



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE BIOQUÍMICA E BIOLOGIA MOLECULAR
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA

ALICE ARAUJO DA SILVA

**ESTRUTURA PRIMÁRIA PARCIAL, MODELAGEM E *DOCKING* MOLECULAR DE
UMA LECTINA PRESENTE EM SEMENTES DE *Vataireopsis araroba* (Aguiar) Ducke**

FORTALEZA

2023

ALICE ARAUJO DA SILVA

ESTRUTURA PRIMÁRIA PARCIAL, MODELAGEM E *DOCKING* MOLECULAR DE UMA
LECTINA PRESENTE EM SEMENTES DE *Vataireopsis araroba* (Aguiar) Ducke

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Ceará, como parte dos requisitos para a obtenção do título de mestre em Bioquímica.

Orientador: Prof. Dr. Benildo Sousa Cavada.

Co-Orientadora: Prof.^a Dra. Kyria Santiago do Nascimento.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- S578e Silva, Alice Araujo da.
Estrutura primária parcial, modelagem e docking molecular de uma lectina presente em sementes de *Vataireopsis araroba* (Aguiar) Ducke / Alice Araujo da Silva. – 2023.
65 f. : il. color.
- Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Bioquímica, Fortaleza, 2023.
Orientação: Prof. Dr. Benildo Sousa Cavada.
Coorientação: Profa. Dra. Kyria Santiago do Nascimento.
1. *Vataireopsis araroba*. 2. Lectinas de Dalbergieae . 3. Estudo estrutural. I. Título.
- CDD 572
-

ALICE ARAUJO DA SILVA

ESTRUTURA PRIMÁRIA PARCIAL, MODELAGEM E *DOCKING* MOLECULAR DE UMA
LECTINA PRESENTE EM SEMENTES DE *Vataireopsis araroba* (Aguiar) Ducke

Dissertação apresentada ao Programa de Pós-Graduação em Bioquímica da Universidade Federal do Ceará, como parte dos requisitos para a obtenção do título de mestre em Bioquímica.

Aprovada em 10/02/2022

BANCA EXAMINADORA

Prof. Dr. Benildo Sousa Cavada (Orientador)
Universidade Federal do Ceará (UFC)

Dr. Messias Vital de Oliveira.
Universidade Federal do Ceará (UFC)

Dr. Vanir Reis Pinto Júnior.
Universidade Federal do Ceará (UFC)

A Deus, pela fidelidade em suas promessas. “Não percais, pois, vossa confiança, à qual está reservada uma grande recompensa. Vós precisais de constância, para que, depois de terdes cumprido a vontade de Deus, alcanceis a promessa” (Hebreus 10,35-39).

Aos meus pais Vanda e Raimundo (*in memoriam*), que nunca mediram esforços para me dar o melhor, sempre, e me apoiar em tudo.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

A Deus, pelo dom da vida, por sempre guiar e iluminar meu caminho e minhas escolhas, dando-me força e sabedoria para realização desta conquista. Por colocar pessoas em meu caminho que contribuíram para minha formação e amadurecimento.

A toda minha família, em especial minha mãe, que me apoiou, me auxiliou e sempre acreditou em mim durante toda esta caminhada. Amo a todos.

Ao meu amigo de todas as horas e namorado, Italo, quero agradecer pelo seu apoio, ajuda, torcida para vencer esta etapa da minha vida. Sou imensamente grata pela sua paciência e incentivo.

Aos meus orientadores, Prof. Dr. Benildo Sousa Cavada e Prof.^a Dra. Kyria Santiago do Nascimento, pelos ensinamentos, paciência, dedicação, e acima de tudo, pelo incentivo durante toda esta jornada da pós-graduação.

A todos os meus professores e professoras do Programa de Pós-Graduação em Bioquímica da Universidade Federal do Ceará, pelos ensinamentos, paciência e disponibilidade. Por tudo que aprendi com vocês, minha eterna gratidão.

Aos Doutores Messias Vital de Oliveira e Vanir Reis Pinto Júnior, pelos ensinamentos, amizade e apoio científico; por aceitarem participar desta banca examinadora e pelas contribuições a este trabalho.

A todos os meus colegas e amigos do Laboratório de Moléculas Biologicamente Ativas (BioMol-Lab), em especial a (o): Valeria, William, Joyce, Edilcarlos, Rebeca, Paula, Paulo, Henrique, Vinícius. Muito obrigada por estarem ao meu lado nos dias bons e ruins, pelo apoio nas dificuldades, pela ajuda nos experimentos, pelas merendas na copa, pelos sorrisos e abraços, pelas conversas descontraídas; enfim, por tudo que fizeram por mim. Amo vocês.

A todas as pessoas que de alguma forma fizeram parte deste percurso, eu agradeço com todo meu coração. Muito obrigada!

RESUMO

Lectinas são uma classe de (glico)proteínas capazes de se ligar reversivelmente a carboidratos específicos. As lectinas da tribo *Dalbergieae* têm despertado um interesse crescente acerca de estudos sobre bioprospecção, caracterização físico-química, estrutural e potencial biotecnológico. Dentre as diversas atividades biológicas relatadas na literatura, temos a atividade pró e anti-inflamatória, nociceptiva, vasorelaxante, antibacteriana, dentre outros. A lectina de *Vataireopsis araroba* (VaL) foi isolada a partir da farinha de sementes em uma única etapa de cromatografia de afinidade e apresentou especificidade para galactosídeos. Este trabalho teve como objetivo realizar estudos de determinação da estrutura primária (por espectrometria de massas) e tridimensional (por bioinformática) da VaL. Os resultados obtidos foram a determinação da sequência de aminoácidos parcial de VaL, cujas análises de alinhamento de sequências mostraram que a lectina apresentou um alto grau de identidade, em relação a sua estrutura primária, e um alto grau de cobertura com outras lectinas da tribo *Dalbergieae* específicas para *N*-acetilgalactosamina/galactose, como a lectina de *Vatairea macrocarpa*. A predição da estrutura secundária apresentou grande predominância de folhas beta, com porcentagem de 41,2% e 1,28% de hélice- α . O perfil de glicosilação apresentou sítios de *N*-glicosilação nos resíduos de asparagina nas posições 111, 114 e 183 da estrutura primária de VaL. A determinação da estrutura tridimensional foi realizada pelo método de modelagem por homologia, onde o protômero apresentou padrão conformacional semelhante a de outras lectinas de *Leguminosae*, sendo constituída pelo motivo *jellyroll*. Os resíduos participantes do sítio de ligação a metais (MBS) e do domínio de reconhecimento a carboidratos (CRD) se mostraram bastante conservados, em comparação com a VML. E por fim, as análises de *docking* molecular revelaram interações favoráveis com o antígeno Tn e *N*-acetilgalactosamina com escores de -45,01 e -37,06, respectivamente. O ligante Tn é um importante biomarcador de câncer, com isto a análise estrutural de novas lectinas que apresentam interações favoráveis com este ligante é de interesse considerável a fim de se elucidar o mecanismo de reconhecimento específico da proteína-carboidrato, bem como para desenvolver novas ferramentas potenciais no diagnóstico do câncer.

Palavras-chave: *Vataireopsis araroba*; lectinas de *Dalbergieae*; estudo estrutural.

ABSTRACT

Lectins are a class of (glyco)proteins capable of reversibly binding to specific carbohydrates. Lectins from the *Dalbergieae* tribe have aroused a growing interest in studies on bioprospecting, physical-chemical, structural characterization and biotechnological potential. Among the various biological activities reported in the literature, we have pro and anti-inflammatory, nociceptive, vasorelaxant, antibacterial activity, among others. *Vataireopsis araroba* lectin (VaL) was isolated from seed flour in a single affinity chromatography step and showed specificity for galactosides. This work aimed to carry out studies to determine the primary (by mass spectrometry) and three-dimensional (by bioinformatics) structure of VaL. The results obtained were the determination of the partial amino acid sequence of VaL, whose sequence alignment analyzes showed that the lectin presented a high degree of identity, in relation to its primary structure, and a high degree of coverage with other lectins of the tribe *Dalbergieae* specific for N-acetylgalactosamine/galactose, such as *Vatairea macrocarpa* lectin. The secondary structure prediction showed a great predominance of beta sheets, with a percentage of 41.2% and 1.28% of α -helix. The glycosylation profile showed N-glycosylation sites on asparagine residues at positions 111, 114 and 183 of the primary structure of VaL. The determination of the three-dimensional structure was carried out by the homology modeling method, where the protomer presented a conformational pattern similar to that of other *Leguminosae* lectins, being constituted by the jellyroll motif. Residues participating in the metal binding site (MBS) and carbohydrate recognition domain (CRD) were highly conserved compared to VML. Finally, molecular docking analyzes revealed favorable interactions with the Tn and N-acetylgalactosamine antigens with scores of -45.01 and -37.06, respectively. The Tn ligand is an important cancer biomarker, therefore the structural analysis of new lectins that present favorable interactions with this ligand is of considerable interest in order to elucidate the mechanism of specific recognition of the protein-carbohydrate, as well as to develop new tools potential in cancer diagnosis.

