performed based on two different strategies: (i) varying the ammonium sulfate concentration from 0.25 to 3.0 *M* combined with different buffers varying from pH 3.5 to 9.0 (strategy supplied by McPherson, 2003) and (ii) varying the sodium formate concentration from 0.5 to 7.0 *M* with Tris–HCl pH 7.6. Both strategies were carried out at 293 K using the hanging-drop vapour-diffusion method in Linbro plates. The drops were composed of equal volumes (2 µl) of protein solution and reservoir solution and were equilibrated against 500 µl of the latter.

Table 1
Crystallization conditions, space groups and number of molecules per asymmetric unit for native CGL and CML and for their complexes with Man(α1-2)Man(α1)OMe, Man(α1-3)Man(α1)OMe and Man(α1-4)Man(α1)OMe.

	Crystallization conditions	Space group	Molecules per ASU
Native CGL	0.1 *M* Tris–HCl pH 8.5, 2 *M* ammonium sulfate	*C*222	4
CGL–Man(α1-2)Man(α1)OMe	0.1 *M* Tris–HCl pH 8.0–9.0, 1.8–2.6 *M* ammonium sulfate	*I*222	1
CGL–Man(α1-3)Man(α1)OMe	4.5–6.5 *M* sodium formate	$P3_2$	4
CGL–Man(α1-4)Man(α1)OMe	4.5–6.5 *M* sodium formate	$P3_2$	4
Native CML	0.1 *M* Na HEPES pH 8.48, 4% PEG 400, 2*M* ammonium sulfate	$P2_12_12$	4
CML–Man(α1-2)Man(α1)OMe	0.1 *M* Tris–HCl pH 8.0–9.0, 1.8–2.6 *M* ammonium sulfate	*I*222	1
CML–Man(α1-3)Man(α1)OMe	4.5–6.5 *M* sodium formate	$P3_2$	4
CML–Man(α1-4)Man(α1)OMe	4.5–6.5 *M* sodium formate	$P3_2$	4

2.2. Data collection and processing

X-ray diffraction data were collected at a wavelength of 1.43 Å using a synchrotron-radiation source (MX1 station, Laboratório

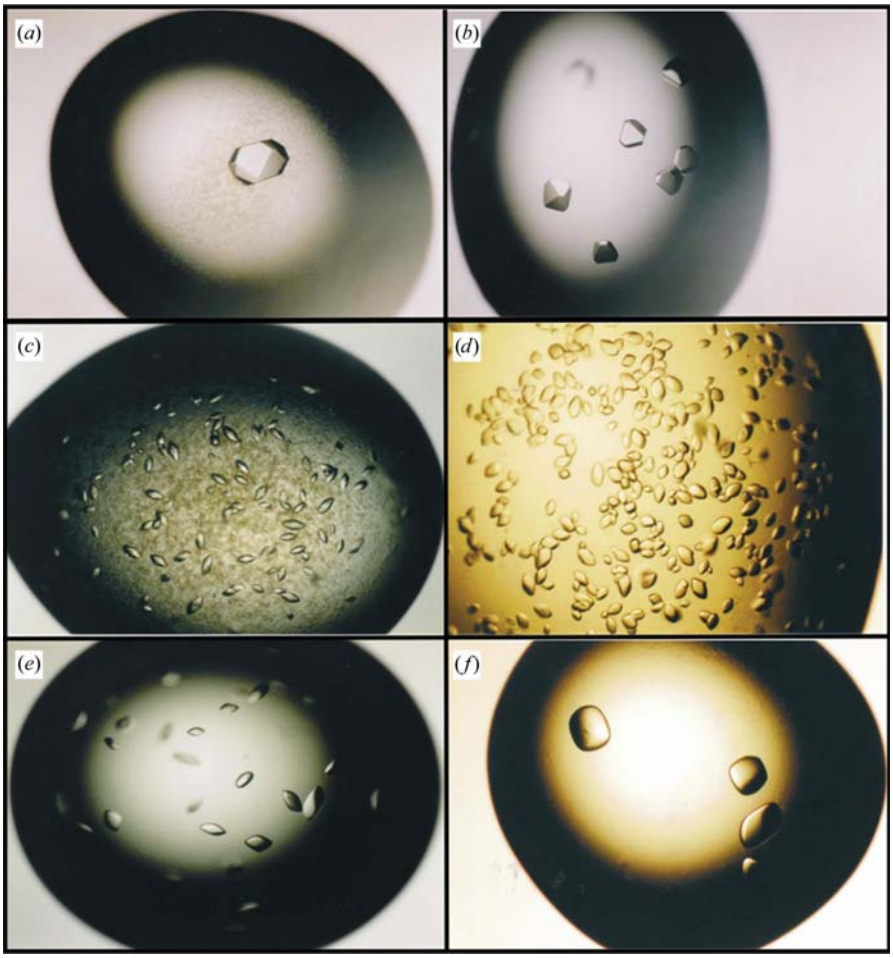


Figure 1
Crystals of (*a*) CGL–Man(α1-2)Man(α1)OMe, (*b*) CML–Man(α1-2)Man(α1)OMe, (*c*) CGL–Man(α1-3)Man(α1)OMe, (*d*) CML–Man(α1-3)Man(α1)OMe, (*e*) CGL–Man(α1-4)Man(α1)OMe and (*f*) CML–Man(α1-4)Man(α1)OMe.

Table 2
Statistics of data collection for CGL complexes.

Values in parentheses are for the highest resolution shell.

Data collection	CGL–Man(α1-2)Man(α1)OMe	CGL–Man(α1-3)Man(α1)OMe	CGL–Man(α1-4)Man(α1)OMe
R_{merge}	6.1 (32.8)	8.8 (40)	6.8 (23.3)
Resolution limit (Å)	21.03–1.50	40.29–2.07	40.096–1.980
$I/\sigma(I)$	7.6 (2.0)	10.7 (2.0)	8.8 (3.2)
Completeness (%)	99.9 (100.0)	91.5 (86)	99.4 (99.4)
Redundancy	19.20	2.5	10.20
Unit-cell parameters (Å)	a = 63.89, b = 86.19, c = 88.73	a = 69.37, b = 69.37, c = 161.21	a = 69.01, b = 69.01, c = 160.44
Matthews coefficient (Å^3 Da^{-1})	2.4	2.2	2.2
Solvent content (%)	48.2	43.6	42.5
Space group	$I222$	$P3_2$	$P3_2$
Wavelength (Å)	1.431	1.431	1.431
Total No. of reflections	760001	119394	607712
Total No. of unique observations	39568	48437	59558

Table 3
Statistics of data collection for CML complexes.

Values in parentheses are for the highest resolution shell.

Data collection	CML–Man(α1-2)Man(α1)OMe	CML–Man(α1-3)Man(α1)OMe	CML–Man(α1-4)Man(α1)OMe
R_{merge}	5.5 (45.2)	9.1 (29.6)	9.9 (34.8)
Resolution limit (Å)	20.96–1.4	40.49–2.10	34.73–2.10
$I/\sigma(I)$	15.5 (2.4)	15.8 (3.0)	15.6 (3.9)
Completeness (%)	97.9 (95.43)	98.9 (98.1)	96.5 (94.8)
Redundancy	4.0 (3.9)	4.2 (4.0)	5.2 (5.1)
Unit-cell parameters (Å)	a = 63.63, b = 85.44, c = 85.71	a = 69.39, b = 69.39, c = 161.29	a = 69.47, b = 69.47, c = 161.52
Mathews coefficient (Å^3 Da^{-1})	2.4	2.2	2.2
Solvent content (%)	48.5	43.6	43.8
Space group	$I222$	$P3_2$	$P3_2$
Wavelength (Å)	1.431	1.431	1.431
Total No. of reflections	183774	212393	252952
Total No. of unique observations	46457	50169	49112

Nacional de Luz Síncrotron, Campinas, Brazil) and a CCD detector (MAR Research) at 100 K. To avoid freezing, crystals were soaked in a cryoprotectant solution containing between 30 and 50% glycerol in mother liquor. Data were processed, indexed and integrated using *MOSFLM* and scaled using *SCALA* (Collaborative Computational Project, Number 4, 1994; Leslie, 1992).

3. Results and discussion

Crystallization conditions, space groups and the number of molecules per asymmetric unit of the dimannoside complexes and native crystals are listed in Table 1. The preliminary crystallographic data for CGL complexes is presented in Table 2 and for CML complexes in Table 3. The crystal complexes are depicted in Fig. 1.

Although the crystallization conditions are different for the crystal complexes and the native proteins, the main cause of the change in the symmetry seems to be the presence of the dimannosides. CGL and CML complexed with Man(α1-2)Man(α1)OMe crystallized in space group $I222$ with one molecule in the asymmetric unit after 72 h, while the crystals of CGL and CML complexed with Man(α1-3)Man(α1)OMe and Man(α1-4)Man(α1)OMe belonged to space group $P3_2$ with a tetramer in the asymmetric unit, crystallizing after 48 h.

The importance of the dimannosides in the crystallization process is noteworthy. Native CGL and CML crystallize under different conditions and in different space groups: 0.1 *M* Tris–HCl pH 8.5, 2.0 *M* ammonium sulfate ($C222$) and 0.1 *M* Na HEPES pH 8.48, 4% PEG 400, 2.0 *M* ammonium sulfate ($P2_12_12$), respectively. In the presence of the carbohydrates, the crystallization conditions become the same for each sugar: for CML and CGL with Man(α1-2)Man(α1)OMe the condition was 0.1 *M* Tris–HCl pH 8.0–9.0, 1.8–2.6 *M* ammonium sulfate, while crystals of CML and CGL complexed with Man(α1-3)Man(α1)OMe and Man(α1-4)Man(α1)OMe only grew in the presence of 4.5–6.5 *M* sodium formate.

It is well established that legume lectins possess three types of hydrophobic sites based on different ligand affinities (Sharon & Lis, 1990). One of these sites is adjacent to the monosaccharide-binding site and participates in interactions involving several hydrophobic sugars. Bouckaert and coworkers described that the O3-linked mannose of Man(α1-3)Man(α1)OMe and the O6-linked mannose of Man(α1-6)Man(α1)OMe bind to the hydrophobic subsite formed by Tyr12, Tyr100 and Leu99 (Bouckaert *et al.*, 1999).

Investigations of the binding of Man(α1-2)Man(α1)OMe, Man(α1-3)Man(α1)OMe and Man(α1-6)Man(α1)OMe to ConA in this same hydrophobic subsite revealed significant differences in their affinity (Moothoo *et al.*, 1999). Based on this, we have crystallized CML and CGL complexed with the dimannosides Man(α1-2)Man(α1)OMe and Man(α1-3)Man(α1)OMe in order to compare them with the previously reported complexes from ConA and correlate their structure and affinity. The crystal complexes of CGL and CML in the presence of Man(α1-4)Man(α1)OMe represent the first ConA-like structure with this carbohydrate. The differences between the affinities of these mannosides may reflect how the protein binds to receptors related to lectin-mediated responses in plants or in other organisms. Therefore, solving the structures of CGL and CML complexed with dimannosides may be important to understanding many of their biological activities. Our data may clarify the understanding of how the interactions between the dimannosides and the hydrophobic subsite formed by Tyr12, Tyr100 and Leu99 occur.

The complex crystals were obtained in space groups that differed from those of the native lectins. Since crystal packing has an influence on the protein conformation (Kanellopoulos *et al.*, 1996), our work may be important in revealing interactions distinct from those in the native structures.

This work was supported by Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do estado de São Paulo (FAPESP), Universidade Regional do Cariri (URCA), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and National Synchrotron Light Laboratory (LNLS), Brazil. BSC and WFA are senior investigators of CNPq.

References

Bouckaert, J., Hamelryck, T. W., Wyns, L. & Loris, R. (1999). *J. Biol. Chem.* **274**, 29188–29195.

Brinda, K. V., Mitra, N., Surolia, A. & Vishveshwara, S. (2004). *Protein Sci.* **13**, 1735–1749.

Ceccatto, V. M, Cavada, B. S., Nunes, E. P., Nogueira, N. A., Grangeiro, M. B., Moreno, F. B., Teixeira, E. H., Sampaio, A. H., Alves, M. A., Ramos, M. V., Calvete, J. J. & Grangeiro, T. B. (2002). *Protein Pept. Lett.* **9**, 67–73.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Delatorre, P., Nascimento, K. S., Melo, L. M., de Souza, E. P., da Rocha, B. A. M., Benevides, R. G., de Oliveira, T. M., Bezerra, G. A., Bezerra, M. J. L. B., da Cunha, R. M. S., Freire, V. N. & Cavada, B. S. (2006). *Acta Cryst.* F**62**, 166–168.

Gadelha, C. A. A., Moreno, F. B. M. B., Santi-Gadelha, T., Cajazeiras, J. B., da Rocha, B. A. M., Rustiguel, J. K. R., Freitas, B. T., Canduri, F., Delatorre, P., de Azevedo, W. F. Jr & Cavada, B. S. (2005). *Acta Cryst.* F**61**, 87–89.

Hamelryck, T. W, Loris, R., Bouckaert, J. & Lode, W. (1998). *Trends Glycosci. Glycotechnol.* **10**, 349–360.

Kanellopoulos, P. N., Tucker, P. A., Pavlou, K., Agianian, B. & Hamodrakas, S. J. (1996). *J. Struct. Biol.* **117**, 16–23.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Lis, H. & Sharon, N. (1998). *Chem. Rev.* **98**, 637–674.

McPherson, A. (2003). *Crystallization of Biological Macromolecules.* New York: Cold Spring Harbor Laboratory Press.

Moothoo, D. N., Canan, B., Field, R. A. & Naismith, J. H. (1999). *Glycobiology*, **9**, 539–545.

Moreira, R. A. & Cavada, B. S. (1984). *Biol. Plant.* **26**, 113–120.

Moreno, F. B. M. B., Delatorre, P., Freitas, B. T., Rocha, B. A. M., Souza, E. P., Faco, F., Canduri, F., Cardoso, A. L. H., Freire, V. N., Lima, J. L., Sampaio, A. H., Calvete, J. J., de Azevedo, W. F. Jr & Cavada, B. S. (2004). *Acta Cryst.* D**60**, 1493–1495.

Sharon, N. & Lis, H. (1990). *FASEB J.* **4**, 3198–3208.

Van Damme, E. J. M., Peumans, W. J., Barre, A. & Rougé, P. (1998). *Crit. Rev. Plant Sci.* **17**, 575–692.

Weis, W. I. & Drickamer, K. (1996). *Annu. Rev. Biochem.* **65**, 441–473.

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

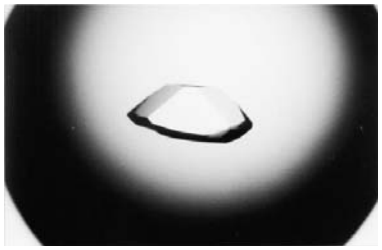
Crystallization and preliminary X-ray diffraction analysis of a lectin from *Canavalia maritima* seeds

Carlos Alberto de Almeida Gadelha,[a] Frederico Bruno Mendes Batista Moreno,[b] Tatiane Santi-Gadelha,[a] João Batista Cajazeiras,[a] Bruno Anderson M. da Rocha,[a] Joane Kathelen Rodrigues Rustiguel,[b] Beatriz Tupinamba Freitas,[a,c] Fernanda Canduri,[b] Plínio Delatorre,[a,c] Walter Filgueira de Azevedo Jr[b]* and Benildo S. Cavada[a]*

[a]BioMol-Lab, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, CE, Caixa Postal 6043, CEP 60455-970, Brazil, [b]Programa de Pós-graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil, and [c]Grupo de Química Biológica, Departamento de Ciências Biológicas, Universidade Regional do Cariri, Crato, CE 63195-000, Brazil

Correspondence e-mail: walterfa@df.ibilce.unesp.br, bscavada@ufc.br

Received 2 September 2004
Accepted 10 November 2004
Online 2 December 2004



A lectin from *Canavalia maritima* seeds (ConM) was purified and submitted to crystallization experiments. The best crystals were obtained using the vapour-diffusion method at a constant temperature of 293 K and grew in 7 d. A complete structural data set was collected to 2.1 Å resolution using a synchrotron-radiation source. The ConM crystal belongs to the orthorhombic space group $P2_12_12$, with unit-cell parameters $a = 67.15$, $b = 70.90$, $c = 97.37$ Å. A molecular-replacement search found a solution with a correlation coefficient of 69.2% and an R factor of 42.5%. Crystallographic refinement is under way.

1. Introduction

Many plants contain sugar-binding proteins commonly known as lectins, designated as carbohydrate-binding proteins of non-immune origin that specifically recognize diverse sugar structures and mediate a variety of biological process (Vijayan & Chandra, 1999).

Plant lectins (Peumans & Van Damme, 1995), especially those purified from species of the Leguminosae family, represent the most well studied group of carbohydrate-binding proteins (Van Damme *et al.*, 1998). Lectins from the Diocleinae subtribe demonstrate a high degree of similarity. Despite being highly analogous, they present significant differences in many biological activities, such as induction of rat paw oedema (Bento *et al.*, 1993), peritoneal macrophage spreading in mouse (Rodriguez *et al.*, 1992), pro- and anti-inflammatory effects (Alencar *et al.*, 1999; Assreuy *et al.*, 1999), capacity for induction of histamine release (Gomes *et al.*, 1994; Ferreira *et al.*, 1996), induction of apoptosis (Barbosa *et al.*, 2001), induction of NO production (Andrade *et al.*, 1999), various renal effects (Havt *et al.*, 2003), mitogenicity (Barral-Neto *et al.*, 1992) and induction of *in vitro* and *in vivo* cytokine production (Cavada *et al.*, 2001).

Despite some minor differences in their primary and three-dimensional structures, it remains clear that this group of proteins diverge considerably in many biological properties, which makes them an excellent model for the study of structure–function relationships (Cavada *et al.*, 2001; Moreno *et al.*, 2004).

The lectin ConM was obtained from *Canavalia maritima*, commonly known as the bay bean, sand bean, beach bean or MacKenzie bean. ConM is a 25.5 kDa protein with 237 residues per monomer. Like other legume lectins, ConM posseses a high amino-acid sequence similarity to the well known concanavalin A (ConA) from *C. ensiformis*, reaching up to 90% identity (Perez *et al.*, 1991).

The present work reports the crystallization and preliminary X-ray diffraction analysis of a lectin from *C. maritima* seeds, a protein that has previously been purified (Perez *et al.*, 1991), tested for histamine-releasing properties in rat peritoneal mast cells (Gomes *et al.*, 1994) and has had its affinity for several monosaccharides determined (Ramos *et al.*, 1996).

2. Materials and methods

2.1. Purification of *C. maritima* seed lectin

Wild mature *C. maritima* seeds were collected in the Ceará state in northeast Brazil. The seeds were ground to a fine powder in a coffee

mill and the soluble proteins were extracted at 298 K by continuous stirring with 0.15 *M* NaCl [1:10(*w*/*v*)] for 1 h, followed by centrifugation at 10 000*g* at 277 K for 20 min. The supernatant was applied onto a Sephadex G-50 column (10 × 50 cm) previously equilibrated with 0.15 *M* NaCl containing 5 m*M* $CaCl_2$ and $MnCl_2$, as described by Cavada *et al.* (1996). The unbound material was eluted with 0.15 *M* NaCl at a flow rate of 45 ml h^{-1} until the absorbance at 280 nm of the effluent stabilized at 0.05. The retained material (a lectin, called ConM) was eluted with 0.1 *M* glycine pH 2.6 containing 0.15 *M* NaCl, dialyzed exhaustively against Milli-Q water and lyophilized. The purity of all ConM preparations was monitored by SDS–PAGE (Laemmli, 1970).

2.2. Crystallization, data collection and processing

ConM was diluted homogeneously to a concentration of 10.0 mg ml^{-1} in 50 m*M* Tris–HCl pH 7.5 contaning 5 m*M* $CaCl_2$ and $MnCl_2$ for all crystallization experiments. Crystallization conditions for ConM were screened using the hanging-drop vapour-diffusion method with Hampton Research Crystal Screens I and II (Hampton Research, Riverside, CA, USA; Jancarik & Kim, 1991) at room temperature (293 K). Microcrystals were obtained using crystallization condition No. 4 of screen I (0.1 *M* Tris–HCl pH 8.5 and 2.0 *M* ammonium sulfate). Improvement of this crystallization condition was obtained by raising the pH and the salt concentration. The best crystals were obtained from drops containing equal volumes of protein (3 µl) and 0.1 *M* Tris–HCl pH 9.0 with 2.2 *M* ammonium sulfate. Crystals grew within a week to maximum dimensions of approximately 0.8 × 0.4 × 0.4 mm (Fig. 1).

X-ray data were collected from a single crystal cooled to a temperature of 100 K. To avoid ice formation, crystals were soaked in a cryoprotectant solution containing 75% 0.1 *M* Tris–HCl pH 9.0 and 25% glycerol and submitted to data collection at a wavelength of 1.4270 Å using a synchrotron-radiation source (CPr station, Laboratório Nacional de Luz Síncrotron-LNLS, Campinas, Brazil). A complete data set was obtained using a CCD (MAR Research) in 120 frames with an oscillation range of 1°. The data set was indexed, integrated and scaled using *MOSFLM* and *SCALA* (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

Several lectins have been crystallized and their structures solved. More than 50 different entries for lectins from the Diocleinae subtribe can be accessed in the Protein Data Bank (Berman *et al.*, 2000); the well known plant lectin ConA represents approximately 90% of these data.

The crystal data were scaled in the range 39.52–2.10 Å and Table 1 shows the data-collection statistics. Assuming the presence of two molecules (474 residues, 25.5 kDa each) in the asymmetric unit, the calculated Matthews coefficient (V_M; Matthews, 1968) was 2.3 $Å^3$ Da^{-1}, indicating a solvent content of 46.5%.

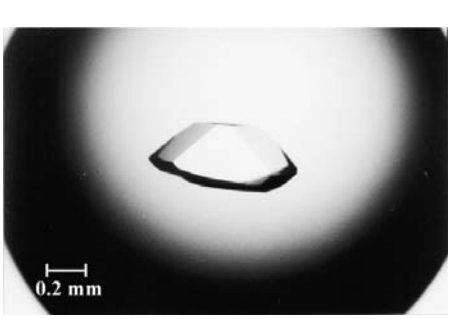


Figure 1
Native crystal of the lectin from *C. maritima* seeds.