Keywords: *Vataireopsis araroba*; *Dalbergieae* lectins; structural study.

LISTA DE ILUSTRAÇÕES

Figura 1 –	Ilustração da classificação das lectinas vegetais com base na sua diversidade estrutural.....	15
Figura 2 –	Estruturas tridimensionais de diferentes lectinas de leguminosas.....	16
Figura 3 –	Estrutura geral das lectinas de leguminosas.....	18
Figura 4 –	Alinhamento da sequência parcial da lectina de <i>Vataireopsis araroba</i> com as sequências de algumas lectinas de <i>Dalbergieae</i>	22
Figura 5 –	Morfologia representativa da espécie <i>Vataireopsis araroba</i>	24
Figura 6 –	Etapas de predição de estrutura tridimensional de proteínas usando o método de Modelagem Comparativa.....	27
Figura 7 –	Perfil cromatográfico e eletroforético da VaL.....	39
Figura 8 –	Cromatograma da VaL, sob condição nativa, em coluna SEC.....	40
Figura 9 –	Alinhamento da sequência parcial da VaL com as sequências obtidas do UniProt de outras lectinas da subfamília <i>Papilionoideae</i> .	43
Figura 10 –	Alinhamento da sequência parcial da VaL com sequências de aminoácidos de outras lectinas que possuem sua estrutura tridimensional resolvida e depositada no banco de dados PDB.....	45
Figura 11 –	Árvore filogenética construída a partir de sequências de aminoácidos, obtidos do UniProt, de lectinas da subfamília <i>Papilionoideae</i>	46
Figura 12 –	Árvore filogenética construída a partir de sequências de aminoácidos de lectinas que possuem sua estrutura tridimensional resolvida e depositada no banco de dados PDB.....	46
Figura 13 –	Predição da estrutura secundária da lectina de <i>Vataireopsis araroba</i>	47
Figura 14 –	Representação do monômero de VaL complexado com o antígeno Tn.....	50
Figura 15 –	Representação tridimensional do MBS da VaL.....	51
Figura 16 –	Representação do CRD da VaL ligado aos carboidratos <i>L</i> -ramnose (A), <i>D</i> -galactose (B), α -lactose (C), <i>N</i> -acetil- <i>D</i> -galactosamina (GalNAc) (D) e o antígeno Tn (E).....	54
Figura 17 –	Sobreposição das estruturas tridimensionais da VaL com a estrutura tridimensional da VML.....	56

LISTA DE TABELAS

Tabela 1 – Sequência dos peptídeos obtidos por espectrometria de massas em tandem (MS/MS).....	42
Tabela 2 – Valores referentes aos parâmetros de validação para a predição da estrutura terciária da lectina de <i>Vataireopsis araroba</i>	49
Tabela 3 – Valores de escores, encaixes, interações de hidrogênio, interações hidrofóbicas e torções obtidas no <i>docking</i> de VaL com diferentes carboidratos.....	52

LISTA DE ABREVIATURAS E SIGLAS

ABA	Aglutininas de <i>Agaricus bisporus</i>
ADL	Lectina de <i>Arachis duranensis</i>
AIL	Lectina de <i>Arachis ipaensis</i>
CML	Lectina de <i>Centrolobium microchaete</i>
CRA	Aglutininas relacionadas à quitinase
CRD	Domínio de Reconhecimento a Carboidratos
CTL	Lectina de <i>Centrolobium tomentosum</i>
EUL	Lectinas relacionadas ao <i>Euonymus europaeus</i>
GNA	Lectinas relacionadas às Aglutininas de <i>Galanthus nivalis</i>
LAL	Lectina de <i>Lonchocarpus araripensis</i>
LSL	Lectina de <i>Lonchocarpus sericeus</i>
LysM	Lectinas do domínio Lysin M
MBS	Sítio de Ligação de Metal
PAL	Lectina de <i>Pterocarpus angolensis</i>
PDB	<i>Protein Data Bank</i>
PELa	Lectina de <i>Platypodium elegans</i>
PNA	<i>Peanut agglutinin</i>
PRL	Lectina de <i>Pterocarpus rotundifolius</i>
SDS	Sodium Dodecyl Sulfate
SL-I	Lectina da haste de <i>Arachis hypogaea</i> I
UFC	Universidade Federal do Ceará
VaL	Lectina de <i>Vataireopsis araroba</i>
VGL	Lectina de <i>Vatairea guianensis</i>
VML	Lectina de <i>Vatairea macrocarpa</i>

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Lectinas vegetais – Considerações iniciais	14
1.2	Lectinas de Leguminosas	16
1.3	Subfamília <i>Papilionoideae</i>	18
1.4	Lectinas da tribo <i>Dalbergieae</i>	19
1.5	Caracterização estrutural das lectinas de <i>Dalbergieae</i>	22
1.6	<i>Vataireopsis araroba</i> (Aguiar) Ducke	24
1.7	Ferramentas de bioinformática utilizadas no estudo das lectinas	25
1.7.1	<i>Modelagem tridimensional</i>	26
1.7.2	<i>Docking molecular</i>	28
1.8	Importância biotecnológica das lectinas de <i>Dalbergieae</i>	29
1.8.1	<i>Atividade anticâncer</i>	30
2	OBJETIVOS	31
2.1	Objetivo Geral	31
2.2	Objetivos Específicos	31
3	MATERIAIS E MÉTODOS	32
3.1	Material vegetal	32
3.2	Extração e purificação da lectina de sementes de <i>Vataireopsis araroba</i>	32
3.3	Determinação da pureza da VaL por Cromatografia de exclusão molecular	33
3.4	Eletroforese em condições desnaturantes	33
3.5	Digestão proteica e sequenciamento dos peptídeos por espectrometria de massa em tandem (MS/MS)	34
3.6	Análise filogenética e predição da estrutura secundária	35
3.7	Modelagem tridimensional	36
3.8	<i>Docking molecular</i>	36
3.9	Avaliação da toxicidade contra náuplios de <i>Artemia salina</i>	37
4	RESULTADOS E DISCUSSÕES	39
4.1	Purificação da lectina de sementes de <i>Vataireopsis araroba</i> (VaL)	39
4.2	Cromatografia de exclusão molecular em sistema UPLC	40
4.3	Estrutura primária parcial de VaL	41

4.4	Alinhamento de sequências e Análise filogenética.....	43
4.5	Predição da estrutura secundária.....	47
4.6	Modelagem tridimensional.....	49
4.7	<i>Docking</i> molecular.....	51
4.8	Toxicidade contra náuplios de <i>Artemia salina</i>.....	56
5	CONCLUSÃO.....	58
	REFERÊNCIAS.....	59

1 INTRODUÇÃO

1.1 Lectinas vegetais – Considerações iniciais

As lectinas são um grupo heterogêneo de proteínas com pelo menos um domínio não catalítico que reconhece e se liga reversivelmente a açúcares livres específicos ou glicanos presentes em glicoconjugados (PEUMANS; VAN DAMME, 1995; LANNOO; VAN DAMME, 2010). É esse reconhecimento altamente específico a carboidratos que permite que essas proteínas se liguem diferentemente a determinadas moléculas e células. E devido a essa seletividade que as lectinas têm sido amplamente utilizadas para a investigação de diversos processos fisiológicos e patológicos, fornecendo evidências de seu amplo espectro de aplicações industriais e farmacológicas (MARTÍNEZ-ALARCÓN *et al.*, 2018)

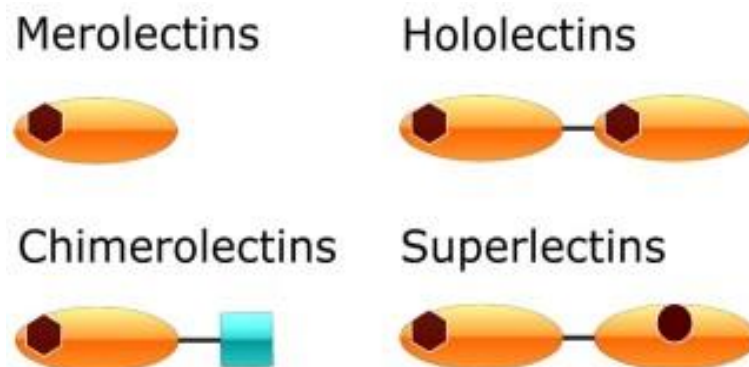
Os primeiros relatos a respeito do descobrimento destas moléculas datam de 1888, quando Peter Hermann Stillmark descobriu uma proteína altamente tóxica das sementes de mamona (*Ricinus communis* L.), chamada de ricina (STILLMARK, 1888). Além desta descoberta, as proteínas também desempenharam, durante experimentos executados, a capacidade de aglutinar eritrócitos (hemácias). Este fenômeno foi mais tarde referido como 'aglutinação', e deu origem à introdução do termo 'aglutinina' ou 'fitohemaglutinina' (ELFSTRAND, 1898). Watkins e Morgan, em 1952, fizeram outra descoberta importante a respeito das propriedades de hemaglutinação das proteínas. Eles conseguiram mostrar que a adição de açúcares específicos para cada tipo de eritrócitos pode bloquear a aglutinação causada pelas lectinas. Desta forma, foi descoberto e comprovado que a atividade de aglutinação das lectinas é baseada no reconhecimento e ligação de estruturas de carboidratos na superfície celular dos glóbulos vermelhos. Com isso o termo 'hemaglutinina' foi substituído por 'lectina', derivado do verbo latino *legere*, que se refere à capacidade da proteína de selecionar ou escolher (WATKINS; MORGAN, 1952 *apud* TSANEVA; VAN DAMME, 2020).