Table 1
Summary of data-collection statistics for ConM.

Values in parentheses are for the highest resolution shell.

X-ray wavelength (Å)	1.4270
Unit-cell parameters	
a (Å)	67.15
b (Å)	70.90
c (Å)	97.37
Space group	$P2_12_12$
Resolution (Å)	39.5–2.1 (2.21–2.10)
No. measurements with $I > 2\sigma(I)$	246192
No. independent reflections	25202
Completeness (%)	91.3 (91.3)
$\langle I/\sigma(I)\rangle$	5.4 (3.0)
R_{sym}†	7.6 (23.7)

† $R_{sym} = \sum_h \sum_i |I(h)_i - \langle I(h)\rangle| / \sum_h \sum_i I(h)_i$, where $I(h)$ is the intensity of reflection h, $\sum_h$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection h.

The preliminary crystal structure of ConM was determined by molecular-replacement methods using the program *AMoRe* (Navaza, 1994). The atomic coordinates of several lectins were used in the search for a structural model. The best result was obtained with the lectin isolated from *C. ensiformis* (PDB code 3enr; Bouckaert *et al.*, 2000), which presented a final correlation coefficient of 69.2% and an *R* factor of 42.5%. Refinement of the structure is in progress.

This work was partly financed by Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico-FUNCAP, Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq, FAPESP, Universidade Regional do Cariri-URCA, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES, National Synchrotron Light Laboratory-LNLS, Brazil and FAPESP (SMOLBNet, 01/07532-0). BSC and WFA are senior investigators of CNPq.

References

Alencar, N. M. N., Teixeira, E. H., Assreuy, A. M., Cavada, B. S., Flores, C. A. & Ribeiro, R. A. (1999). *Mediators Inflamm.* **8**, 107–113.
Andrade, J. L., Arruda, S., Barbosa, T., Paim, L., Ramos, M. V., Cavada, B. S. & Barral-Netto, M. (1999). *Cell. Immunol.* **194**, 98–102.
Assreuy, A. M., Martins, G. J., Moreira, M. E. F., Brito, G. A. C., Cavada, B. S., Ribeiro, R. A. & Flores, C. A. (1999). *J. Urol.* **161**, 1988–1993.
Barbosa, T., Arruda, S., Cavada, B. S., Grangeiro, T. B., Freitas, L. A. R. & Barral-Netto, M. (2001). *Mem. Inst. Oswaldo Cruz*, **96**, 673–678.
Barral-Neto, M., Santos, S. B., Barral, A., Moreira, L. I. M., Santos, C. F., Moreira, R. A., Oliveira, J. T. A. & Cavada, B. S. (1992). *Immunol. Invest.* **21**, 297–303.
Bento, C. A. M., Cavada, B. S., Oliveira, J. T. A., Moreira, R. A. & Barja-Fidalgo, C. (1993). *Agents Actions*, **38**, 48–54.
Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
Bouckaert, J., Loris, R. & Wyns, L. (2000). *Acta Cryst.* D**56**, 1569–1576.
Cavada, B. S., Barbosa, T., Arruda, S., Grangeiro, T. B. & Barral Netto, M. (2001). *Curr. Protein Pept. Sci.* **2**, 123–135.
Cavada, B. S., Moreira-Silva, L. I. M., Grangeiro, T. B., Santos, C. F., Pinto, V. P. T., Barral-Neto, M., Roque-Barreira, M. C., Gomes, J. C., Martins, J. L., Oliveira, J. T. A. & Moreira, R. A. (1996). *Lectins: Biology, Biochemistry, Clinical Biochemistry*, edited by E. Van Driessche, J. Fisher, S. Jeeckmans & T. C. Bog-Hansen, Vol. 11, pp. 74–80. Hellerup, Denmark: Textop.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Ferreira, R. R., Cavada, B. S., Moreira, R. A., Oliveira, J. T. A. & Gomes, J. C. (1996). *Inflamm. Res.* **45**, 442–447.

Gomes, J. C., Ferreira, R. R., Cavada, B. S., Moreira, R. A. & Oliveira, J. T. A. (1994). *Agents Actions*, **41**, 132–135.
Havt, A., Barbosa, P. S. F., Sousa, T. M., Martins, A. M. C., Nobre, A. C. L., Nascimento, K. S., Teixeira, E. H., Pinto, V. P. T., Sampaio, A. H., Fontales, M. C., Cavada, B. S. & Monteiro, H. S. A. (2003). *Protein Pept. Lett.* **10**, 191–197.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Matthews, B. W. (1968). *J. Mol Biol.* **33**, 491–487.
Moreno, F. B. M. B., Delatorre, P., Freitas, B. T., Rocha, B. A. M., Souza, E. P., Facó, F., Canduri, F., Cardoso, A. L. H., Freire, V. N., Lima Filho, J. L., Sampaio, A. H., Calvete, J. J., Azevedo, W. F. Jr & Cavada, B. S. (2004). *Acta Cryst.* D**60**, 1493–1495.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Perez, G., Perez, C., Sousa-Cavada, B., Moreira, R. A. & Richardson, M. (1991). *Phytochemistry*, **30**, 2619–2621.
Peumans, W. J. & Van Damme, E. J. M. (1995). *Plant Physiol.* **109**, 347–352.
Ramos, M. V., Moreira, R. A., Oliveira, J. T. A., Cavada, B. S. & Rouge, P. (1996). *Mem. Inst. Oswaldo Cruz*, **91**, 761–6.
Rodriguez, D., Cavada, B. S., Oliveira, J. T. A., Moreira, R. A. & Russo, M. (1992). *Braz. J. Med. Biol. Res.* **25**, 823–826.
Vijayan, M. & Chandra, N. (1999). *Curr. Opin. Struct. Biol.* **9**, 707–714.
Van Damme, E. J. M., Peumans, W. J., Barre, A. J. & Rougé, P. (1998). *Crit. Rev. Plant Sci.* **17**, 575–692.

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray diffraction analysis of a new chitin-binding protein from *Parkia platycephala* seeds

Benildo S. Cavada,[a]* Rolando E. R. Castellón,[a] Georg G. Vasconcelos,[a] Bruno A. M. Rocha,[a] Gustavo A. Bezerra,[a] Henri Debray,[b] Plínio Delatorre,[a,c] Celso S. Nagano,[a] Marcos Toyama,[d] Vicente P. T. Pinto,[a] Frederico B. M. B. Moreno,[e] Fernanda Canduri[e] and Walter F. de Azevedo Jr[e]*

[a]Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Biomol-Lab, CEP 60451-970, Caixa Postal 6043, Fortaleza-Ceará, Brazil, [b]Laboratoire de Chimie Biologique et Unité Mixte de Recherche du CNRS No. 8576, Université des Sciences et Technologies de Lille, Villeneuve d'Ascq, France, [c]Departamento de Ciências Biológicas, Universidade Regional do Cariri, Crato, CE 63105-000, Brazil, [d]Departamento de Bioquímica, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Campinas, SP, Brazil, and [e]Departamento de Física, UNESP, São José do Rio Preto, SP 15054-000, Brazil

Correspondence e-mail: bscavada@ufc.br, walterfa@ibilce.unesp.br

Received 29 March 2005
Accepted 29 July 2005
Online 31 August 2005

A chitin-binding protein named PPL-2 was purified from *Parkia platycephala* seeds and crystallized. Crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 55.19$, $b = 59.95$, $c = 76.60$ Å, and grew over several days at 293 K using the hanging-drop method. Using synchrotron radiation, a complete structural data set was collected to 1.73 Å resolution. The preliminary crystal structure of PPL-2, determined by molecular replacement, presents a correlation coefficient of 0.558 and an R factor of 0.439. Crystallographic refinement is in progress.

1. Introduction

Chitin, a natural homopolymer composed of β(1–4)-linked *N*-acetylglucosamine $(GlcNAc)_n$, is a major component of the exoskeleton of fungi (comprising up to 30% of fungal cell walls) and invertebrates. It is easily obtained from marine invertebrates, insects and algae (Patil *et al.*, 2000).

The complete enzymatic hydrolysis of chitin to free *N*-acetylglucosamine residues is performed by a chitinolic system and is known to be a continuous reaction. Different organisms produce a wide variety of hydrolytic enzymes that exhibit different substrate specificities. Some of them are called chitinases, which are enzymes that catalyze the hydrolysis of chitin. These proteins are a large and diverse group of enzymes that differ in their molecular structure, substrate specificity and catalytic mechanism (Kasprzewska, 2003). Specificity for chitin oligosaccharide is not a feature that is exclusive to the chitinases. Proteins named 'chitin-binding lectins' or 'hevein-like lectins' also possess affinity for *N*-acetylglucosamine residues, but cannot catalyze the hydrolysis of chitin (Van Damme *et al.*, 1998).

Several chitinases have been found in plants (angiosperms and gymnosperms) and are present in diverse tissues. Most are expressed by stress factors such as infection. Plants use chitinases as a defence against pathogenic fungi, but the enzymes may also perform other functions (Peumans *et al.*, 2002). Some chitinases have industrial and agricultural applications, such as in the biocontrol of pathogenic fungi and insects, as a target for biopesticides and in the production of chitooligosaccharides (Kasprzewska, 2003; Patil *et al.*, 2000).

Plant lectins with chitinase activity are poorly described in the literature. The acidic chitinase from *Brassica juncea* shows a structure that is distinct from those observed for chitinases studied previously. This difference is characterized by the presence of two chitin-binding sites (Zhao & Chye, 1999), which permit this protein to agglutinate cells and may provide an advantage over other chitinases in antimicrobial and antifungal activity (Chye *et al.*, 2005).

Many carbohydrate-binding proteins have been reported, in particular those purified from plants (Moreno *et al.*, 2004; Gadelha *et al.*, 2005). The majority are from the Leguminosae family and comprise lectins and chitinases from diverse sources. Legume lectins have been well studied as a model of carbohydrate recognition. In the subfamily Mimosoideae, however, apart from *Parkia plathycephala* 2 (PPL-2), only the seed lectins from *P. speciosa* (Suvachittanont & Peutpaiboon, 1992), *P. javanica* (Utarabhand & Akkayanont, 1995), *P. platycephala* (Cavada *et al.*, 1997) and *P. discolor* (Cavada *et al.*, 2000) have been isolated and characterized in detail. Moreover, crystal structures are only available for *P. platycephala* lectin (PPL-1)

in native form (PDB code 1zgr) and in complex with 5-bromo-4-chloro-3-indolyl-α-D-mannose (PDB code 1zgs).

Mass-spectrometric analysis indicates that the PPL-2 monomer is not glycosylated and contains six cysteine residues that are involved in three disulfide bonds; PPL-2 gives a main mass peak at 29 407. Functional analysis reveals that PPL-2 recognizes carbohydrates on red blood cells and agglutinates trypsin-treated rabbit erythrocytes (128 haemagglutinating units per millilitre). In addition, PPL-2 can hydrolyze β(1–4)-glycosidic linkages between 2-acetoamido-2-deoxy-β-D-glucopyranoses present in chitin. The exact mechanism of glycoside hydrolysis has been described by Cavada *et al.* (2005) and this mechanism reveals an endochitinase activity to be associated with PPL-2 from the elution times found for the GlcNAc, $(GlcNAc)_2$ and $(GlcNAc)_3$ standards. Hence, PPL-2 is the first and is a remarkable chimerolectin from the Mimosoideae, with the dual property of hydrolyzing chitin and binding sugar moieties on red blood cells (Cavada *et al.*, 2005).

In order to establish the crystal structure of this new member of the chitin-binding proteins, this work reports the crystallization and preliminary X-ray diffraction analysis of a hevamine-like protein from *P. platycephala* seeds, named PPL-2, that has the ability to agglutinate cells and shows inhibitory effects in the growth of bacterial colonies and nematode-egg eclosion (Castellón, 2004; Cavada *et al.*, 2005).

2. Material and methods

2.1. Purification and crystallization

Soluble proteins were extracted from the seeds of *P. platycephala* Benth in an extraction solution (0.1 *M* HCl with 0.1 *M* NaCl). After centrifugation, the supernatant was neutralized with sodium hydroxide (NaOH) and the neutralized solution was submitted to precipitation with ammonium sulfate. The fraction 0/60 was resuspended in 0.05 *M* Tris–HCl buffer with 0.1 *M* NaCl pH 7.0 and exhaustively dialyzed against this buffer. The protein was purified by affinity chromatography on a Red-Sepharose CL-6B (23.0 × 2.5 cm) column equilibrated with the same buffer; elution of the non-interacting material took place using the equilibration buffer and the protein was eluted with 0.05 *M* Tris–HCl with 3.0 *M* NaCl pH 7.0 and finally dialyzed against Milli-Q water and lyophilized (Castellón, 2004; Cavada *et al.*, 2005).

For crystallization trials, the purified lectin was dissolved at a concentration of 7.5 mg ml^{-1} in Milli-Q water. Microcrystals of PPL-2 were grown in Linbro plates at 293 K by the vapour-diffusion/sparse-matrix method (Jancarik & Kim, 1991) in hanging drops using Crystal Screen from Hampton Research. The drops were composed of equal volumes (3 µl) of protein solution and reservoir solution [0.2 *M* ammonium acetate, 0.1 *M* trisodium citrate dehydrate pH 5.6 and 30%(*w*/*v*) polyethylene glycol 4000] and were equilibrated against 500 µl reservoir solution. Microcrystals were seeded into a new drop containing the same crystallization solution and an equal protein volume.

2.2. X-ray data collection

X-ray diffraction data were collected at a wavelength of 1.4727 Å using a synchrotron-radiation source (CPr station, Laboratorio Nacional de Luz Síncrotron, Campinas, Brazil) and a CCD detector (MAR Research) with a crystal-to-detector distance of 70.00 mm at a temperature of 100 K. To avoid freezing, crystals were soaked in a cryoprotectant solution containing 75% mother liquor and 25% glycerol. Using an oscillation range of 1.0° and an exposure time of 30 s per frame, 90 images were collected to a maximum resolution of 1.73 Å. Data were indexed, integrated and scaled using *MOSFLM* and *SCALA* (Collaborative Computational Project, Number 4, 1994).

2.3. Molecular replacement

Sequence-alignment analysis was performed using programs that compared the N-terminal sequence of PPL-2 with those of all the non-redundant bank of proteins deposited in the National Center of Biotechnology Information (NCBI). Local and multiple alignments were carried out using *BLAST* (Altschul *et al.*, 1990) and *CLUSTALW* (Thompson *et al.*, 1994), respectively. To perform multiple alignments, plant chitinases from *Nicotiana tabacum*, *Phytolacca americana*, *Glycine max*, *Zea mays*, *Vitis vinifera*, *Arabdopsis thaliana*, *Vigna unguiculata* and *Hevea brasiliensis* were used.

The molecular-replacement method was used to determine the crystal structure of PPL-2 using the *AMoRe* software (Navaza, 1994). Rotation and translation functions were performed using data in the resolution range 15–3.0 Å. The best solution for each model was selected based on the magnitude of the correlation coefficient and the *R* factor. Four space groups were tested ($P222$, $P2_12_12$, $P222_1$ and $P2_12_12_1$) using the hevamine protein coordinates (PDB code 1hvq, chain *A*; Terwisscha van Scheltinga *et al.*, 1996).

3. Results and discussion

Microcrystals grew in a month using condition No. 9 of Crystal Screen from Hampton Research; when submitted to the seeding experiment, a crystal cluster was observed (Fig. 1*a*) after another month. The drop was perturbed using a fine hair and after about 20 weeks suitable

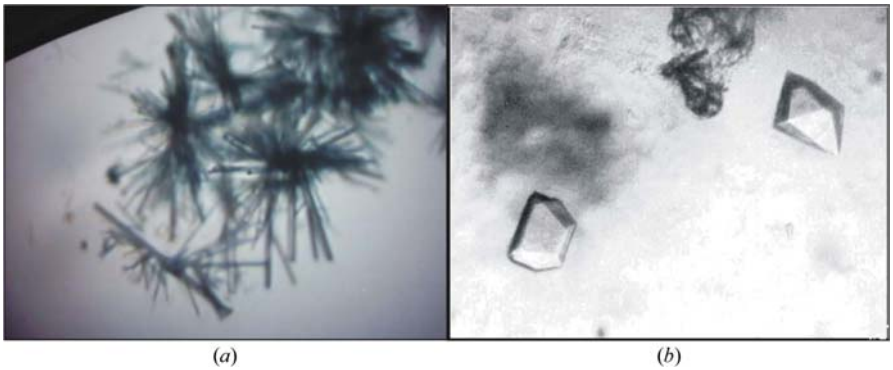


Figure 1
PPL-2 crystals. (*a*) Crystal cluster from PPL-2 seeding. (*b*) Crystals of PPL-2 diffracted to 1.73 Å.

Table 1
X-ray diffraction data collection.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.431
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 55.19, b = 59.95, c = 76.70
Resolution range (Å)	32.27–1.73
Unique reflections	25945
Completeness (%)	95.5 (95.5)
$\langle I/\sigma(I)\rangle$	13.1 (2.4)
R_{sym} (%)	4.0 (22.8)
R_{full} (%)	3.6 (16.3)

Figure 2
Multiple alignment of the N-terminal sequence of PPL-2 with those of plant chitinases. GmCHI, acidic chitinase from *G. max*; ZmCHI, acidic chitinase from *Z. mays*; Hevamina, chitinase/lysozyme from *H. brasiliensis*; VvCHI, precursor of acidic chitinase from *Vitis vinifera*; AtCHI, acidic endochitinase from *A. thaliana*; VuCHI, basic endochitinase type III from *Vigna unguiculata*; PaCHI, chitinase-B from *Phytolacca americana*; NtCHI, basic endochitinase type III from *N. tabacum*.

crystals were obtained (Fig. 1*b*). The best crystals grew to approximate dimensions of 0.3 × 0.2 × 0.3 mm.

Crystals of PPL-2 were grown by the hanging-drop vapour-diffusion method. PPL-2 crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 55.19, b = 59.95, c = 76.70 Å. The volume of the unit cell is 253 757.48 Å^3, which is compatible with one monomer in the asymmetric unit, with a V_M of 2.3 Å^3 Da^{-1} (Matthews, 1968). A summary of the data-collection statistics is given in Table 1.

Sequence-alignment analysis permitted the retrieval of a viable search model to submit these data to molecular replacement. The N-terminal amino-acid sequence from PPL-2 (GGIVVWGQNGGEGTLTSTCESGLYQIVNIAFLSQFGGGRRV) is completely different from that found in PPL-1, a lectin isolated from seeds of *P. platycephala* (SLKGMISVGPWGGSGGNYWSFKANHAITEIVIHVKDNIKS; Cavada *et al.*, 1997). Based on local alignments of PPL-2, similarity has been found with chitinases, proteins that are reported to be related to defence mechanisms in plants. These alignments show that PPL-2 exhibits a high sequence similarity to type III chitinases. One of these proteins is hevamine, a chitinase and lysozyme protein found in latex from *H. brasiliensis*, which has the N-terminal sequence GGIAIYWGQNGNEGTLTQTCSTRKYSYVNIAFLNKFGNGQ. Acidic chitinases extracted from the leaves of *G. max* (Watanabe *et al.*, 1999), *Z. mays* (Didierjean *et al.*, 1996) and *A. thaliana* (Kawabe *et al.*, 1997) also show similarity with PPL-2, as can be observed in Fig. 2.

The N-terminal alignment of PPL-2 and several chitinases demonstrated a degree of similarity estimated at 72% with basic endochitinase type III from *N. tabacum*, 71% with chitinase-B from *Phytolacca americana*, 71% with acidic chitinases from *G. max* and *Z. mays*, 66% with the precursor of the acidic chitinase (Precursor) from *Vitis vinifera*, 60% with acidic endochitinase from *A. thaliana*, 61% with basic endochitinase type III from *Vigna unguiculata* and 64% with hevamine, a chitinase/lysozyme from *H. brasiliensis*. Despite the high similarity observed among all the alignments, hevamine is the only one in which the polypeptidic fragment corresponds to the N-terminal region as in PPL-2 and is thus a good search model for molecular replacement. The best results were obtained for space group $P2_12_12_1$, resulting in a correlation coefficient and R factor of 55.8 and 43.9%, respectively.

Initial crystallographic refinement was performed using rigid-body refinement followed by the maximum-likelihood method with the *REFMAC*5 software (Collaborative Computational Project, Number 4, 1994), resulting in a model with an R factor of 29.7% and an R_{free} of 33.1%. Complete refinement of the structure of PPL-2 is in progress.

This work was partly financed by Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Universidade Regional do Cariri (URCA), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Laboratório Nacional de Luz Síncrotron (LNLS), Campinas, Brazil and FAPESP (SMOLBNet, 01/07532-0). BSC and WFA are senior investigators of CNPq.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Castellón, R. E. R. (2004). Tese de Doutorado apresentada ao Departamento de Fitotecnia, Universidade Federal do Ceará, Fortaleza-Ceará, Brazil.
Cavada, B. S., Castellón, R. E. R., Vasconcelos, G. G., Delatorre, P., Nagano, C. S., Nascimento, K. S., Souza, E. P., Radis-Baptista, G., Leroy, Y., Toyama, M. H., Sampaio, A. H., Pinto, V. P. T. & Debray, H. (2005). Submitted.
Cavada, B. S., Madeira, S. V. F., Calvete, J. J., Sousa, L. A. G., Bomfim, L. R., Dantas, A. R., Lopes, M. C., Grangeiro, T. B., Freitas, B. T., Pinto, V. P. T., Leite, K. B. & Ramos, M. V. (2000). *Prep. Biochem. Biotech.* **30**, 271–280.
Cavada, B. S., Santos, C. F., Grangeiro, T. B., Silva, L. I. M. M., Campos, M. J. O., Sousa, F. A. M. & Calvete, J. J. (1997). *Physiol. Mol. Biol. Plants*, **3**, 109–115.
Chye, M. L., Zhao, K. J., He, Z. M., Ramalingam, S. & Fung, K. L. (2005). *Planta*, **220**, 717–730.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Didierjean, L., Frendo, P., Nasser, W., Genot, G., Marivet, J. & Burkard, G. (1996). *Planta*, **199**, 1–8.
Gadelha, C. A. A., Moreno, F. B. M. B., Santi-Gadelha, T., Cajazeiras, J. B., Rustiguel, J. K. R., Freitas, B. T., Canduri, F., Delatorre, P. & de Azevedo, W. F. Jr (2005). *Acta Cryst.* F**61**, 87–89.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kasprzewska, A. (2003). *Cell Mol. Biol. Lett.* **8**, 809–824.
Kawabe, A., Innan, H., Terauchi, R. & Miyashita, N. T. (1997). *Mol. Biol. Evol.* **14**, 1303–1315.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Moreno, F. B. M. B., Delatorre, P., Freitas, B. T., Rocha, B. A. M., Souza, E. P., Facó, F., Canduri, F., Cardoso, A. L. H., Freire, V. N., Lima-Filho, J. L., Sampaio, A. H., Calvete, J. J., de Azevedo, W. F. Jr & Cavada, B. S. (2004). *Acta Cryst.* D**60**, 1493–1495.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Patil, R. S., Ghormade, V. V. & Deshpande, M. V. (2000). *Enzyme Microb. Technol.* **26**, 473–483.
Peumans, W. J., Proost, P., Swennen, R. L. & Van Damme, E. J. (2002). *Plant Physiol.* **130**, 1063–1072.
Suvachittanont, W. & Peutpaiboon, A. (1992). *Phytochemistry*, **31**, 4065–4070.
Terwisscha van Scheltinga, A. C., Hennig, M. & Dijkstra, B. W. (1996). *J. Mol. Biol.* **262**, 243–257.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Utarabhand, P. & Akkayanont, P. (1995). *Phytochemistry*, **38**, 281–285.
Van Damme, J. M., Peumans, W. J., Barre, A. J. & Rougé, P. J. (1998). *Crit. Rev. Plant Sci.* **17**, 575–692.
Watanabe, A., Nong, V. H., Zhang, D., Arahira, M., Yeboah, N. A., Udaka, K. & Fukazawa, C. (1999). *Biosci. Biotechnol. Biochem.* **63**, 251–256.
Zhao, K. J. & Chye, M. L. (1999). *Plant Mol. Biol.* **40**, 1009–1018.

Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Purification, partial characterization and preliminary X-ray diffraction analysis of a mannose-specific lectin from *Cymbosema roseum* seeds

Benildo S. Cavada,[a]* Emmanuel S. Marinho,[a] Emmanuel P. Souza,[a] Raquel G. Benevides,[a] Plínio Delatorre,[a] Luis A. G. Souza,[b] Kyria S. Nascimento,[a] Alexandre H. Sampaio,[c] Frederico B. M. B. Moreno,[d] Joane K. R. Rustiguel,[d] Fernanda Canduri,[e] Walter F. de Azevedo Jr[f]* and Henri Debray[g]

[a]BioMol-Lab – Department of Biochemistry, Federal University of Ceará, Brazil, [b]INPA, Manaus, Amazonas, Brazil, [c]Biomol-Mar-Fishing Engineering Faculty, Federal University of Ceará, Brazil, [d]Programa de Pós-Graduação em Biofísica Molecular, Departamento de Física, UNESP, São José do Rio Preto, SP, Brazil, [e]Departamento de Morfofisiologia – CCBS – UFMS, Campo Grande-MS, 79070-900, Brazil, [f]Faculdade de Biociências – PUCRS – Porto Alegre-RS, 90619-900, Brazil, and [g]Laboratoire de Chimie Biologique et Unité Mixte de Recherche No. 8576 du CNRS, Université des Sciences et Technologies de Lille, France

Correspondence e-mail: bscavada@ufc.br, wfdaj@uol.com.br

Received 1 December 2005
Accepted 30 January 2006

A lectin from *Cymbosema roseum* seeds (CRL) was purified, characterized and crystallized. The best crystals grew in a month and were obtained by the vapour-diffusion method using a precipitant solution consisting of 0.1 *M* Tris–HCl pH 7.8, 8%(*w*/*v*) PEG 3350 and 0.2 *M* proline at a constant temperature of 293 K. A data set was collected to 1.77 Å resolution at a synchrotron-radiation source. CRL crystals are orthorhombic, belonging to space group $P2_12_12_1$. Crystallographic refinement and full amino-acid sequence determination are in progress.

1. Introduction

Plant lectins are a structurally heterogeneous group of carbohydrate-binding proteins of non-immune origin that exhibit a variety of interactions in cellular processes such as cell communication, host defence, fertilization, parasitic infection and tumour metastasis. These lectins are present in various organisms, animal and vegetal, including leguminous plants. Despite the high level of conservation exhibited in their sequences, leguminous lectins show considerable diversity in their carbohydrate-binding properties and biological effects (Wah *et al.*, 2001).

The Diocleinae lectins, a well studied group of closely related leguminous lectins, exhibit biological effects such as histamine release from rat peritoneal mast cells (Ferreira *et al.*, 1996) and anti- (Assreuy *et al.*, 1997) and pro-oedematogenic effects (Alencar *et al.*, 1999). Minor differences in the ratios of dimeric and tetrameric forms in the lectins, together with differences in the relative orientations of the carbohydrate-binding sites in the quaternary structures, have been hypothesized to contribute to the differences in biological activities exhibited by Diocleinae lectins (Sanz-Aparicio *et al.*, 1997).

The species *Cymbosema roseum* belongs to the Diocleinae subtribe of the Leguminosae family and is widespread throughout the Amazonian forest (Anavilhanas archipelago, Amazonas, Brazil). This work reports the purification, partial characterization, crystallization and preliminary X-ray diffraction analysis of a lectin from *C. roseum* (CRL).

2. Materials and methods

2.1. Purification and crystallization of *C. roseum* lectin (CRL)

C. roseum seeds were ground to a fine powder in a coffee mill. The powder was stirred with 0.15 *M* NaCl [1:10(*w*:*v*)] at room temperature for 4 h and then centrifuged at 10 000*g* for 20 min at 278 K. The resultant supernatant was applied onto a Sepharose-4B-mannose column (0.5 × 10 cm) equilibrated with 0.15 *M* NaCl containing 5 m*M* $CaCl_2$ and 5 m*M* $MnCl_2$. After removing unbound material, the lectin was eluted with 0.1 *M* glycine, 0.15 *M* NaCl pH 2.6. Purified CRL was monitored by SDS–PAGE as described by Laemmli (1970) and was used to perform further characterization. N-terminal sequence analysis was performed using an Applied Biosystems pulsed-liquid phase 477A protein sequencer with a 120A PTH amino-acid analyzer, following the method described by the manufacturer.

Haemagglutinating activity and haemagglutination-inhibition studies were carried out in micro-titration plates using a standard procedure (Faria *et al.*, 2004). The haemagglutinating activity of the lectin was determined using native and enzyme-treated rabbit

erythrocytes. Native cells were prepared by washing the erythrocytes three times with 0.1 *M* NaCl and then resuspending them to a final concentration of 2%(*v*/*v*) in 0.15 *M* NaCl. To prepare enzyme-treated cells, washed packed erythrocytes were incubated with an equal volume of papain or trypsin for 30 min at 310 K. Treated cells were washed three times in 0.15 *M* NaCl and resuspended to a final concentration of 2%(*v*/*v*) in 0.15 *M* NaCl. Haemagglutination tests were performed on serial twofold dilutions of lectin solutions in PBS, each dilution having a final volume of 0.2 ml. A 0.2 ml aliquot of the 2% erythrocyte suspension was added to each dilution. The plates were gently shaken and left for 90 min at room temperature, after which time the degree of macroscopic agglutination was recorded.

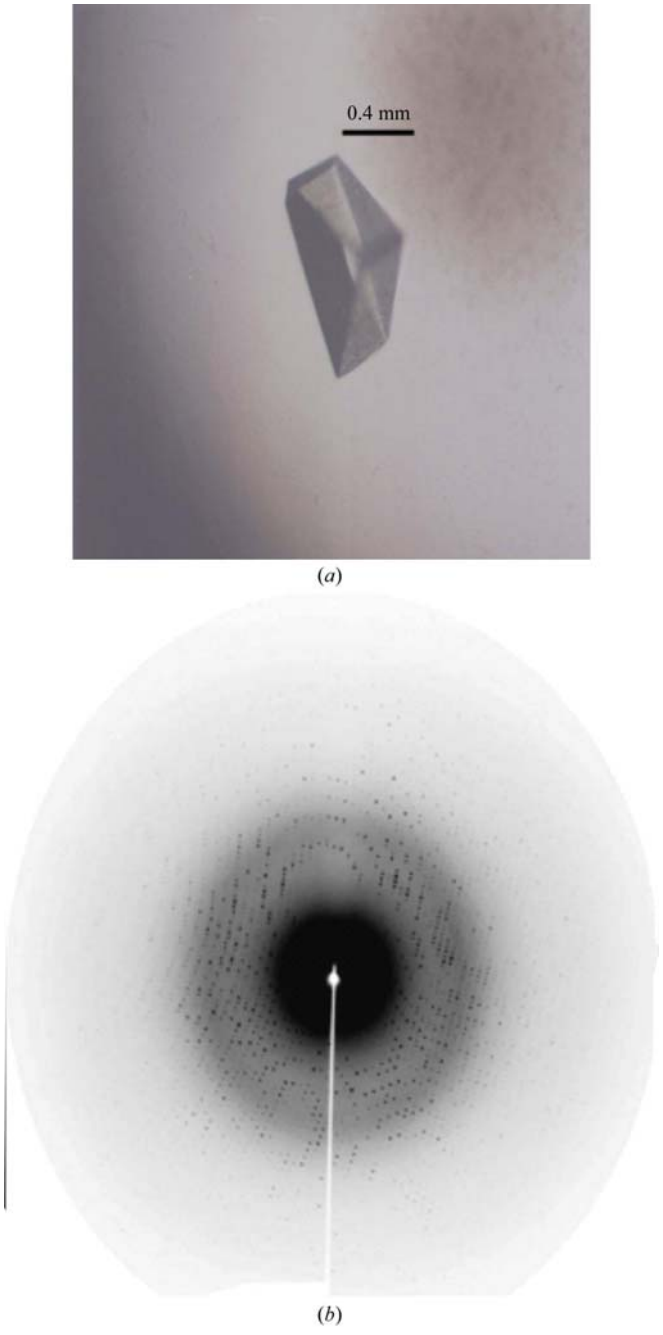


Figure 1
(*a*) Crystal and (*b*) diffraction pattern of *C. roseum* lectin.

Inhibition tests were carried out using stock solutions of sugars and glycoproteins in 0.15 *M* NaCl. A twofold dilution series was prepared for each potentially inhibitory substance in 0.15 *M* NaCl containing 5 m*M* $CaCl_2$ and 5 m*M* $MnCl_2$. Each dilution had a final volume of 0.2 ml. The lectin samples were diluted in 0.15 *M* NaCl to give a solution containing four units of haemagglutinating activity per millilitre (the greatest dilution of the lectin that can agglutinate erythrocytes, *i.e.* the titre, was defined as containing one haem-agglutinating unit per millilitre). 0.2 ml aliquots of the diluted lectin (four units) were added to each well of the diluted inhibitor series. The plates were then left at room temperature for 1 h before 0.2 ml of 2% native or enzyme-treated rabbit erythrocytes was added to each well. The plates were then allowed to stand at room temperature for a further 1 h before being examined for haemagglutination. The haemagglutination-inhibition titre was recorded as the the highest dilution of inhibitor which inhibited the agglutination produced by four haemagglutination units of lectin sample.

The lyophilized purified CRL was dissolved to a concentration of 12 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.0 containing 0.5 m*M* $CaCl_2$ and $MnCl_2$ and used for crystallization trials. Crystallization screening by the hanging-drop vapour-diffusion method was performed in Linbro plates at 293 K using Hampton Research Crystal Screens I and II, SaltRx, Index and PEG/Ion Screens (Hampton Research, Aliso Viejo, CA, USA). The drops were composed of equal volumes (2 µl) of protein solution and reservoir solution and were equilibrated against 500 µl reservoir solution. An example of a crystal of CRL is shown in Fig. 1(*a*).

2.2. X-ray data collection

A crystal was transferred to a cryoprotectant solution consisting of 30% glycerol in the crystallization reservoir solution. Data were collected at 1.42 Å wavelength at a synchrotron-radiation source (beamline MX1, CPr station, Laboratório Nacional de Luz Síncrotron–LNLS, Campinas, Brazil) using a MAR Research CCD imaging plate at a crystal-to-detector distance of 70 mm. A set of 100 1° oscillation images was recorded (an image is shown in Fig. 1*b*). Diffraction data were indexed, integrated and scaled using *MOSFLM* and *SCALA* (Collaborative Computational Project, Number 4, 1994).

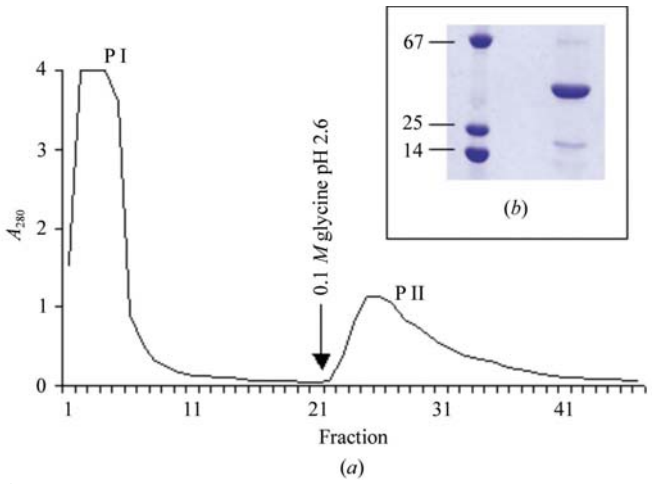


Figure 2
Affinity chromatography and SDS–PAGE of the purified *C. roseum* lectin (CRL). (*a*) Sepharose-4B-mannose chromatography. (*b*) SDS–PAGE showing protein markers (left) and a purified CRL band (right). The molecular-weight markers are bovine serum albumin (67 kDa), chymotrypsinogen (25 kDa) and lysozyme (14 kDa). The gel was Coomassie blue stained.

Table 1
N-terminal sequence alignment of Diocleinae lectins.

Lectin	Sequence
Cymbosema roseum	ADTIVAVELDSYPNTDIGDPSYPH
Dioclea lehmanni I	ADTIVAVELDSYPNTDIGDPSYPH
Dioclea grandiflora	ADTIVAVELNSYPNTDIGDPNYPH
Canavalia ensiformis	ADTIVAVELDTYPNTDIGDPSYPH
Canavalia brasiliensis	ADTIVAVELDSYPNTDIGDPSYPH
Cratylia floribunda	ADTIVAVELDSYPNTDIGDPNYQH

An alignment analysis was performed with *ClustalW* (Thompson *et al.*, 1994) that compared the N-terminal sequence of CRL (which was evidence of the success of the purification) with those of all non-redundant proteins deposited in the National Center of Biotechnology Information (NCBI). Determinations of the full amino-acid sequence and the three-dimensional structure are in progress.

3. Results and discussion

C. roseum lectin (CRL) was purified by a single step using a Sepharose-4B-mannose affinity chromatography column (Fig. 2*a*). All purification steps were monitored by haemagglutinating activity and SDS–PAGE.

CRL showed haemagglutinating activity towards papain-treated, trypsin-treated and untreated rabbit erythrocytes. The minimal concentration of purified protein that agglutinated a 2% rabbit erythrocyte suspension was $<2\ \mu g\ ml^{-1}$. This haemagglutinating property is similar to those of other Diocleinae lectins. From the inhibitory substances tested, mannose was the most potent, with a minimum concentration of 19.5 m*M*. The CRL N-terminal sequence was found to be ADTIVAVELDSYPNTDIGDPSYPH. This sequence is very similar to those of other lectins (Table 1) from the subtribe Diocleinae [100% identity with *Dioclea lehmanni* I (Perez *et al.*, 1990) and *Canavalia brasiliensis* (Moreira & Cavada, 1984); 92% identity with *D. grandiflora* (Moreira *et al.*, 1983), ConA (Hague, 1975) and *Cratylia floribunda* (Oliveira *et al.*, 1991)]. The apparent molecular weight of CRL was determined by SDS–PAGE after heating in the presence of SDS (Fig. 2*b*). The lectin appears to be composed of three polypeptide chains of approximate molecular weights of 30, 18 and 12 kDa. This variety of molecular weights is similar to that found for other lectins from Diocleineae species such as *Canavalia ensiformis* (Hague, 1975). This latter lectin, commonly known as concanavalin A, has been shown to initially be expressed with an initial set of termini, which led to the supposition that the gene product is then cleaved post-translationally to form two chains and a second set of termini and then fused into a single chain again by peptide-bond formation that joins the initial set of termini, a kind of circular permutation of the sequence. In the absence of other evidence, we suppose that the smaller chains are analogous to the two chains of the cleaved gene product and the largest chain is analogous to the fused final product (Cunningham *et al.*, 1979).

Microcrystals were obtained using 0.1 *M* HEPES pH 7.5 containing 10%(*w*/*v*) PEG 3350 and 0.2 *M* proline. Improvement of this crystallization condition was obtaining by increasing the pH and PEG concentration. Suitable crystals were obtained from drops containing 0.1 *M* Tris–HCl pH 7.8, 8%(*w*/*v*) PEG 3350 and 0.2 *M* proline. CRL crystals grew within a month to maximum dimensions of approximately 0.8 × 0.4 × 0.4 mm (Fig. 1*a*). The diffraction data showed the CRL crystals to be orthorhombic, belonging to space group $P2_12_12_1$, with unit-cell parameters $a = 67.8$, $b = 103.1$, $c = 122.1$ Å. CRL crystal data were scaled in the resolution range 34.92–1.77 Å. Statistics of the data collection can be found in Table 2. Assuming the presence of four molecules of 25 kDa (as is standard for Diocleinae lectins that present an apparent weight of 30 kDa in SDS–PAGE, such as ConA and *D. lehmanni* lectin I) in each monomer in the asymmetric unit, the calculated V_M was 2.1 $Å^3\ Da^{-1}$ (Matthews, 1968), indicating a solvent content of 41.9%. Elucidation of the complete amino-acid sequence of CRL, now in progress, will permit the refinement of a three-dimensional CRL structure.

Table 2
Summary of data-collection statistics for CRL.

Values in parentheses are for the highest resolution shell (1.87–1.77 Å).

X-ray wavelength (Å)	1.427
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 67.82$, $b = 103.14$, $c = 122.09$
Resolution limits (Å)	34.92–1.77
Asymmetric unit content	4 molecules
Total reflections measured	286361
Unique reflections measured	80568
Completeness (%)	97.00 (97.0)
R_{merge} (%)	5.4 (32.5)
$\langle I/\sigma(I)\rangle$	7.1 (2.2)

We thank Dr Juan J. Calvete of Instituto de Biomedicina de Valencia for his technical support. This work was partly supported by Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), FAPESP (SMOLBNet, 01/07532-0), Universidade Regional do Cariri (URCA), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Laboratório Nacional de Luz Síncrotron (LNLS) in Brazil. BSC, AHS and WFA are senior investigators of CNPq.

References

Alencar, N. M. N., Teixeira, E. H., Assreuy, A. M. S., Cavada, B. S., Flores, C. A. & Ribeiro, R. A. (1999). *Mediators Inflamm.* **8**, 107–113.
Assreuy, A. M. S., Shibuya, M. D., Martins, G. J., Souza, M. L. P., Cavada, B. S., Moreira, R. A., Oliveira, J. T. A., Ribeiro, R. A. & Flores, C. A. (1997). *Mediators Inflamm.* **6**, 201–210.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cunningham, B. A., Hemperly, J. J., Hopp, T. P. & Gerald, M. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 3218–3222.
Faria, R. A., Andrade-Neto, M., Pinto, L. S., Castellon, R. R., Calvete, J. J. & Cavada, B. S. (2004). *Arch. Latinoam. Nutr.* **54**, 349–353.
Ferreira, R. R., Cavada, B. S., Moreira, R. A., Oliveira, J. T. A. & Gomes, J. G. (1996). *Inflamm. Res.* **45**, 442–447.
Hague, D. R. (1975). *Plant Physiol.* **55**, 636–642.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Moreira, R. A., Barros, A. C. H., Stewart, J. C. & Pusztai, A. (1983). *Planta*, **158**, 63–69.
Moreira, R. A. & Cavada, B. S. (1984). *Biologia Plant.* **26**, 113–120.
Oliveira, J. T. A., Cavada, B. S. & Moreira, R. A. (1991). *Rev. Bras. Bot.* **14**, 63–68.
Perez, G., Hernandez, M. & Mora, E. (1990). *Phytochemistry*, **29**, 1745–1749.
Sanz-Aparicio, J., Hermoso, J., Grangeiro, T. B., Calvete, J. J. & Cavada, B. S. (1997). *FEBS Lett.* **405**, 114–118.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Wah, D. A., Romero, A., Gallego del Sol, F., Cavada, B. S., Ramos, M. V., Grangeiro, T. B., Sampaio, A. H. & Calvete, J. J. (2001). *J. Mol. Biol.* **310**, 885–894.