As lectinas estão distribuídas nos diversos reinos da vida, no entanto o maior número de lectinas estudadas pertence ao reino vegetal. Elas podem ser encontradas em tecidos de armazenamento (ou seja, sementes, bulbos, cascas, rizomas), xilema e floema, e em menor grau em raízes, brotos, folhas e flores. A nível celular, as lectinas estão presentes no núcleo e/ou citoplasma, no vacúolo ou na superfície celular. As lectinas presentes no vacúolo são geralmente muito abundantes e residem em corpos proteicos. Já as lectinas presentes no núcleo e/ou citoplasma geralmente estão em baixas quantidades, mas são reguladas

positivamente após a exposição da planta a estresses bióticos e/ou abióticos (LANNOO; VAN DAMME, 2010; DE CONINCK; VAN DAMME, 2022).

Levando em conta a caracterização estrutural, as lectinas vegetais representam um grande grupo de proteínas com estruturas moleculares altamente diversificadas, com dobras tridimensionais. Uma lectina vegetal que apresenta um domínio de ligação a carboidratos é classificada como uma 'merolectina'. Quando a proteína compreende dois ou mais domínios de ligação a carboidratos, é denominada como "hololectina"; enquanto uma proteína que apresenta uma combinação de um domínio de ligação a carboidrato ligado a pelo menos um outro domínio de ligação com atividade biológica diferente, é classificada como "quimerolectina". Por fim, as lectinas vegetais que são compostas por vários domínios de ligação a carboidratos com diferentes sítios de ligação, neste caso, são chamadas de "superlectinas" (Figura 1) (PEUMANS; VAN DAMME, 1995; TSANEVA; VAN DAMME, 2020).

Figura 1 – Ilustração da classificação das lectinas vegetais com base na sua diversidade estrutural.



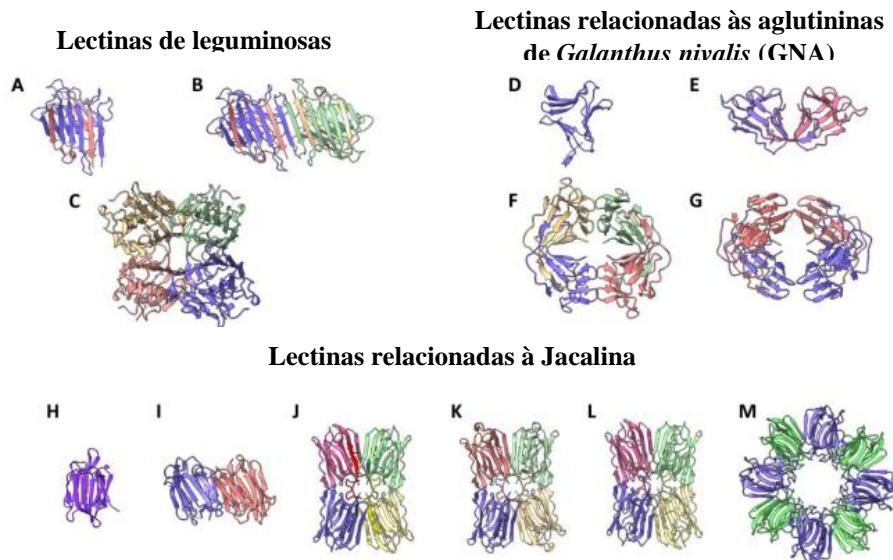
Fonte: TSANEVA; VAN DAMME, 2020.

No que se refere à estrutura primária, uma análise minuciosa das sequências de lectinas disponíveis a partir de estudos do genoma e transcriptoma mostra que todas as lectinas de plantas conhecidas hoje podem ser classificadas em famílias diferentes, com base em suas sequências e na conformação de seu(s) domínio(s) de reconhecimento a carboidratos (*Carbohydrate Recognition Domain* - CRD). Em 1998, o sistema de classificação de lectinas vegetais consistia em 7 famílias (VAN DAMME *et al.*, 1998), mas em 2008 foi estendido para 12 famílias, para acomodar os dados de todas as lectinas de plantas recém-identificadas (VAN DAMME, 2008). As famílias das lectinas recebem o nome da primeira lectina desta família que foi caracterizada em detalhes. Em ordem alfabética, são: a família de aglutininas

de *Agaricus bisporus* (ABA), as Amaranquinas, as aglutininas relacionadas à quitinase (CRA), as lectinas com domínio Cianovirina, as lectinas relacionadas ao *Euonymus europaeus* (EUL), as lectinas relacionadas às aglutininas de *Galanthus nivalis* (GNA), as lectinas com domínio Heveínico, as lectinas relacionadas à Jacalina, as lectinas de leguminosas, as lectinas com domínio Lysin M (LysM), as lectinas relacionadas à aglutinina de *Nicotiana tabacum* (Nictaba) e as lectinas com domínio Ricina-B (TSANEVA; VAN DAMME, 2020).

A diversidade de famílias de lectinas vegetais levou alguns pesquisadores a buscar motivos de lectinas (pequenos padrões de sequência conservada associados à atividade de ligação a carboidratos) em cada família de lectinas e obter uma visão mais profunda da especificidade de diferentes lectinas. A estrutura tridimensional da maioria dos domínios de lectina contém um alto teor de folhas β , e um baixo teor de α -hélices. Estas folhas β se organizam em estruturas β -sanduíche, β -trefoil, β -prisma e β -barril, dependendo do domínio da lectina. As estruturas quaternárias geralmente contêm duas ou quatro subunidades estabilizadas por interações hidrofóbicas, pontes de hidrogênio e interações eletrostáticas (Figura 2) (BONNARDEL *et al.*, 2020).

Figura 2 – Estruturas tridimensionais de diferentes lectinas de leguminosas, lectinas relacionadas às aglutininas de *Galanthus nivalis* (GNA) e lectinas relacionadas à Jacalina.



Fonte: VAN DAMME, 2021.

1.2 Lectinas de Leguminosas

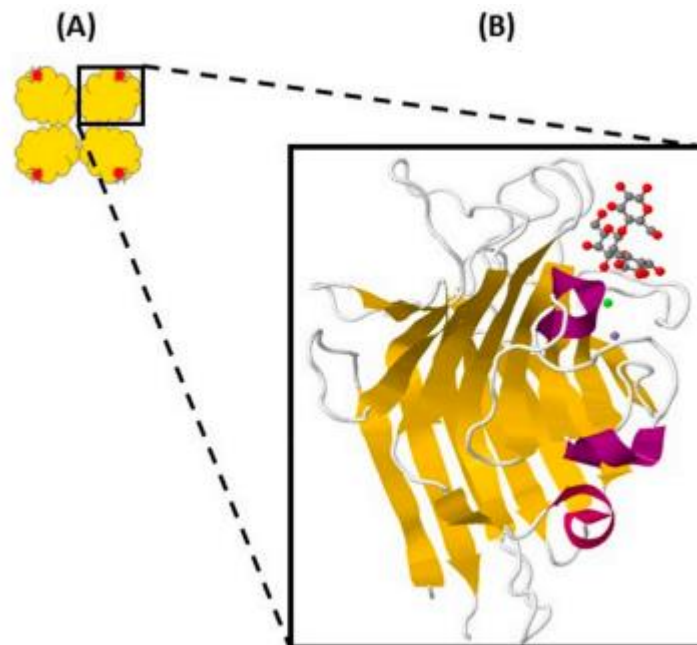
A família *Leguminosae* compreende 727 gêneros e foi subdividida em três subfamílias, *Caesalpinioideae*, *Mimosoideae* e *Papilionoideae*, esta última contendo a maior concentração de estudos sobre lectinas (LAGARDA-DIAZ *et al.*, 2017; CAVADA *et al.*, 2020; NASCIMENTO *et al.*, 2020). No entanto, esta subdivisão foi considerada desatualizada por muitos pesquisadores e em 2017 foi alterada pelo Legume Phylogeny Working Group. Essa nova classificação usa como estrutura as análises filogenéticas mais abrangentes de leguminosas, com base nas sequências de genes (AZANI *et al.*, 2017). Com isso, a classificação da família de *Leguminosae* reconheceu seis subfamílias, incluindo *Duparquetioideae*, *Cercidoideae*, *Detarioideae*, *Dialioideae*, *Papilionoideae* e *Caesalpinioideae*. Gêneros da subfamília *Mimosoideae* agora fazem parte de *Caesalpinioideae*, e alguns gêneros que faziam parte de *Caesalpinioideae* agora estão incluídos em outras subfamílias (AZANI *et al.*, 2017; ESTRELLA *et al.*, 2018).