Available online at www.sciencedirect.com
ScienceDirect
Journal of Structural Biology 160 (2007) 168–176
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

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

Gustavo Arruda Bezerra [a], Taianá Maia Oliveira [a], Frederico Bruno Mendes Batista Moreno [b], Emmanuel Prata de Souza [a], Bruno Anderson Matias da Rocha [a], Raquel Guimarães Benevides [a], Plínio Delatorre [a], Walter Filgueira de Azevedo Jr. [c,*], Benildo Sousa Cavada [a,*]

[a] *Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Biomol-LAB, Campus do Pici S/N, Fortaleza, Ceará, Brazil*
[b] *Curso de Pós-graduação em Biofísica Molecular, Departamento de Física, Universidade Estadual Paulista Júlio de Mesquita Filho, R. Cristovão Colombo, 2265 Nazareth, São José do Rio Preto, São Paulo, Brazil*
[c] *Faculdade de Biociências-PUCRS Av. Ipiranga, 6681 Porto alegre, Rio Grande do Sul, Brazil*

Received 23 April 2007; received in revised form 21 June 2007; accepted 30 July 2007
Available online 16 August 2007

Abstract

Plant lectins, especially those purified from species of the Leguminosae family, represent the best studied group of carbohydrate-binding proteins. The legume lectins from Diocleinae subtribe are highly similar proteins that present significant differences in the potency/efficacy of their biological activities. The structural studies of the interactions between lectins and sugars may clarify the origin of the distinct biological activities observed in this high similar class of proteins. In this way, this work presents a crystallographic study of the ConM and CGL (agglutinins from *Canavalia maritima* and *Canavalia gladiata*, respectively) in the following complexes: ConM/CGL:Man(α1-2)Man(α1-*O*)Me, ConM/CGL:Man(α1-3)Man(α1-*O*)Me and ConM/CGL:Man(α1-4)Man(α1-*O*)Me, which crystallized in different conditions and space group from the native proteins.

The structures were solved by molecular replacement, presenting satisfactory values for R_{factor} and R_{free}. Comparisons between ConM, CGL and ConA (*Canavalia ensiformis* lectin) binding mode with the dimannosides in subject, presented different interactions patterns, which may account for a structural explanation of the distincts biological properties observed in the lectins of Diocleinae subtribe.

Keywords: Legume lectin; Dimannoside; *Canavalia maritima* lectin; *Canavalia gladiata* lectin

1. Introduction

Lectins are carbohydrate-binding proteins or glycoproteins of non-immune origin found in all types of living organisms that decipher the glycocodes encoded in the structure of glycans attached to soluble and integral cell-membrane glycoconjugates (Gabius and Gabius, 1997). Plant lectins are valuable tools for biochemical, analytical, and therapeutic purposes (Lis and Sharon, 1986). Legume lectins in particular have become a model system for protein–carbohydrate recognition because of their wide range of specificities for monosaccharides as well as for complex carbohydrates. Recognition between proteins and carbohydrates is of prime importance in many biological processes,

* Corresponding authors. Fax: +55 85 40089818.
E-mail addresses: walter.junior@pucrs.br (W.F. de Azevedo Jr.), bscavada@ufc.br (B.S. Cavada).

1047-8477/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2007.07.012

such as viral, bacterial, mycoplasmal and parasitic infections, targeting of cells and soluble components, fertilization, cancer metastasis and growth and differentiation (Sharon and Lis, 1990). The specific recognition of an (oligo)saccharide by a protein is a much more complex problem than other biologically relevant recognition processes such as protein–protein or protein–DNA interactions (Loris et al., 2004).

The legume lectins from the Diocleinae subtribe are highly similar proteins that exhibit different biological activities such as human lymphocyte proliferation and interferon γ production (Barral-Netto et al., 1992), peritoneal macrophage stimulation and inflammatory reaction (Rodriguez et al., 1992), induction of paw edema and peritoneal cell immigration in rats (Bento et al., 1993), histamine release from rat peritoneal mast cells (Gomes et al., 1994, 1996), nitric oxide production (Andrade et al., 1999) and *in vivo* lymphocyte activation and apoptosis (Barbosa et al., 2001) among others.

Several factors contribute to these differences in the activities: for example, the pH-dependent oligomerization that some of these lectins present and the relative position of the carbohydrate-binding site (Wah et al., 2001) as well as minor changes in the residues located at key positions regarding to quaternary association (Brinda et al., 2004). The distinct biological activities presented by *Canavalia ensiformis* lectin (ConA) and *Canavalia brasiliensis* lectin (ConBr) have been shown to be a consequence of only one non-conservative substitution that affects the carbohydrate-binding site, making it more open in ConBr than it is in ConA (Sanz-Aparicio et al., 1997).

The structural studies of these proteins are important not only to elucidate the molecular mechanism by which these molecules produce their distinct biological activities, but also to highlight several aspects of the protein–carbohydrate interaction and how lectins decipher the great informative potential encoded into the structures of glycans.

Thus, we have determined the crystallographic structure of six complexes: *Canavalia gladiata* lectin (CGL) and *Canavalia maritima* lectin (ConM) with dimannosides Man(α1-2)Man(α1-*O*)Me, Man(α1-3)Man(α1-*O*)Me and Man(α1-4)Man(α1-*O*)Me, in order to improve the understanding of the structural basis for the affinity of lectins for different dimannosides.

2. Materials and methods

The crystallization procedures, data collection and processing were previously described in detail (Moreno et al., 2006). The phase problem for all the crystal structures was solved by the molecular replacement method using the program MolRep (Vargin and Teplyakov, 1997) from the CCP4 package (CCP4, 1994) (“*Collaborative Computational Project, number 4, 1994*”). ConM (PDB code: 2CWM) and CGL structures (PDB code: 1 WUV) were used as the search model, with all waters removed.

The initial models were afterwards submitted to several cycles of rigid body and restrained refinement, using the program REFMAC5 and monitored using the R_{free} and R_{factor} values (Brunger, 1992). The electron densities were visualized through XtalView (Mcree, 1999), in which dimannosides were fitted and subtle adjustments in the models were carried out, followed by restrained refinement. The PDB entries for all the dimannosides were generated with the program Library/Sketcher from CCP4 package.

The quality of these models was inspected by the program PROCHECK (Laskowski et al., 1993) and the accessible surface areas were calculated using the program NACCESS (Hubbard and Thornton, 1993) with a probe radius of 1.4 Å. Van der Walls contacts were analyzed using the CCP4 program CONTACT with a cut-off distance of 4.5 Å. The crystallographic B-factors (overall, main chains, side chains, waters and the sugars) as well as the Ramachandran plot and root-mean-squares deviations from ideal geometry are presented in Tables 1 and 3. The final refinement statistics for ConM and CGL complexes are presented in Tables 2 and 4, respectively. All figures and superpositions were performed using the program PyMOL (Delano, 2002).

3. Results and discussion

3.1. Overall structure

The four structures, ConM and GGL complexed with Man(α1-3)Man(α1-O)Me (M1-3M) and Man(α1-4)Man(α1-O)Me (M1-4M), exhibit a classic “canonical” tetramer (subunits A, B, C and D) in the asymmetric unit, while the ConM and CGL complexed with Man(α1-2)Man(α1-O)Me (M1-2M) exhibits a single monomer. As previously observed in other legume lectins, the ConM monomer consists of 237 amino acids folded as a β sand-

Table 1
Ramachandran plot and root-mean-square deviations from ideal geometry for ConM structures

	ConM: M1-2M	ConM: M1-3M	ConM: 0M1-4M
Residues distribution			
Most favorable (%)	89.4	86.8	87.8
Allowed (%)	9.7	12.7	11.9
Additionally allowed (%)	1.0	0.4	0.4
Forbidden (%)	0	0.1	0.0
Root-mean-square deviations from ideal geometry			
Bond lengths (Å)	0.02	0.01	0.021
Bond angles (°)	2.143	1.589	2.218
Mean temperature factor (B-factor)			
Main chain (Å^2)	13.497	21.9685	19.437
Side chain (Å^2)	15.596	22.6035	20.1982
Water molecules (Å^2)	20.033	24.597	23.577
Sugar atoms (Å^2)[a]	25.392	29.86	20.542
Overall (Å^2)	18.629	24.75	20.938

[a] In ConM:M1-4M complex, subunits C and D were not accounted for the calculate.

Table 2
Final refinement statistics for ConM structures

Refinement	ConM: M1-2M	ConM: M1-3M	ConM: M1-4M
Space group	I222	P32	P32
Unit-Cell parameters (Å)	$a = 63.63$, $b = 85.44$, $c = 85.71$	$a = 69.39$, $b = 69.39$, $c = 161.29$	$a = 69.47$, $b = 69.47$, $c = 161.52$
Resolution limit (Å)	20.96 (1.4)	40.49 (2.10)	34.73 (2.10)
R_{factor}	0.206	0.187	0.176
Number of reflexions (R_{factor})	30,905	47,581	46,558
R_{free}	0.229	0.255	0.239
Number of reflexions (R_{factor})	1646	2524	2489
Number of residues per asymmetric unit	237	948	948
Number of water molecules	146	347	480

Table 3
Ramachandran plot and root-mean-square deviations from ideal geometry for CGL structures

	CGL: M1-2M	CGL: M1-3M	CGLM: M1-4M
Residues distribution			
Most favorable (%)	86.9	86.1	86.6
Allowed (%)	12.6	12.8	12.7
Additionally allowed (%)	0.5	0.5	0.5
Forbidden (%)	0.0	0.6	0.2
Root-mean-square deviations from ideal geometry			
Bond lengths (Å)	0.011	0.021	0.019
Bond angles (°)	1.542	2.267	2.128
Mean temperature factor (B-factor)			
Main chain (Å^2)	11.097	22.194	18.019
Side chain (Å^2)	13.171	22.946	19.231
Water molecules (Å^2)	19.723	27.759	22.184
Sugar atoms (Å^2)[a]	10.160	29.958	19.166
Overall (Å^2)	12.655	24.464	19.65

[a] In CGL:M1-4M complex, subunits C and D were not accounted for the calculate.

Table 4
Final refinement statistics for CGL structures

Refinement	CGL: M1-2M	CGL: M1-3M	CGL: M1-4M
Space group	I222	P32	P32
Unit-cell parameters (Å)	$a = 63.89$, $b = 86.19$, $c = 88.73$	$a = 69.37$, $b = 69.37$, $c = 161.21$	$a = 69.01$, $b = 69.01$, $c = 160.44$
Resolution limit (Å)	21.03 (1.50)	40.29 (2.07)	40.096 (1.980)
R_{factor}	0.215	0.19	0.18
Number of reflexions (R_{factor})	37,554	49,716	56,114
R_{free}	0.232	0.244	0.236
Number of reflexions (R_{free})	1984	2659	2983
Number of residues per asymmetric unit	237	948	948
Number of water molecules	159	412	467

wich (Srinivas et al., 2001). The complexes do not present significant structural differences when compared to native protein in any parts, other than the carbohydrate-binding region.

The carbohydrate electron densities from M1-3M and M1-2M complexes are very well defined in all subunits, while the ones from M1-4M complexes are well defined in only two subunits (A and B). In the remaining subunits from M1-4M complexes (C and D), the electron density corresponding to the O4-linked mannose is not clear, making impossible to model the dimannoside's second ring. All carbohydrate sites from the six structures are located at the interface of crystallographic symmetric related subunits, establishing numerous contacts with their symmetry mates, with the exception of the C and D subunits from M1-4M complexes, which are partially poorly defined as previously mentioned.

Fig. 1 shows the electron densities for the three dimannosides after the final refinement. In the case of M1-3M and M1-4M complexes, which crystallized as a tetramer, the densities shown are the ones with the best definition.

3.2. M1-3M and M1-4M complexes

3.2.1. Monosaccharide-binding site (first mannose)

In all four complexes, the O1-linked mannose interacts with the conserved monosaccharide site in a way very similar to ConA, with very small differences regarding the distances of these interactions, which are probably due to the crystallographic packing. This binding mode has been already extensively studied previously (Bourne et al., 1990, 1994,). Thus, the sugar is harbored through nine intermolecular H bonds formed between O2, O3, O4, O5 and O6 of the saccharide and the amino acid side chains present in the site.

3.2.2. O3 linked mannose and O4 linked mannose

It is well established that legume lectins possess three types of hydrophobic subsites based on different ligand affinities. One of these subsites is adjacent to the conserved monosaccharide biding site, explaining the fact that hydrophobic glyco/mannosides and other monosaccharide hydrophobic derivatives bind more strongly (10–50 times) than their non-hydrophobic analogs (Sharon and Lis, 1990). Kanellopoulos and co-workers (1996) determined the crystal structures for two ConA complexes: ConA/4′-nitrophenyl-α-D-mannopyranoside (α-PNM, PDB code: 1VAM) and ConA/4′-nitrophenil-α-D-glucopyranoside (α-PNG, PDB code: 1VAL), revealing that the hydrophobic moiety of these molecules interacted with a hydrophobic subsite formed by Tyr12, Leu99 and Tyr100. The O3 linked and the O4 linked mannoses of our structures bind to this same subsite, as depicted in Fig. 2. This binding involves primarily hydrophobic and van der Waals interactions (Table 5).

Both dimannosides, M1-3M and M1-4M, are positioned in the subsite in a similar way, with the side chain of Leu99

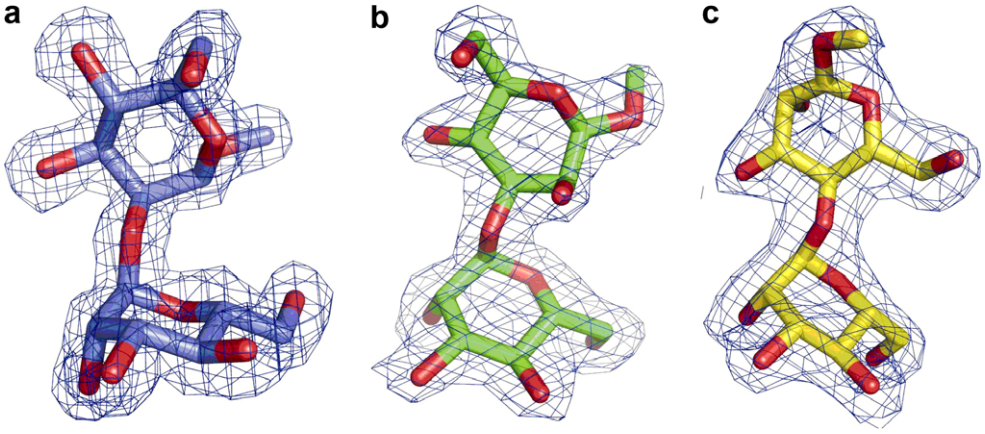


Fig. 1. $2F_{obs} - F_{calc}$ electron densities map from the dimannosides contoured at 1.0σ obtained after the refinement. (a) Man(α1-2)Man(α1-O)Me, (b) Man(α1-3)Man(α1-*O*)Me, (c) Man(α1-4)Man(α1-*O*)Me. Only ConM complexes are shown, CGL complexes are very similar.

stacked against one face of the ring. Despite that, M1-3M is placed more internally, with the anomeric methyl turned to the subsite; evidences for this are both the greater number of van der Waals interactions when compared to the M1-4M and the presence of H bonds in such complexes. This difference can be cleared noticed in Table 5 and in Fig. 2.

Polar interactions, in this case occurring with participation of Tyr12, are only observed for M1-3M. In three subunits, there is an interaction involving the O3 linked mannose and O1 of the sugar and in two, there's an interaction involving O2; in the C subunit, both interactions aforementioned occur simultaneously (Fig. 2). This variation is very likely a consequence of the crystallographic packing. The crystal structures of the complexes of ConA and α-PNG and α-PNM also show this peculiarity, once H bonds occur between Tyr12 and Tyr100 and the O7 and O8 of the carbohydrates, but they vary in both number and strength from subunit to subunit.

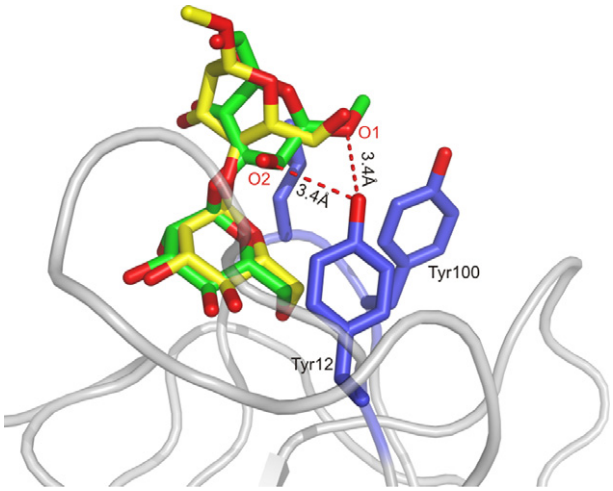


Fig. 2. Superimposition of M1-3M (green) and M1-4M (yellow) in the active site of ConM. The M1-3M more internal localization in the hydrophobic subsite (residues in blue) is evidenced, as well as the hydrogen bonds formed between the Tyr12-OH and O1 and O2. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

The occurrence of H bonds between Tyr12 from ConM and disaccharides has been previously reported to the crystal structures of ConM:Trehalose (PDB code: 2CY6) and ConM:Maltose (PDB code: 2CYF) (Delatorre et al., 2006). This work suggests that the interaction established between the hydroxyl group from Tyr12 and the outer glycose of each complex enhances the lectin ability to bind carbohydrates.

3.3. Comparison between ConM/CGL:M1-3M and ConA:M1-3M

Bouckaert and co-authors (1999) determined the crystal structures of ConA complexed to M1-3M (PDB code: 1QDO) and M1-6M (PDB code: 1QDC). The O3-linked mannose from M1-3M interacts with the subsite formed by Tyr12, Tyr100 and Leu99, while the O6-linked mannose from M1-6M interacts with other subsite formed by Tyr12 and Asp16. In ConA, M1-3M only establishes van der Waals and hydrophobic interactions, while in ConM and CGL it forms two H bonds to Tyr12. Besides that, M1-3M have more van der Waals interactions with ConM's and CGL's subsite than with ConA's (Table 5).

The occurrence of H bonds in ConM and CGL complexed with M1-3M, which are absent in ConA:M1-3M, the greater number of van der Waals contacts and a larger contact surface area upon complexation (ConM: 282 Å^2, CGL: 282 Å^2, ConA: 275 Å^2) should enhance ConM's and CGL's affinity to this dimannoside. In the same way, it is expected that ConM and CGL show greater affinity for M1-3M than for M1-4M (ConM: 261 Å^2, CGL: 261 Å^2). Nevertheless, appearance of extra interactions and a greater contact area do not necessarily imply a higher affinity (Singha et al., 1996). That is the case for the complexes ConA:M1-3M and ConA:M1-6M, where despite M1-6M hydrogen bonds with the

Table 5
Van der Waals contacts of the reducing mannose of following complexes

ConM:M1-3M	ConM:M1-4M	ConA:M1-3M
Tyr12-**CE1—O2**	Tyr12-**OH—C6**	Tyr12 -**OH —C1, C2**
Tyr12-**CZ—O2**	Leu99-**CB—O4, O2, O3**	Leu99-**CG —O1**
Tyr12-**OH—C1, C2**	Leu99-**CG—O2**	Leu99-**CD2—O4, O1**
Leu99-**CB—O3**	Leu99-**CD1—O2**	Tyr100-**CE2—O1**
Leu99-**CG—O4**	Leu99-**CD2—O2**	Tyr100-**OH—C7**
Leu99-**CD1—O1, O4**		
Tyr100-**CZ—O1**		
Tyr100-**OH—CM**		
Tyr100-**CE2—O1**		

hydroxyl group from Tyr12 and has greater buried surface area, M1-3M interacts with a 4 times higher affinity. Similarly, the lectin from *Pterocarpus angolensis* (PAL) presents greater number of interactions to the dimannosides M1-3M, M1-4M and M1-6M than to MeMan, but its affinity is essentially equal for all four sugars (Loris et al., 2004). Thus, to establish real affinity comparisons for the complexes presented here, thermodynamics studies are necessary.

As previously described for ConA, M1-3M's and M1-4M's conformations are essentially equal in PAL, ConM and CGL, despite the different residues forming the sites. In the case of M1-3M complexes, the sugar is slightly dislocated, which happens due to the presence of a Glu221 on PAL in the same place where there is a Leu99 on ConA, ConM and CGL. If the M1-3M conformation in canavalia structures was similar to the one presented on PAL, it would cause a van der Waals clash between Leu99's CD2 and the second's mannose O1. This M1-3M conformation adopted for PAL also allows for OE1 from Glu221 to establish an H bond to the O4 from the second mannose and possibly a polar contact to the O1 from the second mannose.

In the case of the M1-4M complex, the tighter closed site from PAL allows for the establishment of an H bond between the O1 from the second mannose and the Asp136 that occupies the position corresponding to Asp16 from ConM, CGL and ConA. Besides, this conformation is stabilized by a second H bond between O6 from the second mannose and the NE2 from Gln22, which occupies the same position as Tyr100 from ConM, CGL and ConA.

3.4. Waters involved in carbohydrate-binding site

Previous crystallographic study performed with legume lectins determined the occurrence of seven waters highly conserved in all structures, independently of their homology or carbohydrate affinity. A structural function is evident: four waters are linked with calcium and manganese ions and one of them is located at carbohydrate-binding site (Loris et al., 1994).

A structural water at the active site occurs in all ConA-like structures containing at least a monosaccharide bound. This water makes hydrogen bonds with Arg228, Asn14 and Asp16, but not with the sugar. Bouckaert and co-workers (1999) propose that the function of this water is helping to increase the fit between the protein and sugar, rather than to serve as specific hydrogen bond mediators, giving the protein-binding surface a shape complementary to the sugar.

In ConM/CGL:M1-2 M, ConA:M1-2M (PDB code: 1BXH) and ConA:M1-6M (PDB code: 1QDC) this water is perfectly conserved. However, in ConM and CGL complexes with M1-3M and M1-4M its location and interactions patterns are different; there is an additional water molecule making a hydrogen bond with the conserved structural water and Asp16 (Fig. 3a). The reason for this seems to be a difference in conformation of Asp16, which

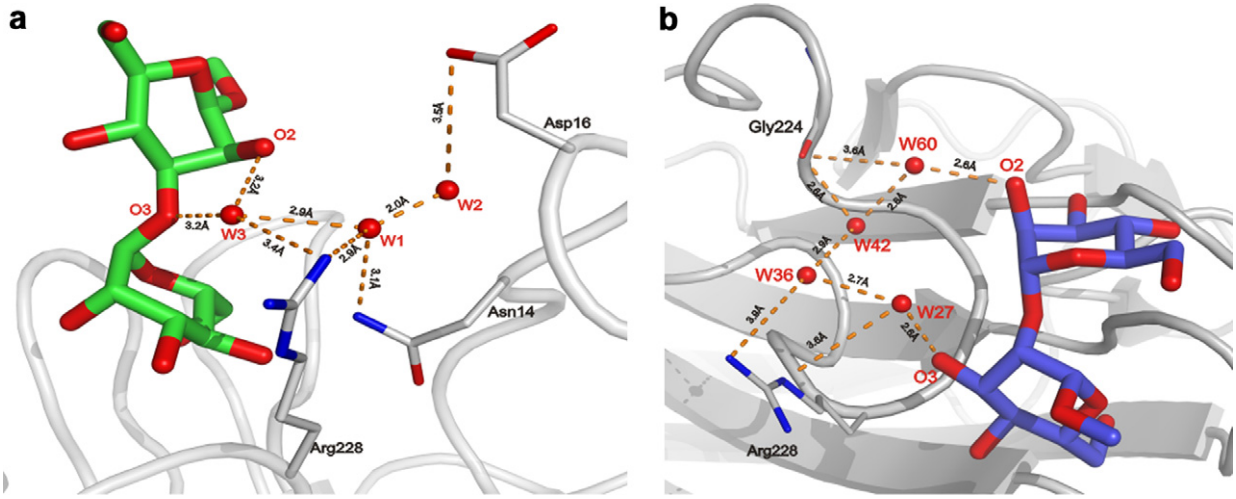


Fig. 3. (a) Representation of the structural waters (W1 and W2) in ConM:M1-3M. W1 and W2 are also present in M1-4M complexes. The interstitial water W3 was observed only in M1-3M complexes and its interactions are expected to enhance the affinity of the protein by this dimannoside. (b) Waters involved in the active site of ConM:M1-2M complex (also observed in CGL complex). This network of interactions formed by the interstitial waters W60 and W27, as well as by structural waters W42 and W36 should contribute to the enhanced affinity calculated for M1-2M.

possess a "turned up" orientation in the complexes containing this additional water.

In M1-3M structures, all the subunits possess the structural water (W1), interacting with Arg228 and Asn14 but only indirectly with Asp16, through W2, which is present in two monomers of ConM (A and B) and of CGL (B and D). ConM:M1-4M complex present W1 and W2 all monomers, with the exception of subunit C, that lacks density for W2. In CGL:M1-4M, the W1 is present in three subunits (A, B and D) while W2 was only detected in subunit A.

These differences in water arrangement among the complexes can reflect an alteration in the affinity for the sugar (Mary et al., 1995), with implications in biological and physiological activities of the protein, or can be simply a consequence of crystallographic packing, since Asp16 have different conformations in the structures.

In M1-3M complexes, three monomers possess an insterstitial water (W3), linking indirectly the Arg228 to O3 and O2 of the dimannoside. Its interaction with W1 characterizes a small network of contacts, which must contribute for an enhanced affinity in the biding of these proteins to M1-3M (Fig. 3a).

3.5. M1-2M(Me) complexes

3.5.1. O2 linked mannose

Unlike the M1-3M and M1-4M complexes, the second ring, the O2 linked mannose (called here as reducing mannose), interacts with the monosaccharide-binding site. In spite of minor modifications, the interaction pattern for this interaction is compatible with the observed for all legume lectins. The Goldstein Rules explain this alternative recognition: the requirements for carbohydrate binding to the monosaccharide-binding site are free equatorial hydroxyl groups at position 3 and 4 and a free primary alcohol group in the 6 position (Goldstein et al., 1974). Therefore, both sugar rings from M1-2M(Me) can be potentially recognized by the monosaccharide-binding site.

3.6. Comparison between ConM/CGL:M1-2M(Me) and ConA:M1-2M complexes

3.6.1. M1-2M(Me) interaction manners

The ConA:M1-2M(Me) crystallographic structure revealed two binding modes to this dimannoside (PDB code: 1BXH). In the subunit A, the O1 linked mannose interacts with the monosaccharide-binding site, while in the subunit D, the O2 linked mannose interacts with this site (reducing mannose). In the subunit B, the two distinct overlapping modes of binding for the disaccharide are present (Moothoo et al., 1999).

3.6.2. Subunit D from ConA:M1-2M(Me)

The interaction of ConM/CGL and M1-2M(Me) is similar to that observed for ConA:M1-2M(Me) subunit D (Moothoo et al., 1999), where the O1 linked mannose lies in a polar site. In ConM, occur hydrogen bonds with Gly98, Leu99, Ser168 and Thr226. In addition, van der Waals interactions with Leu99 are important to stabilize the molecule. In the Subunit D of ConA:M1-2M(Me), the second ring is slightly rotated and there is no hydrogen bond between Gly98-O and Leu99-N with the sugar, in spite of them possess the same orientation (Fig. 4).

In the PAL:M1-2M structure, the dimannoside presents the same orientation verified for the subunit D from ConM, CGL and ConA:M1-2M(Me), although they possess only 48% of similarity. A slight rotation in the saccharide occurs due the presence of a Thr226 in the Canavalia structures, in the same region where PAL possesses Gly104. As a consequence of this M1-2M binding mode, these proteins belong to the select group that makes interactions with the conserved stretch E and the sugar (Loris et al., 2004).

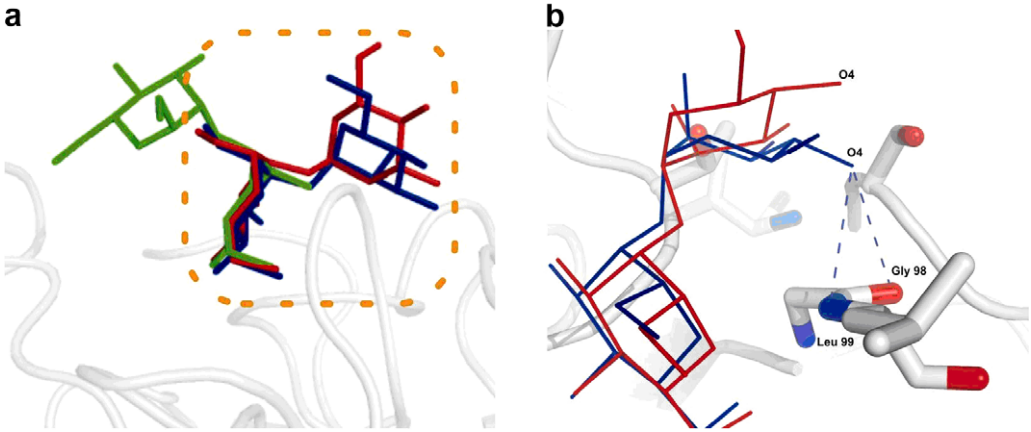


Fig. 4. (a) Superimposition of M1-2M conformations observed in ConM/CGL and in ConA. Green: M1-2M(Me) in subunit A of ConA (PDB code: 1BXH). Red: M1-2M(Me) in the subunit of ConA (PDB code: 1BXH). Blue: M1-2M(Me) in the ConM and CGL complexes. (b) Comparison of the M1-2M(Me) conformations in the ConM/CGL complexes and in the subunit D of ConA. Although very similar, an slight rotation allow the formation of two hydrogen bonds (Leu99 and Gly98 with O4) in ConM/CGL complexes. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

3.6.3. Subunit A from ConA:M1-2M(Me)

The superimposition of ConM/CGL:M1-2M(Me) and ConA:M1-2M(Me) shows that the ConA interaction mode is sterically permitted in ConM, maximizing the interactions between anomeric methyl group and the hydrophobic subsite formed by Tyr12, Leu99 and Ty100, as well as C6 atom of the sugar with the protein.

The crystallographic complex obtained with the *Bowringia milbraedii* lectin (BMA), a protein sharing 48% of similarity with ConA, and Man(α1-2)Man revealed an interaction very similar to that observed for the ConA:M1-2M(Me) subunit A. The only difference is a hydrogen bond between O1 of the reducing mannose and the Gln218 side chain, which should compensate the van der Waals anomeric methyl group interactions present in ConA (Buts et al., 2006).

3.6.4. ConA:M1-2M and waters involved in carbohydrate-binding site

In 2001, Sanders et al. obtained the ConA:Man(α1-2)Man complex (PDB code: 1I3H) at high resolution (1.2 Å). In this structure, the O1 linked mannose binds to the monosaccharide-binding site and the O2 linked mannose interacts to a site formed by Tyr12, Asp16 and Arg228, constituting a distinct mode of M1-2M recognition by ConA. The high resolution of the complex allowed the precise localization of waters involved in carbohydrate-binding site, and of two conformations for Arg228. The ConM:M1-2M(Me) structure also obtained in high resolution (1.4 Å) permits a detailed comparison of the net of waters presents at the biding site.

As observed for the ConA:M1-2M complex, a net of waters bind the sugar to ConM/CGL (Fig. 3b). The water 60 links the O2 of the O1 linked mannose with Gly224. In the same way, the water 27 mediates the interaction between Arg228 and the O3 of O2 linked mannose. Two other water molecules, W36 and W42, seem to perform a structural function, giving to the binding site a shape complementary to the sugar. In addition, they also interact to Gly224, Arg228 and Thr226. This network should result in an enhancement in the dimannoside affinity. Contrasting with ConA:M1-2M, the structural conserved water interacts only with Asn14, Asp16 and Arg228, as conserved for several ConA-like structures complexes with sugars. In CGL, electron density is not observed for W27.

3.6.5. Thermodynamic aspects

ConA binds to M1-2M 3 times stronger than to M1-3M and M1-4M and binds to M1-2M(Me) 3.4 times stronger than to M1-2M. Sanders et al. (2001) propose that the two binding modes of interactions observed for the complex ConA:Man(α1-2)Man(α1-O)Me are responsible for this affinity increase.

PAL and BMA present highest affinity for M1-2M, and only one mode of interaction was verified for these structures. The CGL and ConM:M1-2M structures, although methylated also presents only one interaction mode with M1-2M. In this way, thermodynamics analysis for this protein and the dimannosides in subject should complement this study and permit to elucidate which factors are important to the interaction between ConM, CGL and ConA with Man(α1-2)Man(α1-O)Me.

3.7. The influence of Histidine-205 in the conformation acquired by Man(α1-2)Man(α1-O)Me when interacting with ConA and ConM

ConA:Man(α1-2)Man(α1-O)Me structure presents one mode of interaction in subunit A and a second mode in subunit D, which is the only one observed in for the M1-2M complexes reported here. Although the reason for that remains obscure, the ConM and CGL complexes points toward a likely responsible residue.

In ConA:M1-2M(Me) (subunit D) and ConM/CGL:M1-2M(Me), His205 has a "turned up" conformation, reaching Tyr100, while in ConA:M1-2M(Me) (subunit A), His205 presents a "turned down" conformation, reaching Tyr12. The His205 orientation should provoke conformational changes in the active site, favoring a determined interaction mode (Fig. 5).

Unfortunately, these residues are located at points of crystal contacts, and the precise changes caused by His205 cannot be defined. We proposed that the "turned down" His205 conformation causes an approximation among Tyr12, Tyr100 and Leu99 and the second mannose, maximizing the van der Waals and hydrofobic interactions between the anomeric methyl group and the hydrophobic subsite, the same feature observed in ConA subunit A. In "turned up" His205 conformation (ConA subunit D,

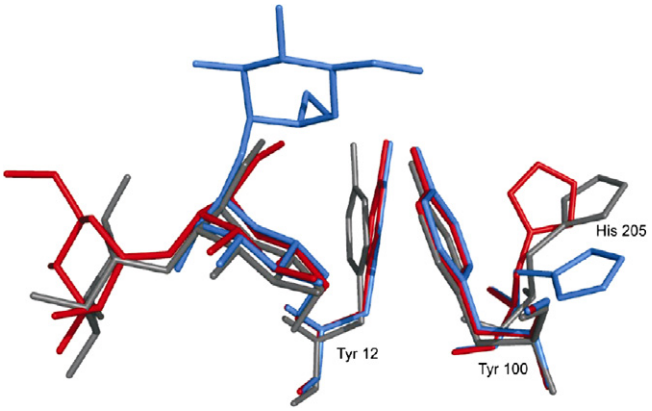


Fig. 5. Superimposition of ConA subunit A (blue), ConA subunit D (red) and ConM (gray) complexed with M1-2M(Me). A double conformation adopted by His205 would be responsible by the two modes of interaction observed for M1-2M(Me). When His205 is "turned down", it provokes an approximation between the hydrofobic subsite (Tyr12, Leu99 and Tyr100) and the second mannose, maximizing the van der Waals interactions of the anomeric methyl group (C7) and the hydrofobic subsite. When His205 is "turned up", the more stable conformation is acquired through polar interactions with Thr226 and Ser168. (For interpretation of the references in color in this figure legend, the reader is referred to the web version of this article.)

ConM and CGL), the residues from hydrophobic subsite do not approximate enough to the second mannose, and the latter acquires a conformation where the four hydrogen bonds formed with Ser168 and Thr226 must compensate for the anomeric methyl group interactions. Therefore, a possible double conformation for His205, as previously identified for Arg228 in the ConA:M1-2M complex and Thr15 in the native ConA at 1.15 Å (PDB code: 1JBC) seems to offer a coherent explanation for the two interactions modes observed in the ConA:Man(α1-2)Man(α1-O)Me.

3.8. Interaction lectin–carbohydrate and its relationship with biological activities

In spite of high degree of similarity (98%), ConA and ConM present significant differences in their biological properties. For instance, ConA and ConM promote endothelium-dependent relaxant effects on aorta, however, ConM possesses a much higher potency than ConA (3 times), resulting in a release of nitric oxide (Gadelha et al., 2005). A recent investigation from our group revealed that in a 125 μg/mL concentration, ConA inhibited 100% the growth of *Streptococcus mutans*and *Streptococcus sanguis*, while ConM was incapable of any inhibition at the same concentration. On the other hand, ConM at 100 μg/mL was able to inhibit significantly the formation of biofilm from the *S. mutans* and *S. sanguis*, while ConA was incapable of inhibition at any concentration tested (*unpublished results*).

Our data show that ConA and ConM have different patterns of interactions with dimannosides in subject, once these dissacharides are present in asparagine linked (N-linked) carbohydrates which must be exposed at cell surface receptors for lectins, these results point toward structural insights for the distinct biological activities verified for the two proteins.

Isothermal titration microcalorimetry studies (Dam et al., 1998) divided Diocleinae lectins in two groups, according to the affinity for biantennary complex oligosaccharides and related to their histamine-releasing properties. ConA belongs to the group with relatively high Kα and ΔH values, which have strong histamine-releasing activity. On the other hand, ConM belongs to the lower relatively Kα and ΔH values group, showing little histamine-releasing effect. Once the recognizing of the N-linked glycans is centered on the trimannoside core (Naismith and Field, 1996; Loris et al., 1996), which possesses Manα1-3Man as a highly specific ConA epitope, our results present a structural explanation regarding these affinity differences.

Atomic coordinates for the structures discussed in this work were deposited in the PDB, with the following PDB access codes:

CGL:M1-2M—2OVU ConM:M1-2M—2OW4,
CGL:M1-3M—2EF6 ConM:M1-3M—2P37,
CGL:M1-4M—2P2K ConM:M1-4M—2P34.

References

Andrade, J.L., Arruda, S., Barbosa, T., Paim, L., Ramos, M.V., Cavada, B.S., Barral-Neto, M., 1999. Lectin-induced nitric oxide production. Cell. Immunol. 194, 98–102.

Barbosa, T., Arruda, S., Cavada, B.S., Grangeiro, T.B., Freitas, L.A.R., Barral-Netto, M., 2001. In vivo lymphocyte activation and apoptosis by lectins of the diocleinae subtribe. Mem. Inst. Oswaldo Cruz 96, 673–678.

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., 1992. Human lymphocyte stimulation by legume lectins from Diocleinar tribe. Immunol. Invest. 21, 297–303.

Bento, C.A.M., Cavada, B.S., Oliveira, J.T.A., Moreira, R.A., Barja-Fidalgo, C., 1993. Rat paw edema and leukocyte immigration induced by plant lectins. Agent Actions 38, 48–54.

Bouckaert, J., Hamelryck, T.W., Wyns, L., Loris, R., 1999. The crystal structures of Man(α1–3)Man(α1-o)Me and Man(α1–6)Man(α1-o)Me in complex with Concanavalin A. J. Biol. Chem. 274, 29188–29195.

Bourne, Y., Roussel, A., Frey, M., Rougé, P., Fontecilla-Camps, J.C., Cambillau, C., 1990. Three-dimensionnal structures of complexes of Lathyrus ochrus isolectin I with glucose and mannose: fine specificity of the monosaccharide-binding site. Prot. Struct. Func. Gen. 8, 365–376.

Brinda, K.V., Mitra, N., Surolia, A., Vishveshwara, S., 2004. Determinants of quaternary association in legume lectins. Protein Sci. 13, 1735–1749.

Brunger, A.T., 1992. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature 355, 472–475.

Buts, L., Garcia-Pino, A., Wyns, L., Loris, R., 2006. Structural basis of carbohydrate recognition by a Man(a1-2)Man specific lectin from *Bowringia milbraedii*. Glycobiology 16, 635–640.

CCP4, 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallog. Sect. D, 50, 760–763.

Dam, T.K., Cavada, B.S., Grangeiro, T.B., Santos, C.F., Sousa, F.A.M., Oscarson, S., Brewer, C.F., 1998. Diocleinae lectins are a group of proteins with conserved binding sites for the core trimannoside of asparagine-linked oligosaccharides and differential specificities for complex carbohydrates. J. Biol. Chem. 273, 12082–12088.

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

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., 2006. Crystal structure of a lectin from *Canavalia maritima* (ConM) in complex with trehalose and maltose reveals relevant mutation in ConA-like lectins. J. Struct. Biol. 154, 280–286.

Ferreira, R.R., Cavada, B.S., Moreira, R.A., Oliveira, J.T.A., Gomes, J.C., 1996. Characteristics of the histamine release from hamster cheek pouch mast cells stimulated by lectins from Brazilian beans and concanavalin A. Inflamm. Res. 45, 442–447.

Gabius, H.J., Gabius, S., 1997. Glycoscience: Status and Pespectives. Chapman & Hall, Weinheim.

Gadelha, C.A.A., Moreno, F.B.M.B., Santi-Gadelha, T., Cajazeiras, J.B., Da 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., De Azevedo Jr., W.F., Cavada, B.S., 2005. Native crystal structure of a nitric oxide-releasing lectin from the seeds of *Canavalia marítima*. J. Struct. Biol. 152, 185–194.

Goldstein, I.J., Reichert, C.M., Misaki, A., 1974. Interaction of concanavalin A with model substrates. Ann. NY Acad. Sci. 234, 283–295.

Gomes, J.C., Rossi, R.R., Cavada, B.S., Moreira, R.A., 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.

Hubbard, S.J., Thornton, J.M., 1993. 'NACCESS', Computer Program. Department of Biochemistry and Molecular Biology, University College, London.

Kanellopoulos, P.N., Pavlou, K., Perrakis, A., Agianian, B., Vorgias, C.E., Mavrommatis, C., Soufi, M., Tucker, P.A., Hamodrakas, S.J., 1996. The crystal structure of the complexes of concanavalin A with 4′-nitrophenyl-alpha-D-mannopyranoside and 4′-nitrophenylalpha-D-glucopyranoside. J. Struct. Biol. 116, 345–355.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. Main-chain bond lengths and bond angles in protein structures. J. Appl. Cryst. 26, 283–291.

Lis, H., Sharon, N., 1986. Lectins as molecules and as tools. Annu. Rev. Biochem. 55, 35–67.

Loris, R., Casset, F., Bouckaert, J., Pletinckx, J., Dao-Thi, M., Poortmans, F., Imberty, A., Pérez, S., Wyns, L., 1994. The monosaccharide binding site of lentil lectin: an X-ray and molecular modelling study. Glycoconj. J. 11, 507–517.

Loris, R., Stasnll, P.P.G., Wyns, L., 1994. Conserved waters in legume lectin crystal structures. J. Biol. Chem. 269, 26722–26733.

Loris, R., Maes, D., Poortmans, F., Wyns, L., Bouckaert, J., 1996. A structure of the complex between concanavalin a and methyl-3,6-di-o-(a-d-mannopyranosyl)-a-d-mannopyranoside reveals two binding modes. J. Biol. Chem. 271, 30614–30618.

Loris, R., Walle, I.V., De Greve, H., Beeckmans, S., Deboeck, F., Wyns, L., Bouckaert, J., 2004. Structural basis of oligomannose recognition by the Pterocarpus angolensisseed lectin. J. Mol. Biol. 335, 1227–1240.

Mary, C., Chervenak, M.C., Toone, E.J., 1995. Calorimetric analysis of the binding of lectins with overlapping carbohydrate-binding ligand specificities. Biochemistry 34, 5685–5695.

Mcree, D.E., 1999. XtalView/Xfit—A versatile program for manipulating atomic Coordinates and Electron Density. J. Struct. Biol. 125, 156–165.

Moothoo, D.N., Canan, B., Field, R.A., Naismith, J.H., 1999. Man alpha1-2 Man alpha-OMe-concanavalin A complex reveals a balance of forces involved in carbohydrate recognition. Glycobiology 9, 539–545.

Moreno, F.B.M.B., Bezerra, G.A., Oliveira, T.M., Souza, E.P., Rocha, B.A.M., Benevides, R.G., Delatorre, P., Cavada, B.S., Azevedo Jr., W.F., 2006. New crystal forms of Diocleinae lectins in the presence of different dimannosides. Acta Crystallog. F 62, 1100–1103.

Naismith, J.H., Field, R.A., 1996. Structural basis of trimannoside recognition by concanavalin A. J. Biol. Chem. 271, 972–976.

Naismith, J.H., Emmerich, C., Habash, J., Harrop, S.J., Helliwell, J.R., Hunter, W.N., Raftery, J., Kalb (Gilboa), A.J., Yariv, J., 1994. Refined structure of concanavalin A complexed with α-methyl-D-mannopyranoside at 2.0 Å resolution and comparison with the saccharide free structure. Acta Cryst. D 50, 847–858.

Rodriguez, D., Cavada, B.S., Oliveira, J.T.A., Moreira, R.A., Russo, M., 1992. The carbohydrate-binding specificity and molecular modelling of *Canavalia maritima* and *Dioclea grandifloralectins*. Br. J. Med. Biol. Res. 25, 823–826.