As lectinas de leguminosas apresentam grande diversidade de especificidades de ligação a carboidratos, maior do que qualquer outro grupo de lectinas vegetais, por isso suas propriedades físico-químicas e biológicas têm sido amplamente estudadas e suas estruturas elucidadas (SHARON *et al.*, 1990 *apud* LAGARDA-DIAZ *et al.*, 2017). A maioria das lectinas de leguminosas são constituídas por protômeros de 25 a 30 kDa, entre 220 e 250 aminoácidos, e formadas por uma ou mais cadeias polipeptídicas. Além disso, elas podem ser *N*-glicosiladas e conter uma ou mais cadeias de glicanos. Com algumas exceções, a maioria dessas lectinas de leguminosas são metaloproteínas, protômeros em lectinas vegetais com um sítio de ligação de metal (*Metal Binding Site* - MBS) apresentando normalmente íons cálcio e manganês fortemente ligados, responsáveis pela manutenção do CRD. As lectinas podem formar dímeros e/ou tetrâmeros e se ligar especificamente a resíduos de manose, galactose, glicose, fucose e seus derivados (CAVADA *et al.*, 2020).

Em se tratando da caracterização estrutural, a estrutura quaternária das lectinas de leguminosas pode ser muito variável, como resultado de modificações como a glicosilação, embora as estruturas primária, secundária e terciária dos monômeros sejam muito semelhantes (CAVADA *et al.*, 2019; LAGARDA-DIAZ *et al.*, 2017). O protômero das lectinas de leguminosas tem uma dobra terciária do tipo *jellyroll* que consiste em duas folhas β (beta) principais: uma folha β parcialmente estendida de seis fitas e uma folha β curvada de sete fitas, antiparalelas e unidas entre si por alças de tamanhos variados, formando um sanduíche β , estabilizado pela presença de dois núcleos hidrofóbicos (Figura 3) (NECO *et al.*, 2018). Embora a estrutura tridimensional das lectinas de leguminosas seja conservada, as alças importantes para a ligação de carboidratos são bastante variáveis, explicando por que o

domínio das lectinas de leguminosas pode se ligar a toda uma gama de estruturas de carboidratos (LAGARDA-DIAZ *et al.*, 2017).

Figura 3 – Estrutura geral das lectinas de leguminosas. (A) tetrâmero com domínios de reconhecimento de carboidratos (vermelho); (B) imagem ampliada do monômero com folhas β (amarelo), sítios de ligação de metal (área com esferas verdes e cinzas) e o domínio de reconhecimento a carboidratos (área ocupada pela molécula cinza e vermelha).



Fonte: LAGARDA-DIAZ *et al.*, 2017.

1.3 Subfamília *Papilionoideae*

Papilionoideae constitui-se a maior subfamília de *Leguminosae*, sendo formada por 28 tribos, 478 gêneros e 13.800 espécies (LEWIS *et al.*, 2005). Esta subfamília distingue-se vegetativamente das outras subfamílias, pelos caracteres florais e do fruto, sendo considerada monofilética. Além disso, a subfamília *Papilionoideae* apresenta uma ampla distribuição territorial, ocorrendo em regiões que abrangem desde as florestas pluviais até as regiões desérticas, com climas secos e frios (POLHILL; RAVEN, 1981). Várias sementes de leguminosas, da subfamília *Papilionoideae*, são bastante conhecidas por seu grande impacto econômico na agricultura e na indústria alimentícia, como a soja (*Glycine max*), o feijão (*Phaseolus spp.* e *Vicia faba L.*), o amendoim (*Arachis hypogaea L.*), a lentilha (*Lens*

culinarius Medik.), a ervilha (*Pisum sativum* L.) e o grão-de-bico (*Cicer arietinum* L.) (CARDOSO *et al.*, 2013).

Taxonomicamente, a subfamília *Papilionoideae* apresenta uma variação na morfologia floral. Gêneros de tribos mais primitivas, como *Swartzieae* e *Sophoreae*, exibem flores com maior simetria radial do que simetria bilateral, e pétalas pouco diferenciadas com estames livres. A tribo *Swartzieae* possui 17 gêneros, predominantemente neotropicais, compreendendo cerca de 258 espécies. Já a tribo *Sophorea*, apresenta 45 gêneros e cerca de 396 espécies, distribuídas em todo o mundo (CARDOSO *et al.*, 2013).

1.4 Lectinas da tribo *Dalbergieae*

Dalbergieae é uma das tribos da subfamília *Papilionoideae*, que compreende 49 gêneros e mais de 1300 espécies (LEWIS *et al.*, 2005). Estudos acerca da filogenética molecular desta tribo propuseram classificações baseadas em características, como morfologia da planta e comparação de sequências de genes. A tribo *Dalbergieae* é a base da evolução da subfamília *Papilionoideae*, estando dentro de uma escala filogenética distante de tribos mais evoluídas como *Phaseoleae*. Assim, ao longo da diferenciação dessas espécies, as pequenas variações nos genomas das espécies plesiomórficas podem ter levado a diferenças no consenso filogenético (POLHILL; RAVEN, 1981).

As lectinas desta tribo têm despertado um interesse crescente acerca de estudos sobre caracterização estrutural, propriedades físico-químicas e atividades biológicas. Até o momento, um total de 27 lectinas vegetais desta tribo foram purificadas, abrangendo os seguintes gêneros: *Andira*, *Arachis*, *Centrolobium*, *Lonchocarpus*, *Machaerium*, *Platypodium*, *Platymiscium*, *Pterocarpus*, *Vatairea* e *Vataireopsis* (NASCIMENTO *et al.*, 2020). O primeiro relato na literatura dessas lectinas data de 1975, com a descrição da purificação de PNA (*Peanut agglutinin*), a primeira lectina isolada de amendoim da espécie *Arachis hypogaea* (LOTAN *et al.*, 1975). Outro trabalho pioneiro envolve a purificação da lectina de *Vatairea macrocarpa* (VML). A elucidação da sequência de aminoácidos da VML foi fundamental para compreender como se dá o processamento pós-traducional das lectinas de *Dalbergieae* (CAVADA *et al.*, 1998).

Estruturalmente, as lectinas de *Dalbergieae* são bastante semelhantes em seu enovelamento tridimensional, mas apresentam diferenças significativas nos padrões de oligomerização e na conservação do domínio de reconhecimento de carboidratos. A caracterização dessas lectinas permitiu a divisão em três grupos, com base na especificidade a

carboidratos: as lectinas específicas para manose/glicose, as lectinas específicas para *N*-acetilgalactosamina/galactose e as lectinas específicas para *N*-acetilglucosamina (NASCIMENTO *et al.*, 2020). A maioria das lectinas isoladas desta tribo apresentam especificidade para glicose/manose e seus derivados de carboidratos, como por exemplo, as lectinas de *Pterocarpus angolensis* (PAL), *Platypodium elegans* (PELa) e *Centrolobium tomentosum* (CTL), que se ligam a resíduos de oligomanose e *N*-glicanos complexos (LORIS *et al.*, 2003; ARARIPE *et al.*, 2017; ALMEIDA *et al.*, 2016). No caso das lectinas que apresentaram especificidade para *N*-acetilgalactosamina/galactose, as lectinas do gênero *Vatairea* e *Arachis* são as principais representantes desse grupo (MARQUES *et al.*, 2017; SOUSA *et al.*, 2016). E por fim, as lectinas de *Lonchocarpus sericeus* (LSL) e *Lonchocarpus araripensis* (LAL) foram as únicas lectinas de *Dalbergieae* que mostraram especificidade para *N*-acetilglucosamina (ALENCAR *et al.*, 2005; PIRES *et al.*, 2016).

As propriedades físico-químicas para a maioria das lectinas de *Dalbergieae* levam em conta sete propriedades funcionais relacionadas à estabilidade, que são: especificidade de ligação ao açúcar, tipo de eritrócitos testados, glicosilação de proteínas, dependência de íons metálicos, faixa de pH ideal, termoestabilidade e perfil eletroforético. Com relação à especificidade de ligação a carboidratos, esta propriedade foi citada anteriormente. Já com relação ao perfil de glicosilação, as lectinas de *Dalbergieae* apresentaram porcentagens de carboidratos em suas estruturas variando de 2% a 40% (NASCIMENTO *et al.*, 2020).

No que se refere à dependência de íons metálicos, das lectinas testadas, algumas se mostraram dependentes de íons catiônicos, como Ca^{2+} , Mn^{2+} e Mg^{2+} , para sua atividade hemaglutinante. Em se tratando do perfil eletroforético das lectinas de *Dalbergieae*, foram observadas diferenças na quantidade de bandas de proteína e peso molecular em géis de poli-acrilamida (SDS-PAGE). Para esta tribo, foi observado na literatura pelo menos dois grupos de lectinas categorizados de acordo com suas diferentes especificidades. Primeiro, as lectinas específicas à glicose/manose (maioria) apresentaram uma única cadeia polipeptídica contendo um único sítio de glicosilação e um peptídeo sinal, mas sem existência evidente de processamento pós-traducional. O segundo grupo, são lectinas que compartilham características de especificidade para galactose e seus derivados, como o gênero *Vatairea*, consistindo de pelo menos três cadeias polipeptídicas (α , β e γ) e dois sítios de glicosilação. A eletroforese em gel desnaturante das lectinas do gênero *Vatairea* apresenta um perfil composto por uma banda dupla de 32–34 kDa referente ao polipeptídeo da cadeia α com glicosilações duplas e simples. E duas bandas menores de 22 e 13 kDa que correspondem à

cadeia β e γ , respectivamente. Essas duas últimas bandas resultam do processamento pós-traducional do polipeptídeo da cadeia α (ALVES FILHO *et al.*, 2013).