Sanders, D.A.R., Mothoo, D.V., Raftery, J., Howard, A.J., helliwell, J.R., Naismith, J.H., 2001. The 1.2 A resolution structure of the cona-dimannose complex. J. Mol. Biol. 310, 875–884.

Sanz-Aparicio, J., Hermoso, J., Graneiro, T.B., Calvete, J.J., Cavada, B.S., 1997. The crystal structure of *Canavalia brasiliensis* lectin suggest a correlation between its quaternary conformation and its distinct biological properties from concanavalin A. FEBS Lett. 405, 114–118.

Sharon, N., Lis, H., 1990. Legume lectins—a large family of homologous proteins. FASEB J. 4, 3198–3208.

Singha, N.C., Surolia, N., Surolia, A., 1996. Titration calorimetric studies to elucidate the specificity of the interactions of polymyxin B with lipopolysaccharides and lipid A. Biosci. Rep. 16, 1–10.

Srinivas, V.R., Reddy, G.B., Ahmad, N., Swaminathan, C.P., Mitra, N., Surolia, A., 2001. Legume lectin family, the ''natural mutants of the quaternary state'', provide insights into relationship between protein stability and oligomerization. Biochim. Biophys. Acta 1527, 102–111.

Vargin, A., Teplyakov, A., 1997. MOLREP: an automated program for molecular replacement. J. Appl. Crystallog. 30, 1022–1025.

Wah, D.A., Romero, A., Gallego del Sol, F., Cavada, B.S., Ramos, M.V., Grangeiro, T.B., Sampaio, A.H., 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.

the FEBS Journal

cDNA cloning and 1.75 Å crystal structure determination of PPL2, an endochitinase and *N*-acetylglucosamine-binding hemagglutinin from *Parkia platycephala* seeds

Benildo S. Cavada[1], Frederico B. B. Moreno[2], Bruno A. M. da Rocha[1,3], Walter F. de Azevedo Jr[4], Rolando E. R. Castellón[1], Georg V. Goersch[1], Celso S. Nagano[5], Emmanuel P. de Souza[1], Kyria S. Nascimento[1], Gandhi Radis-Baptista[1], Plínio Delatorre[3], Yves Leroy[6], Marcos H. Toyama[7], Vicente P. T. Pinto[8], Alexandre H. Sampaio[9], Domingo Barettino[5], Henri Debray[6], Juan J. Calvete[5] and Libia Sanz[5]

1 BioMol-Laboratory, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil
2 Departamento de Física, Universidade Estadual Paulista, UNESP, São José do Rio Preto, São Paulo, Brazil
3 Departamento de Ciências Físicas e Biológicas, Universidade Regional do Cariri, Fortaleza, Ceará, Brazil
4 Faculdade de Biociências, Centro de Pesquisas em Biologia Molecular e Funcional, PUCRS, Porto Alegre, Rio Grande do Sul, Brazil
5 Instituto de Biomedicina de Valencia, CSIC, Spain
6 Laboratoire de Chimie Biologique et Unité Mixte de Recherche No. 8576 du CNRS, Université des Sciences et Technologies de Lille, France
7 Departamento de Bioquímica, Instituto de Biologia, Universidade Estadual de Campinas (UNICAMP), Campinas, SP, Brazil
8 Faculdade de Medicina, Universidade Federal do Ceará, Sobral, Brazil
9 Laboratorio de Bioquímica Marinha, Departamento de Engenharia de Pesca, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil

Keywords
endochitinase; glycosyl hydrolase family 18; Mimosoideae; *Parkia platycephala*; X-ray crystal structure

Correspondence
B. S. Cavada, BioMol-Laboratory, Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil
Fax/Tel: +55 8540089818
E-mail: bscavada@ufc.br
H. Debray, Laboratoire de Chimie Biologique et Unité Mixte de Recherche du CNRS N°8576, Université des Sciences et Technologies de Lille, bâtiment C-9, 59655 Villeneuve D'Ascq Cedex
Fax: +33 320436555
Tel: +33 320410108
E-mail: henri.debray@univ-lille1.fr

(Received 22 May 2006, revised 26 June 2006, accepted 28 June 2006)

doi:10.1111/j.1742-4658.2006.05400.x

Parkia platycephala lectin 2 was purified from *Parkia platycephala* (Leguminosae, Mimosoideae) seeds by affinity chromatography and RP-HPLC. Equilibrium sedimentation and MS showed that *Parkia platycephala* lectin 2 is a nonglycosylated monomeric protein of molecular mass 29 407 ± 15 Da, which contains six cysteine residues engaged in the formation of three intramolecular disulfide bonds. *Parkia platycephala* lectin 2 agglutinated rabbit erythrocytes, and this activity was specifically inhibited by *N*-acetylglucosamine. In addition, *Parkia platycephala* lectin 2 hydrolyzed β(1–4) glycosidic bonds linking 2-acetoamido-2-deoxy-β-D-glucopyranose units in chitin. The full-length amino acid sequence of *Parkia platycephala* lectin 2, determined by N-terminal sequencing and cDNA cloning, and its three-dimensional structure, established by X-ray crystallography at 1.75 Å resolution, showed that *Parkia platycephala* lectin 2 is homologous to endochitinases of the glycosyl hydrolase family 18, which share the $(\beta\alpha)_8$ barrel topology harboring the catalytic residues Asp125, Glu127, and Tyr182.

Abbreviations
CTAB, cetyl triethylammonium bromide; GlcNac, N-acetyl-D-glucosamine; GSP, gene-specific forward primer; HPAEC-PAD, high-pH anion exchange chromatography with pulsed amperometric detection; PE, pyridylethylated; PPL1, *Parkia platycephala* lectin 1; PPL2, *Parkia platycephala* lectin 2; PTC, phenylisothiocyanate; PTH, phenylthiohydantoin.

Lectins comprise a heterogeneous class of (glyco)proteins that possess one noncatalytic domain that binds carbohydrates in a specific and reversible manner without altering their covalent structure [1]. Lectins decipher the glycocodes encoded in the structure of glycans in processes such as cell communication, host defense, fertilization, development, parasitic infection, tumor metastasis, and plant defense against herbivores and pathogens [2]. Mechanisms for sugar recognition have evolved independently in a restricted number of protein folds (e.g. jelly roll domain, C-type lectin fold, β-propeller, β-trefoil motif, β-prism I and II domains, Ig domains, β-sandwich, mixed αβ structure, and hevein domain) [1,3] (for a complete catalog of carbohydrate-binding protein domains, please consult the 3D Lectin Database at http://www.cermav.cnrs.fr/lectines). In plants, most of the currently known lectins can be placed in seven families of structurally and evolutionarily related proteins [1]. The seed lectins of leguminous plants constitute the largest and most thoroughly studied lectin family. These lectins have represented paradigms for establishing the structural basis [4–9] and thermodynamics [10–13] of selective sugar recognition.

Most studies on lectins from Leguminosae involve members of the Papilionoideae subfamily, whereas investigations on lectins of the other two subfamilies, Caesalpinoideae and Mimosoideae, are scarce. Indeed, to date, the only lectins from the Mimosoideae that have been functionally and structurally characterized are those from seeds of species of the genus *Parkia*, including *Parkia speciosa* [14], *Parkia javanica* [15], *Parkia discolor* [16] and the glucose/mannose-specific lectin from *Parkia platycephala* seeds [17–21]. *Parkia* (Leguminosae, Mimosoideae), regarded as the most primitive group of leguminous plants [22], is a pantropical genus of trees comprising about 30 species found in the neotropics from Honduras to south-eastern Brazil, West Africa, the northern part of Malaysia and the south of Thailand. *Parkia platycephala* is an important forage tree growing in parts of north-eastern Brazil. The seed lectin from *Parkia platycephala* is a 47.9-kDa single-chain nonglycosylated mosaic protein composed of three tandemly arranged jacalin-related β-prism domains [19,20].

The sugar-binding specificity of *Parkia platycephala* lectin towards mannose, an abundant building block of surface-exposed glycoconjugates of viruses, bacteria, and fungi, suggests a role for the *Parkia platycephala* lectin in defense against plant pathogens [1]. Moreover, the *Parkia platycephala* lectin also shows sequence similarity with stress-upregulated and pathogen-upregulated defense genes of a number of different plants, suggesting a common ancestry for jacalin-related lectins and inducible defense proteins [19]. In addition to using lectins, whose precise role in plant defense remains to be determined [23,24, and references cited], plants defend themselves against pathogens (i.e. fungi) secreting pathogenesis-related enzymes, such as xylanases and chitinases, which degrade the pathogen's cell wall [25–27]. In a previous article we have reported the presence of an endochitinase in *Parkia platycephala* seeds [28]. Now, we have determined its complete amino acid sequence by a combination of Edman degradation and cDNA cloning, and report its biochemical characterization and the determination of its crystal structure. Our results show that this protein, termed *Parkia platycephala* lectin 2 (PPL2), is homologous to endochitinases of the glycosyl hydrolase family 18 that exhibit rabbit erythrocyte-agglutinating, *N*-acetylglucosamine-binding and chitin-hydrolyzing activities.

Results and Discusion

PPL2, a nonglycosylated and monomeric GlcNAc-binding hemagglutinin

PPL2 was purified from *Parkia platycephala* seeds by affinity chromatography on either Red-Sepharose (Fig. 1A) or chitin-Sepharose. The protein agglutinated trypsin-treated rabbit erythrocytes (128 hemagglutinating units mg^{-1}), and this activity was abolished by 19 mM *N*-acetyl-D-glucosamine (GlcNac). Other sugars, such as glucose, mannose, galactose, fucose and *N*-acetyl-D-galactosamine, displayed only partial hemagglutination inhibitory activity at much higher concentrations (> 75 mM) than GlcNac. Moreover, the glycoproteins bovine thyroglobulin, ovine submaxillary mucin, bovine fetuin and bovine asialofetuin were devoid of hemagglutination inhibitory activity. Bovine thyroglobulin contains nine complex glycosylation sites and four high-mannose oligosaccharides [29]. Ovine submaxillary mucin is a glycoprotein bearing a high density of O-linked oligosaccharides expressing sialyl Tn antigens and sialyl core 3 sequences [30]. Bovine fetuin contains three N-linked glycosylation sites occupied with trisialylated, tetrasialylated or pentasialylated trianntennary structures, and three monosialylated or disialylated O-linked saccharides [31–33]. We thus concluded that PPL2 represented an *N*-acetylglucosamine-binding hemagglutinin.

The apparent molecular masses of both native and reduced PPL2 determined by SDS/PAGE were 30 kDa (Fig. 1A, insert). The molecular mass of native PPL2, measured by MALDI-TOF MS, was 29 407 ± 15 Da (Fig. 1A). This value was not altered upon incubation of the denatured, but nonreduced,

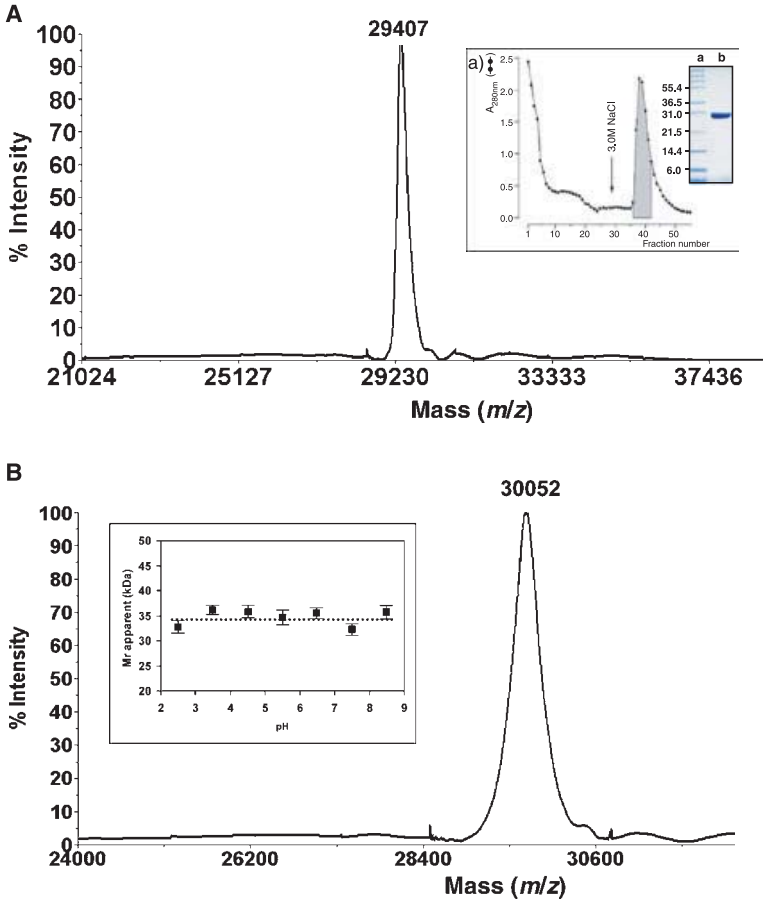


Fig. 1. Purification and molecular mass determination of PPL2. (A) MALDI-TOF mass determination of native PPL2 purified by affinity chromatography as illustrated in the insert. Insert: the fraction of a *Parkia platycephala* seed homogenate precipated with 60% saturation ammonium sulfate was resuspended in 50 mM Tris, pH 7.0, containing 100 mM NaCl, and applied to a Red-Sepharose column. Retained material was eluted with 3 M NaCl. Fractions exhibiting hemagglutinating activity (gray area) were pooled. Right panel: SDS/PAGE of the pooled hemagglutinin termed PPL2. Lane a, molecular mass makers: glutamic dehydrogenase (55.4 kDa), lactate dehydrogenase (36.5 kDa), carbonic anhydrase (31.0 kDa), trypsin inhibitor (21.5 kDa), lysozyme (14.4 kDa), aprotinin (6.0 kDa). Lane b, reduced PPL2. (B) MALDI-TOF mass determination of reduced and pyridylethylated PPL2. Insert: apparent molecular masses of native PPL2 determined by equilibrium sedimentation analytic centrifugation in solutions with different pH values.

protein with the alkylating reagent 4-vinylpyridine. On the other hand, the same treatment after reduction of the protein with dithiothreitol changed the molecular mass of PPL2 to 30 052 ± 15 Da (Fig. 1B). The mass increment of about 645 Da indicated that PPL2 had incorporated six pyridylethyl groups. The combined data clearly showed that PPL2 contained six cysteine residues engaged in the formation of three intramolecular disulfide bonds. Amino acid compositional analysis of the purified protein (Table 1) was in agreement with this conclusion.

The estimated apparent molecular mass for PPL2 on a calibrated size-exclusion chromatographic column was 12 kDa, indicating that the protein had an anomalous elution profile. Molecular mass determinations by size-exclusion chromatography are dependent on the hydrodynamic properties of the molecule, and, in addition, interaction of the protein with the matrix may also introduce large errors into the estimated molecular mass. Thus, we carried out a more rigorous analysis of the aggregation state of PPL2 employing

Table 1. Amino acid composition [mol·(mol protein)$^{-1}$] of PPL2. Asx, aspartic acid and asparagine; Glx, glutamic acid and glutamine.

Amino acid	PPL2
Asx	34
Glx	16
Gly	22
Ser	27
His	2
Arg	5
Thr	13
Ala	20
Pro	11
Tyr	9
Val	12
Met	1
Cys	6
Ile	13
Leu	23
Phe	11
Lys	9
Trp	7
Total	241

analytic ultracentrifugation equilibrium sedimentation, a technique that is firmly based in thermodynamics and does not therefore rely on calibration or on making assumptions concerning the shape of the protein. Using this approach, the apparent molecular mass of the PPL2 lectin in solutions with pH in the range 2.5–8.5 was 34 ± 3 kDa (Fig. 1B, insert). This figure, in conjunction with the MS analyses, showed that the protein behaved as a pH-independent monomeric protein.

Carbohydrate analysis performed by GLC (data not shown) failed to show the presence of any amino or neutral monosaccharide, strongly indicating that PPL2 was a nonglycosylated protein.

PPL2 displays chitinase activity

Edman degradation analysis of reduced and pyridylethylated protein yielded the first 42 amino acid residues of PPL2: GGIVVYWGQNGGEGTLTSTCESGL YQIVNIAFLSQFGGGRRP. A BLAST analysis (http://www.ncbi.nlm.nih.gov/blast/) revealed extensive (up to approximately 75%) similarity with a large number of plant chitinase sequences deposited in the publicly accessible protein databases, such as the basic chitinase III from *Nicotiana tabacum* (P29061), an acidic chitinase from *Glycine max* (BAA77677), chitinase b from *Phytolacca americana* (Q9S9F7), chitinase from *Psophocarpus tetragonolobus* (BAA08708), chitinases from *Vitis vinifera* (CAC14014), basic chitinase from *Vigna unguiculata* (Q43684), and chitinase B from leaves of pokeweed (Q9S9F7). All of these proteins are poly [1,4-(*N*-acetyl-β-D-glucosaminide)] glycanhydrolases of the glycosyl hydrolase family 18 (EC 3.2.1.14) [34] (http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00704), whose prototype is hevamine, isolated from the rubber tree [35,36].

The possible chitinase activity of PPL2 was investigated by quantitative GC determination of the amount of GlcNac released using chitin as substrate. PPL2 released 3 μg of GlcNac·h^{-1}·(mg protein)$^{-1}$. In comparison, commercial *Streptomyces griseus* chitinase exhibited an activity of 80 μg of GlcNac·h^{-1}·(mg protein)$^{-1}$, and the GlcNac-specific agglutinins from wheat germ (WGA) and *Urtica dioica* (UDA) did not show any chitinase activity. Peracetylated GlcNac (retention time 33.60 min) was observed in the reaction mixtures containing PPL2 or *Streptomyces griseus* chitinase but not in those reaction mixtures to which WGA or UDA were added. These results demonstrated that PPL2 was indeed an active chitinase able to hydrolyze the β(1–4) glycosidic bond linking the GlcNac units of chitin. In order to determine whether PPL2 presented chitinase activity only for the nonreducing end of chitin (exochitinase activity) or also had the ability to hydrolyze internal β(1–4) glycosidic linkages (an endochitinase activity), 40 μL of the reaction mixture used for the chitinase assay were analyzed by Dionex high-pH anion exchange chromatography using a CarboPac PA-100 column. The elution times of three major analytes present in the reaction mixture (3.93, 4.84 and 5.58 min) matched those of the standard carbohydrates GlcNac, $(GlcNac)_2$ and $(GlcNac)_3$ (3.86, 4.84 and 5.58 min, respectively). This result demonstrated an endochitinase activity for PPL2. The exact mechanism of glycoside hydrolysis (e.g. with retention or not of the β-anomeric configuration of the products) remains to be established, however.

The finding that PPL2 exhibited GlcNac-dependent hemagglutination and endochitinase activities was striking but not without precedent. The acidic chitinase BjCHI1 from *Brassica juncea* showed hemagglutination ability [37]. However, BjCHI1 is a unique chitinase with two chitin-binding domains, and both chitin-binding domains are essential for agglutination [38]. On the other hand, PPL2 is a single-domain protein. Hence, PPL2 may possess at least two carbohydrate-binding sites. One of them probably corresponds to the catalytic site, whereas the other one(s) remain to be characterized.

Plant chitinases constitute a class of pathogenesis-related proteins that play an important role in defense against pathogens through degradation of chitin present in the fungal cell wall and in insect cuticles [37,39]. The first characterization of a chitinase in the Mimosoideae subtribe, an antifungal chitinase from *Leucaena leucocephala* has been reported only recently [40]. This protein belongs to the class I chitinases of the glycosyl hydrolase family 19, and is, thus, structurally unrelated to PPL2.

It is noteworthy that the seeds of *Parkia platycephala* contain two different lectins: the mannose/glucose-specific PPL1 [19,21] and the GlcNac-binding lectin with chitinase activity, PPL2, described here. The fact that mannose is an abundant building block of surface-exposed glycoconjugates of viruses, bacteria and fungi supports the view that PPL, and other mannose-recognizing lectins, play a role in plant defense against pathogens [1]. Specifically, the planar array of carbohydrate-binding sites on the rim of the toroid-shaped structure of the *Parkia platycephala* lectin dimer [21] immediately suggested a mechanism to promote multivalent interactions leading to cross-linking of carbohydrate ligands as part of the host strategy against phytopredators and pathogens. The presence of two unrelated lectins in plant seeds has

been also reported in *Canavalia ensiformis* (Leguminosae): concanavalin A, a prototypic glucose/mannose-specific legume lectin built by the jellyroll fold [1,7], and concanavalin B, which, although it shares about 40% sequence identity with plant chitinases belonging to glycosyl hydrolase family 18, has not been shown to have any chitinase activity [41]. The lack of chitinase activity of concanavalin B can be explained by differences in the loops that form the substrate-binding cleft [42].

Sequencing of cDNA and genomic DNA for PPL2

Conserved amino acid sequences from glycosyl hydrolase family 18 were used to design two degenerate primers that allowed us to PCR-amplify a specific product of approximately 500 bp (pPPL2). Its sequence was then used to design a gene-specific forward primer (GSP-PPL2) to extend the sequence analysis of the PPL2 cDNA by 3′RACE. Using the GSP-PPL2 and Qo primers, the sequence was extended in the 3′ direction by PCR walking. From these sequences (pPPL2 and 3′RACE), two specific primers (PPL2f and PPL2r) were designed that amplified a fragment of 800 bp corresponding to the stretch between the conserved N-terminal sequence ^{6}YWGQNGG^{12} and the STOP codon (Fig. 2). Using primers designed from the cDNA sequence, the PPL2 gene was amplified from genomic DNA of *Parkia platycephala* seedlings. The size of the amplified genomic DNA was identical to that of the cDNA, indicating that the PPL2 gene was devoid of introns, as observed for other class III chitinase genes [43].

The complete amino acid sequence of PPL2 determined by the combination of N-terminal sequencing and cDNA cloning contains 271 amino acid residues, including the six conserved cysteine residues of class III chitinases, and the putative catalytic residues of class III plant chitinases, which in PPL2 correspond to amino acid positions 125 (Asp) and 127 (Glu). The calculated isotope-averaged molecular mass of the PPL2 sequence is 29 490.1 Da, which is about 86 ± 15 Da greater than the molecular mass determined by MALDI-TOF MS, suggesting that the native protein may lack the C-terminal valine residue.

Overall three-dimensional structure of PPL2

Figure 3 displays the structure of PPL2. The $2F_o - F_c$ density map contoured at 1σ showed that, with the exception of a small loop between the α_4 and β_5 regions corresponding to residues from Asn144 to Lys149, the majority of the protein residues were well

```
|---------------------------- N-terminal----------------------------------------
 G   G   I   V   V   Y   W   G   Q   N   G   G   E   G   T   L   T   S   T   C      20
                     TAT TGG GGC CAG AAT GGA GGA GAA GGC ACT TTG ACA TCA ACA TGT      60
                     |PPL2f ->

--------------------------------N-terminal-------------------------------------
 E   S   G   L   Y   Q   I   V   N   I   A   F   L   S   Q   F   G   G   G   R      40
GAA AGT GGC CTC TAT CAA ATT GTT AAC ATA GCT TTT CTA TCA CAA TTT GGG GGT GGT CGA    120
------->
 R   P   Q   I   N   L   A   G   H   C   D   P   A   N   N   G   C   R   T   V      60
CGA CCT CAA ATA AAC CTC GCT GGC CAC TGT GAT CCT GCT AAT AAT GGT TGC CGC ACG GTG    180

 S   D   G   I   R   A   C   Q   R   R   G   I   K   V   M   L   S   I   G   G      80
AGC GAT GGC ATC AGA GCC TGC CAA AGA AGA GGC ATC AAG GTC ATG CTC TCC ATC GGA GGC    240

 G   A   G   S   Y   S   L   S   S   V   Q   D   A   R   S   V   A   D   Y   I     100
GGC GCC GGA AGC TAC TCG TTG TCG TCG GTC CAG GAT GCC AGA AGC GTA GCA GAT TAC ATA    300

 W   N   N   F   L   G   G   R   S   S   S   R   P   L   G   D   A   V   L   D     120
TGG AAC AAC TTT CTA GGT GGA CGA TCA TCC TCA AGA CCT CTA GGC GAC GCC GTT TTG GAT    360

 G   V   D   F   D   I   E   H   G   G   A   Y   Y   D   A   L   A   R   R   L     140
GGA GTA GAT TTT GAC ATT GAA CAT GGT GGT GCA TAC TAT GAT GCT CTT GCA AGA AGA CTA    420

 S   E   H   N   R   G   G   K   K   V   F   L   S   A   A   P   Q   C   P   F     160
TCA GAG CAT AAC CGA GGA GGT AAA AAA GTG TTC TTA TCT GCT GCA CCA CAA TGT CCT TTT    480
                                                |-GSP-PPL2->

 P   D   Q   S   L   N   K   A   L   S   T   G   L   F   D   Y   V   W   V   Q     180
CCA GAT CAG TCT CTC AAT AAG GCA CTT TCC ACT GGG CTC TTT GAT TAT GTG TGG GTT CAA    540

 F   Y   N   N   P   Q   C   E   F   N   S   G   N   P   S   N   F   R   N   S     200
TTC TAT AAC AAC CCT CAG TGT GAG TTT AAT TCT GGA AAC CCT AGC AAC TTT CGG AAC TCA    600

 W   N   K   W   T   S   S   F   N   A   K   F   Y   V   G   L   P   A   S   P     220
TGG AAC AAG TGG ACT TCA AGT TTC AAT GCC AAG TTT TAT GTT GGG CTT CCT GCC TCA CCT    660

 E   A   A   G   S   G   Y   V   P   P   Q   Q   L   I   N   Q   V   L   P   F     240
GAA GCA GCT GGG AGT GGG TAT GTG CCA CCA CAA CAA CTT ATA AAT CAA GTG TTG CCT TTT    720

 V   K   R   S   P   K   Y   G   G   V   M   L   W   D   R   F   N   D   L   K     260
GTT AAA AGA TCT CCC AAG TAT GGA GGA GTC ATG CTT TGG GAT AGG TTC AAT GAT CTG AAA    780

 T   K   Y   S   S   K   I   K   P   S   V   *                                     271
ACT AAG TAC AGC TCC AAA ATT AAG CCC AGT GTT TGA                                   816
                                <- PPL2r -|
```

Fig. 2. cDNA and amino acid sequence of PPL2. The nucleotide and the amino acid sequences are numbered on the right side. The underlined nucleotide sequences correspond to primers used to clone and sequence the full-length PPL2. The underlined amino acid sequences 6–12 and 178–185 represent the conserved polypeptide stretches from which degenerate primers were initially designed. The N-terminal amino acid sequence determined by Edman degradation is labeled. The six conserved cysteines of class III chitinases are shadowed, and the conserved residues of the active site of family 18 of glycosyl hydrolases are boxed.

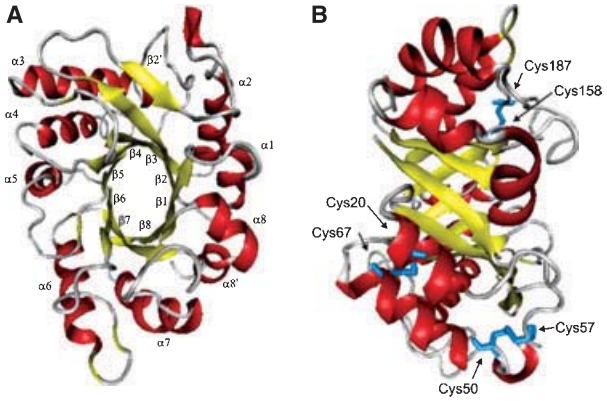


Fig. 3. Crystal structure of PPL2. (A) and (B) show two views of the $(\alpha\beta)_8$ barrel fold of PPL2. The α-helices (red) and β-strands (yellow) are labeled from 1 to 8. Disulfide bonds are depicted in blue. In (B), the active site cleft loops are located at the right face of the model.

fitted. The PPL2 model has good overall stereochemistry (Table 2), with no amino acid residues in the disallowed region of the Ramachandran plot. The PPL2 structure consists of a compact $(\beta/\alpha)_8$ barrel with dimensions of approximately $50 \times 40 \times 25$ Å, including three disulfide bonds (Cys20–Cys67, Cys50–Cys57 and Cys158–Cys187) and five *cis* peptide bonds. Two of the *cis* peptide bonds of PPL2 (Gly147–Lys148 and Lys148–Lys149) are located in a region of poor density, whereas the remaining three (Ala31–Phe32, Phe160–Pro161 and Trp253–Asp254) are well defined at the electron density. With the exception of four sulfate ions (Fig. 4), which presumably remained bound to PPL2 throughout its purification protocol, as the protein was precipitated by ammonium sulfate to separate it from pigments, no metal ions or ligands were detected. Sulfate ions were assigned according to Copley and Barton [44].

Structural comparison and analysis of conserved motifs

The overall structural features of the PPL2 model are conserved in other GH18 plant chitinases, i.e. hevamine (*Hevea brasiliensis*) (PDB code 2HVM), the

Table 2. Statistics of data collection, refinement and quality of the structure.

	Overall resolution dataset	Highest resolution dataset
Data collection		
Total number of observations	95 262	12 669
Total number of unique observations	25 805	3521
R_{merge}	0.040	0.228
Highest resolution limit (Å)	1.73	1.73
Lowest resolution limit (Å)	32.31	1.83
Completeness (%)	95.5	90.4
Multiplicity	3.7	3.6
I/σ (I)	13.1	2.4
Wavelength (Å)	1.431	
Space group	$P2_12_12_1$	
Cell parameters (Å)	$a = 55.19$, $b = 59.95$, $c = 76.70$	
Refinement		
Resolution range (Å)	1.73–32.31	
R_{factor} (%)	16.88	
R_{free} (%)	19.87	
Number of nonhydrogen atoms in protein structure	2086	
Number of sulfate ions	4	
Number of water molecules	249	
Root mean square deviations from ideal values		
Bond lengths (Å)	0.012	
Bond angles (degrees)	1.48	
Temperature factors		
Average *B*-value for whole protein chain (Å^2)	13.26	
Average *B*-value for sulfate ions (Å^2)	41.97	
Average *B*-values for water molecules (Å^2)	24.29	
Ramachandran plot		
Residues in most favored regions	195 (87.8%)	
Residues in additional allowed regions	26 (11.7%)	
Residues in generously allowed regions	1 (0.5%)	

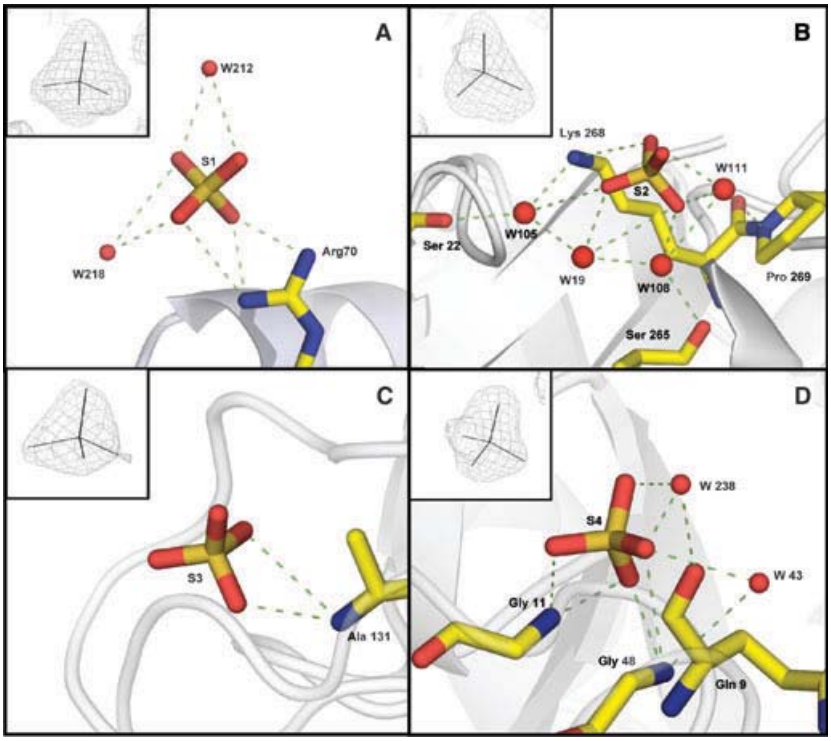


Fig. 4. Sulfate ions bound to crystallized PPL2. (A)–(D) display details of the binding of sulfate ions (S) 1–4 within the crystal structure of PPL2. In each panel, the electron density assigned to the sulfate ions is displayed in an insert. W, water molecule.

xylanase inhibitor XIP-I from *Triticum aestivum* (1TE1), and ConB (*Canavalia ensiformis*) (1CNV), with which PPL2 shares 68%, 40% and 40% sequence similarity, respectively (Fig. 4A). The three-dimensional structure of PPL2 can be superimposed onto those of hevamine, XIP-I and ConB, with root mean square deviation (r.m.s.d.) for all Cα atoms of 0.90 Å, 1.01 Å and 1.14 Å, respectively. In particular, the two consensus motifs described for the glycosyl hydrolase family 18, e.g. the presence of the absolutely conserved strands β_3 and β_4 (Fig. 4A, boxed), and the hydrogen bond network between residues Asp120 and Gly121 and Val74 (Fig. 4A,B) [33], are also conserved in PPL2. On the other hand, the largest structural divergence is associated with the active site cleft loops, which comprise the residues linking neighbor β-strands in the $(\alpha\beta)_8$ barrel. Thus, whereas with the exception of the $\beta_6\alpha_6$ loop, all the active site cleft loops of PPL2 are highly conserved in hevamine, and only few structural differences are evident when comparing the $\beta_2\alpha_2$ and $\beta_7\alpha_7$ loops from PPL2 and ConB, the active site cleft loops from XIP-I significantly depart from those of PPL2.

The PPL2 chitin-binding site

X-ray studies have suggested that enzymes of the GH18 family showing chitinase activity have conserved Asp125, Glu127 and Tyr183 amino acids (hevamine numbering) in their active sites. Their significance for catalysis is not well understood, although it has been suggested that Glu127 may act as a proton donor to the cleavable glycosidic bond, and Asp125 and Tyr183 would contribute to the stabilization of the oxazolinium intermediate [45]. In PPL2, these residues correspond to Asp125, Glu127 and Tyr182 (Figs 2 and 4A). Asp125 and Glu127 are located in the $\beta_4\alpha_4$ loop, and Tyr182 at loop $\beta_6\alpha_6$. The highly conserved, functionally relevant, structural features that are common to PPL2 and hevamine suggest that these two chitinases may share essentially the same catalytic mechanism. In addition, our data showing that PPL2 strongly binds GlcNac would support a hypothetical mechanism by which the lectin hydrolyzes a chitin polymer by cycles of anchoring, cleavage and being released from a GlcNac-binding site, and anchoring to another GlcNac-binding site. Clearly, detailed molecular and structural studies are required to investigate this.

Experimental procedures

Isolation of PPL2

Mature seeds from *Parkia platycephala* were collected in the state of Ceará (north-eastern Brazil) and ground in a coffee mill. The flour was defatted with *n*-hexane, air-dried at room

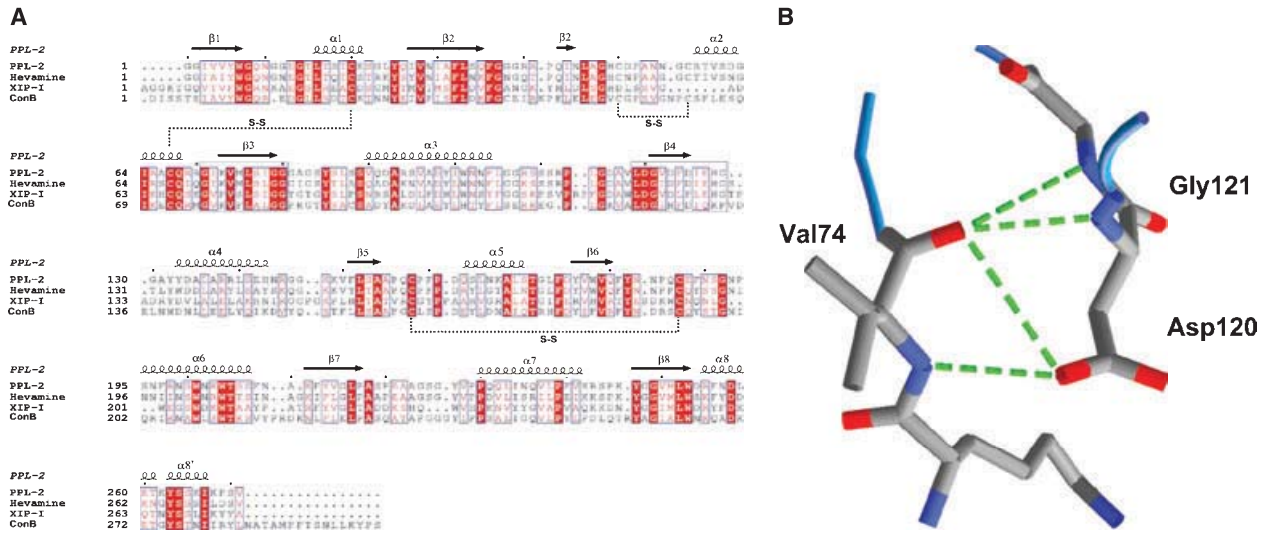


Fig. 5. Structural features of PPL2 and the GH18 family. (A) Multiple sequence alignment of PPL2, hevamine, XIP-I and ConB. Absolutely conserved residues in the four proteins are shown in white over a red background. Conservative substitutions or residues conserved in at least two proteins are depicted in pale red and boxed. Cysteine residues engaged in the formation of disulfide bonds (S–S) are conected by discontinuous lines. The secondary structure elements of PPL2 are shown on top of the sequence alignment: arrows represent β-strands and springs denote α-helices. (B) Detail of the network of hydrogen bonds between PPL2 residues Asp120, Gly121 and Val74, which represent a conserved structural motif of the GH18 family.

temperature and kept dry for further use. Soluble proteins were extracted overnight at room temperature by continuous stirring with 1 : 15 (w/v) 500 mM HCl solution, containing 150 mM NaCl. Insoluble material was separated by centrifugation (Ultracentrifuga Beckman modelo XL-1, Palo Alto, CA) at 10 000 ***g*** for 20 min at 5 °C. The supernatant was adjusted to pH 7.0 and left for 12 h at 4 °C. Precipitated pigments were removed by centrifugation (Ultracentrifuga Beckman modelo XL-1), and the supernatant was subjected to precipitation with 60% saturated ammonium sulfate. After centrifugation (Ultracentrifuga Beckman modelo XL-1), the pellet was resuspended in a small volume of 50 mM Tris, pH 7.0, containing 100 mM NaCl, dialyzed against this buffer, and subjected to affinity chromatography on a Red-Sepharose CL-4B column (26 × 1.5 cm) (Sigma-Aldrich, São Paulo, Brazil) equilibrated with the same buffer as described previously for GlcNAc-specific enzymes [46]. Unbound material was eluted by washing the column with equilibration buffer, and the retained fraction was desorbed with 3 M NaCl in buffer, dialyzed against equilibrium buffer, and assayed for hemagglutinating activity following a standard procedure with trypsin-treated rabbit red blood cells [47]. To this end, a two-fold dilution was prepared for each sugar (1 M starting concentration) solution in 0.15 M NaCl containing 5 mM $CaCl_2$ and 5 mM $MnCl_2$. Each dilution had a final volume of 0.2 mL. The purified lectin was diluted in 0.15 M NaCl to achieve 4 units of hemagglutinating activity per mL. The lowest concentration of inhibitor exhibiting agglutinating activity was termed the minimum inhibitory concentration. Aliquots of 0.2 mL of the 4 unit solution of the lectin were used for hemagglutination inhibition assay. Monosaccharides (mannose, glucose, galactose, *N*-acetylglucosamine, *N*-acetylgalactosamine, fucose) and glycoproteins (bovine thyroglobulin, ovine submaxillary mucin, bovine fetuin, and asialofetuin) were tested for hemagglutination inhibitory activity.

Purification of PPL2

The protein fraction retained in the Red-Sepharose CL-4B column was further fractionated by RP-HPLC and by chitin affinity chromatography. For RP-HPLC, 3 mg of total proteins was dissolved in 250 μL of 0.1% trifluoroacetic acid (solution A) and centrifuged (Ultracentrifuga Beckman modelo XL-1) at 4500 ***g*** for 2 min. The supernatant was applied on a μBondapack C18 analytic column (3.9 × 300 mm) (Waters, Milford, MA, USA) equilibrated in solution A, and the column was developed using the following chromatographic conditions: 100% buffer A for 5 min, followed by gradients of 0–30% of solution B (66.6% acetonitrile in A) for 5 min, 30–40% B for 30 min, 40–70% B for 5 min, 70–80% B for 10 min, 80–100% B for 5 min, and 100% B for 10 min. The elution was monitored at 280 nm. Fractions were collected manually, lyophilized and stored at −70 °C until used. For affinity chromatography, the protein fraction retained in the Red-Sepharose column was applied overnight to a chitin column (2 × 5 cm) (Sigma-Aldrich) equilibrated in 50 mM Tris/HCl, 150 mM NaCl, pH 7.2. Unbound material was eluted by washing the column with equilibration buffer,

and the retained fraction was desorbed with 50 mM Tris/HCl, 3 M NaCl, pH 7.2.

Molecular mass determinations

Tricine-PAGE in a discontinuous gel and buffer system [48] was used to estimate the apparent molecular mass of the proteins. Samples were denatured for 10 min in sample buffer containing 2.5% (w/v) SDS before electrophoresis. After the run, the gels were stained with Coomassie Brilliant Blue G (0.2%) in methanol/acetic acid/water (4 : 1 : 6, v/v) and destained in the same solution. Protein molecular weight markers (GE Healthcare Biosciences AB, Uppsala, Sweden) were included in each run.

The molecular masses of the native, reduced and carbamidomethylated proteins were determined by MALDI-TOF MS using an Applied Biosystems (Foster City, CA, USA) Voyager PRO-STR instrument operating at an accelerating voltage of 25 kV in the linear mode and using 3,5-dimethoxy-4-hydroxycinnamic acid (10 mg·mL^{-1} in 50% acetonitrile) as the matrix.

The apparent molecular mass of the *Parkia platycephala* lectin 2 in solutions of different pH was determined by size-exclusion chromatography and by analytic ultracentrifugation equilibrium sedimentation using a Beckman XL-A centrifuge with UV absorption scanner optics. For size-exclusion chromatography, PPL2 (2 mg·mL^{-1}) was applied to a Superose-12 HR10/30 column connected to an ÄKTA HPLC system (GE-Healthcare Bioscience). The column was equilibrated and eluted with 20 mM sodium phosphate buffer, pH 7.2, containing 150 mM NaCl at a flow rate of 0.5 mL·min^{-1}. Elution was monitored at 280 nm. Equilibrium sedimentation experiments were carried out at 20 °C and 13 000 r.p.m. using an AN-50 Ti rotor. The protein was dissolved at about 0.1 mg·mL^{-1} in the following buffers, each containing 100 mM NaCl, 1 mM Cl_2Mn, and 1 mM Cl_2Ca: 20 mM sodium citrate pH 2.5; 20 mM sodium citrate, pH 3.5; 20 mM sodium citrate, pH 4.5; 20 mM Mes, pH 5.5; 20 mM Mes, pH 6.5; 20 mM Tris/HCl, pH 7.5; and 20 mM Tris/HCl, pH 8.5.