As lectinas da tribo *Dalbergieae* possuem propriedades físico-químicas de pH e temperatura muito semelhantes. A maioria das lectinas mantém sua atividade na faixa de pH 4-10 com atividade hemaglutinante máxima na faixa de pH 7-8. O perfil de termoestabilidade mostrou que as lectinas mantêm uma atividade ótima nas temperaturas entre 40 e 50 °C, sendo que a maioria pode atingir atividade máxima em temperaturas até 60 °C, porém acima dessa temperatura, em torno de 70–100 °C, ocorre desnaturação e conseqüentemente perda total ou parcial da atividade hemaglutinante (NASCIMENTO *et al.*, 2020).

Com relação aos tipos de eritrócitos testados, várias espécies da tribo *Dalbergieae* apresentaram atividade hemaglutinante com eritrócitos de coelho, humano (ABO), rato, vaca, ovelha, e suínos, tanto eritrócitos nativos quanto tratados com enzimas proteolíticas. Em geral, os dados obtidos mostraram que as lectinas específicas a manose têm preferência por eritrócitos de coelho em comparação aos humanos, enquanto as lectinas específicas à galactose normalmente demonstram preferência por eritrócitos humanos. Isso pode ser explicado pela abundância de galactose e porções *N*-acetilgalactosamina na superfície deste último (NASCIMENTO *et al.*, 2020).

O estudo acerca da caracterização estrutural das lectinas de *Dalbergieae* se iniciou em 1975, quando primeira lectina desta tribo teve sua composição parcial de aminoácidos analisada, a lectina específica à galactose presente em semente de *Arachis hypogaea* (LOTAN *et al.*, 1975). Em 1985, Lauwereys e colaboradores (LAUWEREYS *et al.*, 1985) conseguiram determinar a estrutura primária parcial da lectina de *Arachis hypogaea* (PNA), em um trabalho que apresentava quase 60% da sequência de aminoácidos. A sequência completa de 236 resíduos, no entanto, foi elucidada somente em 1991 (YOUNG *et al.*, 1991). Mais tarde, em 1998, a sequência completa de aminoácidos da lectina de *Vatairea macrocarpa* (VML) foi elucidada (CALVETE *et al.*, 1998). Em 2012, uma nova lectina específica a galactose das sementes de *Vatairea guianensis* (VGL) foi purificada e teve sua estrutura primária determinada por espectrometria de massas. A VGL apresentou 95% de similaridade de sequência com VML (SILVA *et al.*, 2012). Com o alinhamento das sequências de aminoácidos obtidas na literatura, foi observado que a identidade de sequência das lectinas de *Dalbergieae* específicas a manose/glicose varia de um mínimo de 66,12% entre a lectina presente na haste de *Arachis hypogaea* I (SL-I) e a lectina de *Pterocarpus angolensis* (PAL), e um máximo de 99% entre a lectina de *Arachis ipaensis* (AIL) e SL-I. Já as lectinas específicas

para galactose apresentam uma identidade significativamente menor, entre os gêneros *Vatairea* e *Arachis*, com valores de 43% entre a VML e a PNA (NASCIMENTO *et al.*, 2020).

Apesar da existência de dados estruturais de algumas lectinas desta tribo, apenas um estudo, registrado na literatura, foi realizado acerca da purificação, caracterização físico-química e biológica da lectina presente em sementes de *Vataireopsis araroba* (Aguiar) Ducke (OSTERNE *et al.*, 2022). Com isso, a caracterização estrutural de novas lectinas contribuem para um melhor entendimento acerca da relação evolutiva entre as diferentes tribos da subfamília *Papilionoideae*, do ponto de vista bioquímico, e para uma melhor compreensão dos papéis biológicos e possíveis aplicações das lectinas.

1.5 Caracterização estrutural das lectinas de *Dalbergieae*

O estudo realizado por Osterne e colaboradores (OSTERNE *et al.*, 2022) apresentou a estrutura primária parcial da lectina de *Vataireopsis araroba*. A sequência parcial determinada apresentou um total de 109 resíduos de aminoácidos e corresponde a cerca de 45,41% da sequência total prevista, a julgar pelos alinhamentos de sequência com outras lectinas da tribo *Dalbergieae*. O alinhamento no banco de dados não redundante (BLASPP) revelou grau de similaridade acima de 70% com as sequências das lectinas de *Vatairea macrocarpa* (VML) (UNIPROT ID: P81371), *Vatairea guianensis* (VGL) (UNIPROT ID : P86893), *Arachis hypogaea* (PNA) (UNIPROT ID: P02872) e aglutinina I de *Robinia pseudoacacia* (RPbAI) (UNIPROT ID: Q41162), todas elas são lectinas de *Dalbergieae* específicas para galactose (Figura 5) (OSTERNE *et al.*, 2022).

Figura 5 – Alinhamento da sequência parcial da lectina de *Vataireopsis araroba* (VaL) com as sequências de algumas lectinas de *Dalbergieae*. Os aminoácidos destacados em vermelho representam resíduos semelhantes.

```

1      10      20      30      40      50
Arachis_hypogaea_PNA      AATVSNFNFNSSEGHPAIFNFGDVTVLNNTQLTNK...VMVGRVLYAMIVRITKS
Robinia_pseudoacacia_RPbAI  TGSLSFPPFAPNQPYIINQGDALVTSITVQLQLINVNGVFSKSLGRALYAAFFQIMD
Vataireopsis_araroba_Val    SEVVSFPFTFNPFPK.....
Vatairea_macrocarpa_VML     SEVVSFPFTFNPFPKDIILGGDALVTSIKKRLQLTKRDGEPVDHSLGRALYAAFIHIMD
Vatairea_guianensis_VGL    SEVVSFPFTFNPFPKDIILGGDALVTSIKKRLQLTKRDGEPVDHSLGRALYAAFIHIMD
consensus>70              .e.vSfSf.kf.pn.....qgd.v.s.g..qit.....s.gr.ly..p..iw

60     70     80     90     100    110
Arachis_hypogaea_PNA      SAAGNVASFLTSFSEMKDTRKYDFADGLFFIAPFDTCIPAGSIGGGHIGVSD...TKG
Robinia_pseudoacacia_RPbAI  STAGNVASFLTSFPIIQAFNPATADGLFFIAPFDTCIPAGSIGGGHIGVSD...TKG
Vataireopsis_araroba_Val    .....SIVVEATDESSTADGLIAPFDTCIPAGSIGGGHIGVSD...TKG
Vatairea_macrocarpa_VML     DSTIDRVASFAISFVVEATDESSTADGLIAPFDTCIPAGSIGGGHIGVSD...TKG
Vatairea_guianensis_VGL    DSTIDRVASFAISFVVEATDESSTADGLIAPFDTCIPAGSIGGGHIGVSD...TKG
consensus>70              ..t..vasf.tsfsfvveapde..tADGLiFFIAP.DTQp.....GG.LG.fnd...nk

120    130    140    150    160    170
Arachis_hypogaea_PNA      AGHFVVEFDTSNSENVDPEIDNVGIDVNSVDSVXTVPMNSVSAVVMVTVIYDSSTKT
Robinia_pseudoacacia_RPbAI  SNQIVVVEFDTSNNGDW.DPKGRHLGIVNSIEIINTVPMNWTNDEVANVFIISYEAATKS
Vataireopsis_araroba_Val    ...VVVEFDTSN.TW.DPSAR.....
Vatairea_macrocarpa_VML     SIQIVVVEFDTSN.TW.DPSARHIGIVNSIEIMRYIKGWENDEVANVYISYEAATKT
Vatairea_guianensis_VGL    SIQIVVVEFDTSN.TW.DPSARHIGIVNSIEISRYIKGWENDEVANVYISYEAATKT
consensus>70              ...VaVEFDTSN..w.DP..rh.ginvsies.k.v.w...g.v..v.i.ye.stk.

180    190    200    210    220    230
Arachis_hypogaea_PNA      LSVAVTNDVSGDITTAQVVDLRAKLPKVVVFGFSATGEGCG..RQIHLIRSHPTTILI
Robinia_pseudoacacia_RPbAI  LTASLVYPLLEYSFIIIDAINVVKIIMLPKVVVFGFSATGEGCG..RQIHLIRSHPTTILI
Vataireopsis_araroba_Val    .....ALPKVVVFGFSATGEGCG..RQIHLIRSHPTTILI
Vatairea_macrocarpa_VML     LTASLVYPLNATSYIVSANVQLKSAKLPKVVVFGFSATGEGCG..RQIHLIRSHPTTILI
Vatairea_guianensis_VGL    LTASLVYPLNATSYIVSANVQLKSAKLPKVVVFGFSATGEGCG..RQIHLIRSHPTTILI
consensus>70              l.....i...vd.k..LPEwVr.GFSAt.gl...v#thd!l.ws..s...

Arachis_hypogaea_PNA      TITRRS.
Robinia_pseudoacacia_RPbAI  G.....
Vataireopsis_araroba_Val    .....
Vatairea_macrocarpa_VML     APSDDSN
Vatairea_guianensis_VGL    ANSDYT.
consensus>70              .....

```

Fonte: OSTERNE *et al.*, 2022

Algumas lectinas de *Dalbergieae* tiveram suas estruturas tridimensionais resolvidas por métodos cristalográficos e, atualmente, cinco lectinas estão depositadas no Banco de Dados de Proteínas (PDB), incluindo PNA (BARNERJEE *et al.*, 1996), PAL (LORIS *et al.*, 2003), PELa (CAVADA *et al.*, 2018), VML (SOUSA *et al.*, 2016) e CTL (ALMEIDA *et al.*, 2016). Além disso, a SL-1 (PATHAK *et al.*, 2006), VGL (MARQUES *et al.*, 2017), CML (NECO *et al.*, 2018), AIL e ADL (NASCIMENTO *et al.*, 2018) também tiveram suas estruturas terciárias determinadas por modelagem molecular.

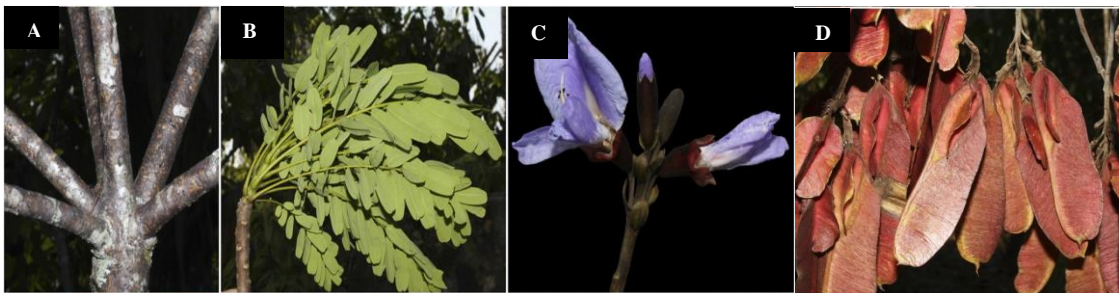
Os monômeros dessas lectinas de *Dalbergieae* podem se associar e formar dímeros e tetrâmeros. Embora esses monômeros apresentem a mesma dobra β -sanduíche, que pode ser chamado de motivo *jellyroll* ou dobra de lectina leguminosa, sua estrutura quaternária é bastante diversa e seus padrões de oligomerização diferenciados. Dessa forma, a determinação de novas estruturas de lectinas, especialmente as de grupos menos conhecidos como o gênero *Vataireopsis*, é de grande importância, pois essas novas estruturas de lectinas podem corresponder a propriedades distintas e possíveis aplicações que ainda não foram

identificadas nas lectinas dos demais gêneros, além de possibilitar um estudo em escala filogenética e um alinhamento com lectinas próximas, da tribo *Dalbergieae*, podendo-se levar em consideração as diferenças no consenso filogenético.

1.6 *Vataireopsis araroba* (Aguiar) Ducke

A espécie *Vataireopsis araroba* (Aguiar) Ducke é endêmica da floresta tropical brasileira, localizada no sul da Bahia e no norte do Espírito Santo. Pertence à tribo *Dalbergieae*, da subfamília *Papilionoideae* (família Leguminosae), e é uma árvore caducifólia alta, que pode atingir entre 20–35 m de altura. Sua madeira é amarela ou marrom escura com densidade de 0,60 g/cm³, é considerada fétida e de sabor amargo. Suas folhas apresentam até 50 folíolos e suas flores, em panículas eretas, são agrupadas na extremidade dos ramos (Figura 4). Seu período de floração se estende entre os meses de maio a julho, período que começa com a queda das folhas. E a frutificação, período subsequente, estende-se ao longo do ano (LIMA, 1980). Dentre as aplicações referentes à espécie *Vataireopsis araroba*, estudos realizados com compostos presentes em suas cascas mostraram sua eficácia contra a psoríase (KERKHOF *et al.*, 2006).

Figura 4 – Morfologia representativa da espécie *Vataireopsis araroba* (Aguiar) Ducke; (A) Caule; (B) folhas fasciculadas terminais; (C) flor bilateralmente simétrica com pétalas onduladas; (D) sâmaras pardo-acastanhadas com aproximadamente 8,5-10,5 cm de comprimento.



Fonte: CARDOSO *et al.*, 2013.

O primeiro relato na literatura com a descrição da purificação e caracterização da primeira lectina isolada das sementes de *Vataireopsis araroba* (Aguiar) Ducke data de 2022 (OSTERNE *et al.*, 2022). Segundo o trabalho desenvolvido, a lectina de *Vataireopsis araroba*, denominada VaL, foi purificada em uma única etapa de cromatografia de afinidade utilizando matriz de goma guar e apresentou especificidade para *N*-acetilgalactosamina, D-

galactose e açúcares relacionados. A respeito da atividade hemaglutinante, a lectina VaL mostrou capacidade de aglutinar diferentes tipos de células sanguíneas humanas (A, B e O) e de coelho, tanto nativos quanto tratados com papaína e tripsina. No que se refere ao perfil de glicosilação, a lectina apresentou um teor de açúcar de 0,6% por miligrama de proteína (OSTERNE *et al.*, 2022).

Em se tratando das demais características e propriedades físico-químicas, a lectina mostrou-se estável em temperaturas de até 60 °C e em uma faixa de pH de 7 a 8. Além disso, foi observada a dependência a cátions divalentes, Ca^{2+} e Mn^{2+} , para sua atividade. Estas características são compartilhadas com outras lectinas de *Dalbergiæae* do mesmo grupo, como a lectina de amendoim da espécie *Arachis hypogaea*, PNA (LAW *et al.*, 1988), a VML e a VGL (CAVADA *et al.*, 1998; SILVA *et al.*, 2012). E por fim, com relação ao seu perfil eletroforético, a lectina VaL apresentou uma única banda em gel de poliacrilamida (SDS-PAGE), com peso molecular aparente de aproximadamente 35 kDa. Já em condições nativas, por cromatografia de exclusão molecular por tamanho, a lectina apresentou uma massa molecular de 79 kDa, indicando que a lectina é um homodímero (OSTERNE *et al.*, 2022).

1.7 Ferramentas de bioinformática utilizadas no estudo das lectinas

As funções desempenhadas pelas lectinas nos organismos estão diretamente relacionadas à sua estrutura nativa e à sua dinâmica de interação com diferentes ligantes, como os carboidratos, por exemplo. Atualmente, a utilização de ferramentas de bioinformática permite o estudo estrutural da lectina (como: alinhamento de sequências, modelagem molecular), a sua afinidade por um determinado ligante (*docking* molecular) e como a interação dessa lectina com o ligante se comporta ao longo de uma faixa de tempo (dinâmica molecular) (PATEL *et al.*, 2019).

Todas as estruturas tridimensionais de proteínas são modelos, com níveis variáveis entre dados experimentais e predição computacional (SCHWEDE, 2013). Geralmente, para se obter as coordenadas atômicas com uma certa precisão são necessárias técnicas experimentais, como a cristalografia de Raios-X, Ressonância Magnética Nuclear (RMN) e Crio-Microscopia Eletrônica (cryo-EM) (CARRONI; SAIBI, 2016; Klebe, 2013). Os dados oriundos dessas técnicas dependem em sua maioria de ferramentas computacionais para a interpretação espacial dos dados, construção e refinamento dos modelos moleculares (SCHWEDE, 2013). Dentre as diversas ferramentas de bioinformática que podem ser utilizadas na caracterização estrutural das proteínas, as técnicas de modelagem tridimensional

e *docking* molecular foram uma das aplicadas no presente estudo estrutural da lectina de *Vataireopsis araroba*.

1.7.1 Modelagem tridimensional

Os modelos tridimensionais das proteínas podem ser construídos por métodos de modelagem comparativa (por homologia), de predição de enovelamento (ou *threading*) e predição por primeiros princípios (ou *ab initio*). No método de modelagem comparativa, utilizado no desenvolvido do estudo estrutural da lectina de *Vataireopsis araroba*, a proteína de interesse (alvo) teve sua estrutura tridimensional predita usando como referência a estrutura tridimensional de outra proteína similar (também chamada de molde, e na maioria das vezes evolutivamente relacionada) (GINALSKI, 2006).

A modelagem comparativa é o método empregado mais frequentemente, e seu limite de predição está intrinsecamente relacionado com o grau de similaridade entre as estruturas alvo e molde. Geralmente, consideram-se como limites mínimos de aplicabilidade do método valores de 25 a 30% de identidade, obtidos através do alinhamento entre a estrutura primária da proteína alvo e de uma ou mais proteínas molde. A modelagem comparativa pode ser dividida em cinco etapas descritas resumidamente na Figura 6 (VERLI, 2014).