Quantitation of free cysteine residues and disulfide bonds

For quantitation of free cysteine residues and disulfide bonds, the purified proteins dissolved in 10 μL of 50 mM Hepes, pH 9.0, 5 M guanidine hydrochloride containing 1 mM EDTA were heat-denatured at 85 °C for 15 min, allowed to cool at room temperature, and incubated with either 10 mM 4-vinylpyridine for 15 min at room temperature, or with 10 mM 1,4-dithioerythritol (Sigma-Aldrich) for 15 min at 80 °C; this was followed by addition of 4-vinylpyridine at 25 mM final concentration and incubation for 1 h at room temperature. The pyridylethylated (PE) protein was freed from reagents using a C18 Zip-Tip pipette tip (Millipore Ibérica S.A., Madrid, Spain) after activation with 70% acetonitrile (ACN) and equilibration in 0.1% trifluoroacetic acid. Following protein adsorption and washing with 0.1% trifluoroacetic acid, the PE-protein was eluted onto the MALDI-TOF plate with 1 μL of 70% ACN and 0.1% trifluoroacetic acid and subjected to MS analysis as above.

The number of free cysteine residues (N_{SH}) was determined from Eqn (1):

$$N_{SH} = (M_{PE} - M_{NAT})/105.1 \quad (1)$$

where M_{PE} is the mass of the denatured but nonreduced protein incubated in the presence of 4-vinylpyridine, M_{NAT} is the mass of the native, HPLC-isolated protein, and 105.1 is the mass increment (in Da) due to the pyridylethylation of one thiol group.

The number of total cysteine residues (N_{Cys}) can be calculated from Eqn (2):

$$N_{Cys} = (M_{Alk} - M_{NAT})/105.1 \quad (2)$$

where M_{Alk} is the mass (in Da) of the fully reduced and alkylated protein.

Finally, the number of disulfide bonds N_{S-S} can be calculated from Eqn (3):

$$N_{S-S} = (N_{cys} - N_{SH})/2 \quad (3)$$

Amino acid analysis and N-terminal amino acid sequence determination

Amino acid analysis was performed on a Pico-Tag amino acid analyzer (Waters) as described [49]. One nanomole of purified protein was hydrolyzed in 6 M HCl/1% phenol at 106 °C for 24 h. The hydrolyzate was reacted with 20 μL of fresh derivatization solution (methanol/triethylamine/water/phenylisothiocyanate, 7 : 1 : 1 : 1, v/v) for 1 h at room temperature, and the phenylisothiocyanate (PTC)-amino acids were identified and quantitated on an RP-HPLC column calibrated with a mixture of standard PTC-amino acids (Pierce, Rockford, IL, USA). Cysteine residues were determined as cysteic acid.

N-terminal sequencing of reduced and carboxymethylated proteins was performed in an Applied Biosystems model Procise 491 gas–liquid protein sequencer. The phenylthiohydantoin (PTH) derivatives of the amino acids were identified with an Applied Biosystems model 450 microgradient PTH analyzer.

Genomic DNA and RNA isolation, and cDNA cloning

Genomic DNA from fresh leaves of 2-week-old seedlings of *Parkia platycephala* grown from mature seeds was extracted using the cetyl triethylammonium bromide (CTAB) procedure [50].

For RNA isolation, young *Parkia platycephala* buds were immediately ground to a powder with a pestle in liquid nitrogen. Total cellular RNA was isolated with Concert Plant RNA reagent (Invitrogen S.A., Barcelona, Spain). Single-stranded cDNAs were synthesized by reverse transcription using oligo-dT_{17} and MMLV reverse transcriptase (Promega Biotech Ibérica, Madrid, Spain). Degenerated primers were designed from conserved amino acid sequences of plant chitinases YWGQNGG and WVQFY NNP (sense primer 5′-TAY TGG GAR AAY GGN GG-3′, and antisense primer 5′-GG RTT RTT RTA RAA YTG NAC CCA-3′; the nomenclature follows the IUPAC code for degeneracies). PCR amplification was performed with 1 U (International unit) of Taq DNA polymerase (HF, Roche Diagnostics S.L., Barcelona, Spain) using the following conditions: DNA was denatured at 94 °C for 4 min, and this was followed by 30 cycles of denaturation (30 s at 94 °C), annealing (30 s at 50 °C) and extension (30 s at 72 °C), followed by a final extension for 10 min at 72 °C. The amplified DNA fragment was cloned into the pGEM-T vector (Invitrogen). The inserted DNA fragments were subjected to sequencing on an Applied Biosystems model 377 DNA sequencing system using T7 and SP6 primers, and this sequence was used for designing specific oligonucleotides for completing the sequence by 3′RACE. 3′RACE was done as described [51] using the Qt primer (5′-CCA GTG AGC AGA GTG ACG AGG ACT CGA GCT CAA GCT_{16}-3′) for reverse transcription, and the sense primer GSP-PPL2 (5′-CTG CTG CAC CAC AAT GTC CTT TTC-3′) and the antisense primer Qo (5′-CCA GTG AGC AGA GTG ACG-3′) for PCR amplification. The 3′RACE reaction conditions were as those for cDNA amplification, except that annealing was done at 60 °C. Using this information, two specific primers were designed, PPL2-forward (5′-TAT TGG GGC CAG AAT GGA G-3′) and PPL2-reverse (5′-TCAA ACA CTG GGC TTA ATT TTG G-3′) for amplifying and sequencing the full-length ORF of PPL2.

Assay for chitinase activity

Chitinase enzymatic assays were performed in Pyrex tubes (7 mL) with Teflon-lined screw caps. The reaction mixtures (total 1250 μL) contained 0.05 M sodium acetate buffer (pH 5.5), 5 mg of washed chitin powder (blank), and either 25 μL of a PPL2 solution (1 $mg{\cdot}mL^{-1}$) or 10 μL (0.5 μU) of *Streptomyces griseus* family 19 chitinase (Sigma) (one unit will liberate 1.0 mg of GlcNac from chitin per hour at pH 6.0 at 25 °C in a 2 h assay) as positive control, both in sodium acetate buffer. The negative control consisted of the same reaction mixture, except that sodium acetate buffer replaced the protein sample. Twenty-five microliters of 1 $mg{\cdot}mL^{-1}$ solutions of two GlcNac-specific lectins, the agglutinins from wheat germ (WGA) and *Urtica dioica* (UDA), which are devoid of chitinase activity, were also included in the assays as specificity controls. For calibration and quantitation, a mixture of 1 μg of each, mannose and GlcNac in sodium acetate buffer was used. The reaction mixtures were incubated at 37 °C for 3 h and lyophilized. GlcNac production was monitored and quantitated as peracetylated GlcNac by GC [Varian 3400 gas chromatograph equipped with a flame ionization detector, a Ross injector and a 30 m × 0.25 mm capillary column $EC.Tm^{-1}$ (100% methylsilicone apolar phase of column, $EC.Tm^{-1}$, 0.25 μm film phase, Altech), 0.25 μm film phase (Altech, Flemington, NJ, USA)]. The injector and detector temperature was 250 °C, and the oven temperature program was 3 $C°{\cdot}min^{-1}$ from 120 to 250 °C. The carrier gas helium pressure was 1 bar. Briefly, released GlcNac was peracetylated by addition of 0.5 mL of acetic anhydride to the lyophilized samples, followed by incubation for 4 h at 100 °C. Samples were then evaporated to dryness under a stream of nitrogen and mild heating with a hair dryer. To eliminate salts and proteins from the reaction mixture, 1.5 mL of chloroform and 1 mL of distilled water were added to each tube. After thorough vortexing, the aqueous upper phase was discarded and the lower chloroform phase was extracted four times with 1 mL of distilled water. The chloroform phases were freed of water by filtration through small columns made of a Pasteur pipette filled with anhydrous sodium sulfate. The filtrates were collected in Pyrex tubes (7 mL) and evaporated to dryness under a stream of nitrogen. Chloroform (40 μL) was added to each tube, and 4 μL was injected in the gas chromatograph for analysis.

GlcNac production (retention time 33.60 min) was also monitored by GC/MS analysis performed on a Carlo Erba GC 8000 gas chromatograph equipped with a 25 m × 0.32 mm CP-Sil 5CB low-bleed MS capillary column, 0.25 μm film phase (Chrompack France, Les Ullis, France). The temperature of the Ross injector was 250 °C and the samples were analyzed using the following temperature program: 120 °C for 3 min, then 3 $C°{\cdot}min^{-1}$ until 250 °C. The column was coupled to a Finnigan Automass II mass spectrometer. The analyses were performed either in the electron impact mode (ionization energy 70 eV, source temperature 150 °C) or in the chemical ionization mode in the presence of ammonia (ionization energy 150 eV, source temperature 100 °C). Detection was performed for positive ions.

High-pH anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD)

HPAEC-PAD was performed with a Dionex Series DX30 HPLC system (Dionex Corporation, Voisins Le Bretonneux, France) equipped with a pulsed electrochemical detector, operating in the pulsed amperometric detection mode with a gold working electrode and an Ag/AgCl reference electrode. Electrode potential settings were E1 + 0.05 V, E2 + 0.6 V and E3 − 0.6 V, with 500, 3 and 7 ms applied durations, respectively, and an integrated time period of

0.10–0.48 s. Detection was set with a range of detection of 300 nC. Detector response was analyzed with a C-R8A chromatopac integrator (Shimadzu, Kyoto, Japan). A standard sample consisted of 0.1 $\mu g \cdot \mu L^{-1}$ of each GlcNac, *N*,*N*′-diacetylchitobiose or *N*,*N*′,*N*″-triacetylchitotriose dissolved in water. Samples (12.5 μL) were injected in a Dionex CarboPac PA-100 pellicular anion exchange column running at a flow rate of 0.8 $mL \cdot min^{-1}$. Elution was performed with buffer A (100 mM NaOH) for 1 min followed by a linear gradient of 0–40% buffer B (100 mM NaOH and 1 M sodium acetate) over 40 min.

Crystallization and structure determination

PPL2 was crystallized by the hanging drop vapor diffusion method at 20 °C as described [28]. The crystals belong to the $P2_12_12_1$ space group with one monomer in the asymmetric unit. Crystals soaked in a cryoprotectant solution containing 75% of mother liquor [0.2 M ammonium acetate, 0.1 M trisodium citrate dehydrate, pH 5.6, and 30% (w/v) PEG 4000] and 25% of glycerol were flash-frozen at 100 K in a liquid nitrogen stream. X-ray diffraction data were collected at 1.73 Å at the synchrotron radiation source of Cpr station Laboratório Nacional de Luz Síncrotron (Campinas, Brazil). The data were processed and scaled using MOSFLM and SCALA [52], respectively. Crystallographic data are summarized in Table 2.

The PPL2 crystal structure was determined by molecular replacement using the AMORE software [52], using data in the resolution range 15–3.0 Å, and the hevamine coordinates (PDB accession code 2HVM) as the search model. Rotation and translation functions revealed one molecule in the asymmetric unit. The position and orientation of the molecule, as a single rigid body entity, were refined for 20 cycles with REFMAC [52], using reflections in the resolution range 32–1.73 Å. Appropriate amino acid changes were carried out to convert the molecular model of hevamine into PPL2. Several steps of rebuilding, interspersed with restrained refinement, using REFMAC, yielded the current model at 1.73 Å resolution. Sulfate ion molecules were placed by inspection of the $F_o - F_c$ map. For each cycle of refinement, the stereochemistry of the model was monitored with the PROCHECK incorporated into the CCP4 package [52]. Finally, water molecules were placed in the model over several steps of refinement with ARP/WARP and inspected manually. The atomic coordinates, fitted with XTALVIEW [52], are accesible from the Protein DataBank (http://www.rcsb.org/pdb/) under code 2GSJ.

Acknowledgements

This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), CAPES, FUNCAP, PADCT and Program CAPES/COFECUB no. 336/01, and grant BFU2004-01432/BMC from the Ministerio de Educación y Ciencia, Madrid, Spain. B. S. Cavada, W. F. De Azevedo Jr and A. H. Sampaio are senior investigators of CNPq.

References

1 Van Damme EJM, Peumans WJ, Barre A & 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.
2 Gabius H-J & Gabius S (1997) *Glycoscience. Status and Perspectives*. Chapman & Hall, Weinheim.
3 Dodd RB & Drickamer K (2001) Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity. *Glycobiology* **11**, 71–79.
4 Rini JM (1995) Lectin structure. *Annu Rev Biomol Struct* **24**, 551–577.
5 Weis WI & Drickamer K (1996) Structural basis of lectin-carbohydrate recognition. *Annu Rev Biochem* **65**, 441–473.
6 Elgavish S & Shaanan B (1997) Lectin–carbohydrate interactions: different folds, common recognition principles. *Trends Biochem Sci* **22**, 462–467.
7 Loris R, Hamelryck T, Bouckaert J & Wyns L (1998) Legume lectin structure. *Biochim Biophys Acta* **1383**, 9–36.
8 Bouckaert J, Hamelryck T, Wyns L & Loris R (1999) Novel structures of plant lectins and their complexes with carbohydrates. *Curr Opin Struct Biol* **9**, 572–577.
9 Vijayan M & Chandra N (1999) Lectins. *Curr Opin Struct Biol* **9**, 707–714.
10 Chervenak MC & Toone EJ (1995) Calorimetric analysis of the binding of lectins with overlapping carbohydrate binding. *Biochemistry* **34**, 5685–5695.
11 Dam TK, Cavada BS, Grangeiro TB, Santos CF, de Sousa FAM, Oscarson S & Brewer CF (1998) Diocleinae lectins are a group of proteins with conserved binding sites for the core trimannoside of asparagine-linked oligosaccharides and differential specificities for complex carbohydrates. *J Biol Chem* **273**, 12082–12088.
12 Dam TK, Cavada BS, Grangeiro TB, Santos CF, Ceccatto VM, de Sousa FAM, Oscarson S & Brewer CF (2000) Thermodynamic binding studies of lectins from the diocleinae subtribe to deoxy analogs of the core trimannoside of asparagine-linked oligosaccharides. *J Biol Chem* **275**, 16119–16126.
13 Dam TK, Roy R, Das SK, Oscarson S & Brewer CF (2000) Binding of multivalent carbohydrates to concanavalin A and Dioclea grandiflora lectin. Thermodynamic analysis of the 'multivalency effect'. *J Biol Chem* **275**, 14223–14230.

14 Suvachittanont W & Peutpaiboon A (1992) Lectin from *Parkia speciosa* seeds. *Phytochemistry* **31**, 4065–4070.
15 Utarabhand P & Akkayanont P (1995) Purification of a lectin from *Parkia javanica* beans. *Phytochemistry* **38**, 281–285.
16 Cavada BS, Madeira SVF, Calvete JJ, Sousa LAG, Bomfim LR, Dantas AR, Lopes MC, Grangeiro TB, Freitas BT, Pinto VPT *et al.* (2000) Purification, chemical, and immunochemical properties of a new lectin from Mimosoideae (*Parkia discolor*). *Prep Biochem Biotech* **30**, 271–280.
17 Cavada BS, Santos CF, Grangeiro TB, Moreira da Silva LIM, Campos MJO, de Sousa FAM & Calvete JJ (1997) Isolation and partial characterization of a lectin from *Parkia platycephala* Benth seeds. *Physiol Mol Biol Plant* **3**, 109–115.
18 Ramos MV, Cavada BS, Bomfim LR, Debray H, Mazard A-M, Calvete JJ, Grangeiro TB & Rougé P (1999) Interaction of the seed lectin from *Parkia platycephala* (*Mimosoideae*) with carbohydrates and complex glycans. *Prot Pept Lett* **6**, 215–222.
19 Mann K, Farias CM, Gallego del Sol FG, Santos CF, Grangeiro TB, Nagano CS, Cavada BS & Calvete JJ (2001) The amino-acid sequence of the glucose/mannose-specific lectin isolated from *Parkia platycephala* seeds reveals three tandemly arranged jacalin-related domains. *Eur J Biochem* **268**, 4414–4422.
20 Gallego del Sol F, Gómez J, Hoos C, Nagano CS, Cavada BS, England P & Calvete JJ (2005) Energetics of 5-bromo-4-chloro-3-indolyl-α-D-mannose binding to the *Parkia platycephala* seed lectin and its use for MAD phasing. *Acta Cryst F* **61**, 326–331.
21 Gallego del Sol F, Nagano CS, Cavada BS & Calvete JJ (2005) The first crystal structure of a *Mimosoideae* lectin reveals a novel quaternary arrangement of a widespread domain. *J Mol Biol* **353**, 574–583.
22 Heywood VH (1971) Chemotaxonomy of the *Leguminosae* (Harborne JB & Boulter D, eds), pp. 1–29. Academic Press, London.
23 Chrispeels MJ & Raikhel NV (1991) Lectins, lectin genes, and their role in plant defense. *Plant Cell* **3**, 1–9.
24 Wang X & Ma Q (2005) Characterization of a jasmonate-regulated wheat protein related to a β-glucosidase-aggregating factor. *Plant Physiol Biochem* **43**, 185–192.
25 Collinge DB, Kragh KM, Mikkelsen JD, Nielsen KK, Rasmussen U & Vad K (1993) Plant chitinases. *Plant J* **3**, 31–40.
26 Hamel F, Boivin R, Tremblay C & Bellemare G (1997) Structural and evolutionary relationships among chitinases of flowering plants. *J Mol Evol* **44**, 614–624.
27 Kasprzewska A (2003) Plant chitinases – regulation and function. *Cell Mol Biol Lett* **8**, 809–824.
28 Cavada BS, Castellón RER, Vasconcelos GG, Rocha BAM, Bezerra GA, Debray H, Delatorre P, Nagano CS, Toyama M, Pinto VPT *et al.* (2005) Crystallization and preliminary X-ray diffraction analysis of a new chitin-binding protein from *Parkia platycephala* seeds. *Acta Crystallogr F* **61**, 841–843.
29 Rawitch AB, Pollock HG & Yang S-X (1993) Thyroglobulin glycosylation: location and nature of the N-linked oligosaccharide units in bovine thyroglobulin. *Arch Biochem Biophys* **300**, 271–279.
30 Hill HD Jr, Reynolds JA & Hill RL (1977) Purification, composition, molecular weight, and subunit structure of ovine submaxillary mucin. *J Biol Chem* **252**, 3791–3798.
31 Spiro RG & Bhoyroo D (1974) Structure of the O-glycosidically linked carbohydrate units of fetuin. *J Biol Chem* **249**, 5704–5717.
32 Green ED, Adelt G, Baenziger JU, Wilson S & Van Halbeek H (1988) The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharide by 500-megahertz ^{1}H NMR spectroscopy. *J Biol Chem* **263**, 18253–18268.
33 Rohrer JS, Cooper GA & Townsend RR (1993) Identification, quantitation, and characterization of glycopeptides in reversed-phase HPLC separations of glycoprotein proteolytic digests. *Anal Biochem* **212**, 7–16.
34 Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* **280**, 309–316.
35 Jekel PA, Hartmann BH & Beintema JJ (1991) The primary structure of hevamine, an enzyme with lysozyme/chitinase activity from *Hevea brasiliensis* latex. *Eur J Biochem* **200**, 123–130.
36 Van Scheltinga ACT, Kalk KH, Beintema JJ & Dijkstra BW (1994) Crystal structures of hevamine, a plant defense protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure* **2**, 1181–1189.
37 Chye ML, Zhao KJ, He ZM, Ramalingam S & Fung KL (2005) An agglutinating chitinase with two chitin-binding domains confers fungal protection in transgenic potato. *Planta* **220**, 717–730.
38 Tang CM, Chye ML, Ramalingam S, Ouyang SW, Zhao KJ, Ubhayasekera W & Mowbray SL (2004) Functional analyses of the chitin-binding domains and the catalytic domain of *Brassica juncea* chitinase BjCHI1. *Plant Mol Biol* **56**, 285–298.
39 Robertus JD & Monzingo AF (1999) The structure and action of chitinases. *EXS* **87**, 125–135.
40 Kaomek M, Mizuno K, Fujimura T, Sriyotha P & Cairns JR (2003) Cloning, expression, and characterization of an antifungal chitinase from *Leucaena leucocephala* de Wit. *Biosci Biotechnol Biochem* **67**, 667–676.
41 Hennig M, Jansonius JN, Van Scheltinga ACT, Dijkstra BW & Schlesier BJ (1995) Crystal structure of concanavalin B at 1.65 A resolution. An 'inactivated' chitinase

from seeds of *Canavalia ensiformis*. *J Mol Biol* **254**, 237–246.

42 Van Scheltinga ACT, Hennig M & Dijkstra BW (1996) The 1.8 Å resolution structure of hevamine, a plant chitinase/lysozyme, and analysis of the conserved sequence and structure motifs of glycosyl hydrolase family 18. *J Mol Biol* **262**, 243–257.

43 Lawton KA, Beck J, Potter S, Ward E & Ryals J (1994) Regulation of cucumber class III chitinase gene expression. *Mol Plant-Microbe Interact* **7**, 48–57.

44 Copley RR & Barton GJ (1994) A structural analysis of phosphate and sulphate binding sites in proteins. Estimation of propensities for binding and conservation of phosphate binding sites. *J Mol Biol* **242**, 321–329.

45 Bokma E, Rozeboom HJ, Sibbald M, Dijkstra BW & Beintema JJ (2002) Expression and characterization of active site mutants of hevamine, a chitinase from the rubber tree *Hevea brasiliensis*. *Eur J Biochem* **269**, 893–901.

46 Pastuszak I, Drake R & Elbein AD (1996) Kidney *N*-acetylgalactosamine (GalNAc)-1-phosphate kinase, a new pathway of GalNAc activation. *J Biol Chem* **271**, 20776–20782.

47 Ainouz IL, Sampaio AH, Benevides NMB, Freitas ALP, Costa FHF, Carvalho MR & Pinheirojoventino F (1992) Agglutination of enzyme treated erythrocytes by Brazilian marine algal extracts. *Bot Mar* **35**, 475–479.

48 Schägger H & von Jagow G (1987) Tricine–sodium dodecyl sulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* **166**, 368–379.

49 Henrikson RL & Meredith SC (1984) Amino acid analysis by reversed-phase high-performance liquid chromatography: precolumn derivatization with phenylisothiocyanate. *Anal Biochem* **136**, 65–71.

50 Steenkamp J, Wiid I, Lourens A & van Helden P (1994) Improved method for DNA extraction from Vitis vinifera. *Am J Enol Vitic* **45**, 102–106.

51 Frohman MA & Martin GR (1989) Rapid amplification of cDNA ends using nested primers. *Techniques* **1**, 165–170.

52 Collaborative Computational Project Number 4 (1994). *Acta Cryst* **D50**, 760–763.