Figura 6 – Etapas de predição de estrutura tridimensional de proteínas usando o método de Modelagem Comparativa.



Fonte: VERLI, 2014.

Algumas lectinas de *Dalbergiæ* tiveram sua estrutura tridimensional resolvida por modelagem por homologia. A lectina de *Vatairea guianensis* (VGL) teve sua estrutura tridimensional resolvida utilizando o programa MODELLER 9.16. A análise da sequência da VGL pelo programa BLASTp demonstrou 93% de identidade com a sequência da VML (4U36), e por isso a VML foi utilizada como modelo para a modelagem por homologia da VGL (MARQUES *et al.*, 2017). A lectina de *Centrolobium microchaete* (CML) teve sua estrutura tridimensional resolvida por modelagem por homologia, utilizando a estrutura da CTL (PDB id: 5EYY) como modelo. A estrutura da CTL tem 96% de identidade com a da

CML com base nas pontuações obtidas no BLASTp (NECO *et al.*, 2018). As lectinas do gênero *Arachis*, *Arachis duranensis* (ADL) e *Arachis ipaensis* (AIL), também tiveram seus modelos tridimensionais resolvidos pelo *software* MODELLER v 9.17, usando a lectina de *Centrolobium tomentosum* (CTL, PDB id: 5EYY) como modelo para a modelagem por homologia (NASCIMENTO *et al.*, 2018).

1.7.2 Docking molecular

Uma segunda ferramenta computacional que têm recebido grande destaque em pesquisas associadas à saúde e a biotecnologia, como na descoberta de fármacos, é o *docking* molecular ou atracamento molecular. O *docking* molecular pode ser definido como uma ferramenta computacional que realiza a predição da conformação e orientação de uma molécula ligante dentro do sítio de ligação de uma proteína alvo (YURIEV; HOLIEN; RAMSLAND, 2015).

Os algoritmos do *docking* e as funções de pontuação são capazes de gerar estruturas complexas de receptor-ligante (YURIEV; HOLIEN; RAMSLAND, 2015), possibilitando assim a predição do modo de ligação entre estes, além de também estimar a afinidade entre um ligante e um receptor alvo (ALMEIDA, 2016; VERLI, 2014). O grau de afinidade e de especificidade do ligante pela proteína são definidos pelas características físico-químicas e interações intermoleculares existentes entre eles, que incluem as ligações de hidrogênio, as interações provenientes do efeito hidrofóbico, interações de Van der Waals, interações eletrostáticas e ligações covalentes que podem ser formadas durante o processo de interação proteína-ligante (VERLI, 2014).

Em simulações de *docking* rígido, o algoritmo de pesquisa explora diferentes posições para o ligante no local ativo do receptor, utilizando os graus de liberdade de translação e de rotação (VERLI, 2014). Durante a execução do *docking*, os receptores e ligantes são encaixados e os complexos formados são analisados por meio de cálculos estatísticos e funções de pontuação, sendo as melhores pontuações entre os complexos que mais se aproximam das condições experimentais. O objetivo da função de pontuação é delinear as posições corretas das incorretas em um tempo de computação razoável (MENG *et al.*, 2011). O alto desempenho dos algoritmos possibilita o ancoramento de milhares de ligantes de maneira simultânea em uma simulação por *docking*, tornando-se uma ferramenta extremamente útil para a indústria farmacêutica, por permitir um *screening* de moléculas-alvo

de maneira efetiva e com baixo custo (MENG *et al.*, 2011; PAGADALA; SYED; TUSZYNSKI, 2017).

Alguns trabalhos envolvendo as lectinas de *Dalbergieae* e *docking* molecular são encontrados na literatura, como o estudo realizado por Neco e colaboradores (2018) onde foi analisado as interações da lectina de *Centrolobium microchaete* (CML) com monossacarídeos, dissacarídeos e *N*-glicanos, por *docking* molecular. Os resultados obtidos no estudo demonstraram que a CML provavelmente se liga a α -metil-D-manosídeo, D-manose, *N*-acetil-D-glucosamina, D-glicose e sacarose, corroborando com os resultados obtidos nos ensaios de inibição de hemaglutinação (NECO *et al.*, 2018). Outra lectina em que foi aplicado o *docking* molecular para verificar a capacidade de interação com diversos açúcares foi a VGL, lectina de *Vatairea guianensis*. O *docking* proteína-ligante revelou interações favoráveis com *N*-acetil-D-galactosamina, D-galactose e açúcares relacionados, bem como vários *N*- e *O*-glicanos biologicamente relevantes (MARQUES *et al.*, 2017). Com isso, de maneira geral, o *docking* molecular pode ser utilizado como uma ferramenta eficiente na compreensão da afinidade entre lectinas e diferentes açúcares, bem como na exploração de novos ligantes que possam ser reconhecidos por estas proteínas.

1.8 Importância biotecnologia das lectinas de *Dalbergieae*

As lectinas vegetais estudadas, da tribo *Dalbergieae*, têm sido testadas para diversas atividades biológicas, sendo aplicável em diferentes áreas da biotecnologia e da medicina, como atividade pró e anti-inflamatória (ARARIPE *et al.*, 2017; ALENCAR *et al.*, 2003; MARQUES *et al.*, 2021), nociceptivas (FREITAS PIRES *et al.*, 2019), propriedades anticancerígenas (AVICHEZER; ARNON, 1996; LI *et al.*, 2015), vasorelaxante, antibacteriana, antiviral, dentre outras (NASCIMENTO *et al.*, 2020). Além das atividades biológicas citadas acima, algumas lectinas de *Dalbergieae* desempenham um efeito inseticida interessante, provavelmente devido a interações das lectinas com glicoconjugados nas células de insetos, afetando assim processos biológicos essenciais do organismo (REYES-MONTAÑO; VEGA-CASTRO, 2018), como foi observado nos ensaios biológicos da lectina de *Vataireopsis araroba* usando três linhagens celulares de insetos: células CF1 do intestino médio de lagarta da espécie *Choristoneura fumiferana*, células S2 do embrião de mosca da espécie *Drosophila melanogaster* e células GutAW do intestino médio de lagarta da espécie *Helicoverpa zea*. A lectina VaL apresentou efeito biostático reduzindo o crescimento celular e induzindo aglutinação celular em células CF1, devido a sua interação com glicanos na

superfície celular; já nas demais células, S2 e GutAW, não foi observado efeito biostático (OSTERNE *et al.*, 2022).

1.8.1 Atividade anticâncer

Há na literatura estudos relatando a aplicação das lectinas de *Dalbergieae* para o câncer, tanto para o diagnóstico do câncer, pela capacidade de algumas lectinas desse grupo de interagirem com marcadores específicos de células cancerígenas, como para o efeito citotóxico direto. Como um exemplo de lectina muito estudada, que possui diversas atividades biológicas relatadas na literatura, podemos citar a lectina PNA (*Peanut agglutinin*), a primeira lectina isolada do amendoim da espécie *Arachis hypogaea*. Um dos trabalhos desenvolvidos demonstra o potencial da PNA na identificação de carboidratos específicos em cortes de tecidos e células, como foi observado nas células de glóbulos vermelhos (RYDER *et al.*, 1992), intestino grosso humano (JUAN *et al.*, 2017), sistema linfóide humano e célula de mama humana. Estes tecidos/células que se ligam a PNA e possibilitam identificação de carboidratos específicos (AVICHEZER, *et al.*, 1996). Outros trabalhos demonstram a aplicação da lectina PNA como biomarcador de câncer, principalmente como sensor para detecção do antígeno Tn (SILVA, 2018).

O antígeno Tn (GalNAc-O-Ser), é uma das estruturas associadas a tumores humanos mais específicas (HAKOMORI, 1989; ZHAO *et al.*, 2008). Presente em *O*-glicanos do tipo mucina, o antígeno Tn tem origem epitelial e, embora esteja presente em tecidos saudáveis e doentes benignos, foi observado em aproximadamente 90% dos carcinomas, incluindo câncer de mama, próstata, pulmão e pâncreas, onde foi demonstrada uma correlação direta entre a agressividade do carcinoma e a densidade desse antígeno (DESAI, 2000; SPRINGER, 1997). O antígeno Tn é um biomarcador útil, bem como um potencial alvo terapêutico contra o câncer, tendo em vista sua capacidade de induzir a resposta das células T (CIPOLLA *et al.*, 2002). Além da lectina PNA, foi verificado no estudo desenvolvido por Sousa e colaboradores (2015 e 2016) a interação da VML com o antígeno Tn, através da ligação de resíduos com a *N*-acetilgalactosamina, caracterizando assim outro importante marcador de câncer (SOUSA *et al.*, 2015; SOUSA *et al.*, 2016).

Desse modo, com a purificação e caracterização de novas lectinas da tribo *Dalbergieae*, espera-se a realização de mais estudos acerca de suas atividades biológicas, devido ao seu alto potencial de aplicabilidade nas diversas áreas da biotecnologia e biomedicina.

2 OBJETIVOS

2.1 Objetivo Geral

Determinar a estrutura primária de uma lectina isolada em sementes de *Vataireopsis araroba* (VaL), bem como determinar sua estrutura tridimensional e *docking* molecular com diferentes carboidratos.

2.2 Objetivos Específicos

- Purificar a lectina de sementes de *Vataireopsis araroba* por cromatografia de afinidade;
- Avaliar o grau de pureza da lectina por Cromatografia de Exclusão Molecular acoplada a um sistema de ultra-performance (UPLC);
- Determinar a massa molecular aparente e o grau de pureza da lectina de *Vataireopsis araroba* por eletroforese em gel de poliacrilamida, SDS-PAGE;
- Gerar peptídeos da lectina VaL com diferentes enzimas proteolíticas para o sequenciamento;
- Realizar o sequenciamento de seus respectivos peptídeos por espectrometria de massa em tandem (MS/MS);
- Obter a sequência de peptídeos da lectina VaL utilizando o *software* ProteinLynx (Waters) e realizar alinhamento múltiplo com ClustalOmega e ESPript3;
- Realizar modelagem tridimensional da lectina de *Vataireopsis araroba* por bioinformática;
- Analisar simulações da lectina com diferentes carboidratos por *docking* molecular;
- Avaliar sua toxicidade contra náuplios de *Artemia salina*.

3 MATERIAIS E MÉTODOS

3.1 Material vegetal

As sementes maduras de *Vataireopsis araroba* foram obtidas pelo Laboratório de Moléculas Biologicamente Ativas (BioMol-Lab) da Universidade Federal do Ceará (UFC), Campus do Pici, e devidamente identificadas no herbário Prisco Bezerra (EAC) no departamento de Biologia da UFC.

3.2 Extração e purificação da lectina de sementes de *Vataireopsis araroba* (VaL)

Todo o processo de extração e purificação da lectina foi realizado seguindo a metodologia descrita por Osterne *et al.* (2022). Inicialmente, as sementes de *Vataireopsis araroba* foram descascadas e trituradas em um moedor elétrico até a obtenção de um pó fino. O processo de extração de proteínas foi realizado submetendo a farinha das sementes com NaCl 150 mM (1:10 g/mL) sob agitação constante por 4 h à temperatura ambiente. Em seguida, o extrato bruto foi centrifugado a 10.414 x g por 30 min a 4 °C. O sobrenadante foi filtrado com papel filtro e em seguida submetido a cromatografia de afinidade em matriz de goma guar reticulada, conforme relatado (SILVA *et al.*, 2012). A matriz foi pré-equilibrada com NaCl 150 mM, CaCl₂ 5 mM e MnCl₂ 5 mM. O extrato foi aplicado e mantido em contato com a matriz por 4 h. Posteriormente, todas as substâncias que não se ligaram à matriz foram eluídas com a solução de equilíbrio NaCl 150 mM, CaCl₂ 5 mM e MnCl₂ 5 mM. Em seguida, a lectina de interesse, que ficou retida na matriz, foi eluída com a solução de glicina-HCl 100 mM pH 2,6, contendo NaCl 150 mM. Para este protocolo, foi mantido um fluxo constante de 1 mL por minuto durante toda a cromatografia, usando uma bomba peristáltica. E após todo esse processo de eluição, a fração contendo a lectina de interesse foi dialisada exaustivamente contra água destilada e liofilizada. Todas as frações foram monitoradas por espectrofotometria a 280 nm. A atividade da lectina foi detectada por meio de ensaios de atividade hemaglutinante, utilizando eritrócitos de coelho e sua pureza foi monitorada por cromatografia de exclusão molecular e eletroforese em gel de poliacrilamida (SDS-PAGE) (LAEMMLI, 1970).

3.3 Determinação da pureza da VaL por cromatografia de exclusão molecular

A pureza da amostra foi avaliada por cromatografia de exclusão molecular acoplada a um sistema de cromatografia líquida de ultra-performance (UPLC) em modo isocrático. A coluna utilizada foi do tipo BioSuite™ 250,5 µm HR SEC 7,8 x 300 mm. Inicialmente, ela foi pré-equilibrada com Tris-HCl 50 mM pH 7,6 contendo NaCl 150 mM. A lectina foi solubilizada na mesma solução de Tris-HCl 50 mM pH 7,6 contendo NaCl 150 mM a uma concentração final de 2 mg/mL e em seguida filtrada em filtro de 0,45µm. Após a filtração, a solução lectínica foi aplicada na coluna e eluída com a mesma solução de equilíbrio da coluna. As frações foram coletadas em um volume de aproximadamente 1 mL, sob um fluxo constante de 0,6 mL/min. O processo cromatográfico foi realizado com base nos parâmetros utilizados no estudo desenvolvido por Osterne *et al.* (2022).

3.4 Eletroforese em condições desnaturantes

A pureza da lectina isolada foi monitorada por SDS-PAGE, seguindo a metodologia desenvolvida por Laemmli (1970). Para isto, foi preparado um gel de separação usando acrilamida/bisacrilamida 12,5% em tampão Tris-HCl 0,1 M pH 8,8 contendo SDS, TEMED (concentrado) e persulfato de amônio (APS) 100 mg/mL. O gel de empilhamento (*stacking gel*) foi preparado usando acrilamida/bisacrilamida 4% em tampão Tris-HCl 0,1 M pH 6,8 contendo SDS, persulfato de amônio 100 mg/mL e TEMED (concentrado). A amostra foi solubilizada a uma concentração de 4 mg/mL utilizando tampão de amostra contendo Tris-HCl 0,0625 M pH 6,8, glicerol 12,5% e SDS 2%. Em seguida, foi realizado o aquecimento da amostra a 100 °C por 10 minutos. Logo depois, a amostra foi resfriada à temperatura ambiente (aproximadamente 25 °C) e 15 µL desta preparação foi aplicada em cada poço do gel de empilhamento.

A corrida eletroforética foi realizada em um sistema Mini-protean II mini-gel (Bio-Rad; Milão, Itália) com uma voltagem variando em até 150 V, potência de até 10W, e amperagem constante de 30 mA. O tampão de corrida foi preparado utilizando Tris 0,025M pH 8,8; Glicina 0,192M e SDS 0,1%. Após a corrida eletroforética, o gel de separação foi corado por 24 horas em *Coomassie Brilliant Blue G-250* a 0,05% dissolvido em metanol, ácido acético e água a uma proporção 1:3,5:8 (v/v/v). A retirada do excesso do corante foi feita utilizando água destilada. E por fim, a massa aparente da lectina foi determinada em comparação com as massas moleculares de proteínas conhecidas presentes no marcador

molecular TrueColor High Range Protein Marker (SINAPSE INC) com bandas de proteína variando de 10 a 245 kDa.

3.5 Digestão proteica e sequenciamento dos peptídeos por espectrometria de massa em tandem (MS/MS)

A estrutura primária da VaL foi obtida seguindo a metodologia de Shevchenko *et al.* (2006). Primeiramente as bandas em torno de 30 kDa, correspondentes à lectina isolada em gel SDS-PAGE, foram excisadas e descoradas em bicarbonato de amônio 100 mM e acetonitrila (1:1 v/v). Posteriormente, as bandas foram desidratadas em acetonitrila 100% e secas em um Speedvac (LabConco Kansas City, MO, EUA). Soluções de endoproteínase Asp-N, quimotripsina, tripsina e Glu-C (Sigma-Aldrich) em bicarbonato de amônio 50 mM foram utilizadas, separadamente, para a clivagem proteolítica na proporção especificada no manual do fabricante. As enzimas permaneceram incubadas com a amostra por 12 h a 37 °C para a clivagem proteolítica. Após a incubação, a reação foi interrompida com a adição de ácido fórmico a 5% e acetonitrila 100% (1:2 v/v). As soluções contendo os peptídeos digeridos foram transferidas para novos tubos e secas em concentrador à vácuo (SpeedVac).

Os peptídeos secos, resultantes da digestão, foram então solubilizados em ácido trifluoroacético (TFA) a 0,1% e separados em uma Coluna C18 de fase reversa (100 µm × 100 mm), utilizando um sistema nanoAcquity™ (Waters Corp). A eluição ocorreu em um gradiente linear de acetonitrila (10-85%) contendo ácido fórmico a 0.1% a 600 µL/min. Os peptídeos separados foram transferidos a um *nanoelectrospray* acoplado ao espectrômetro de massa. O espectrômetro de massa foi operado em modo positivo, com uma temperatura de fonte de 80 °C e voltagem de 3,5 kV no capilar. A calibração foi realizada com fragmentos de íon duplo protonado [Glu1] - fibrinopeptídeo B (Sigma-Aldrich) (m/z 785,84). Os íons precursores foram fragmentados usando o método de Decomposição Induzida por Colisão (Collision Induced Decomposition, CID). A fragmentação ocorreu usando argônio como gás de colisão, e a energia de colisão em rampa variou de acordo com o estado de carga do íon precursor selecionado.

Os dados obtidos foram processados e analisados usando ProteinLynx (Waters), e o parâmetro de busca foi o padrão de fragmentação peptídica. Os espectros CID foram interpretados manualmente usando a ferramenta Peptide Sequencing do software MassLynx 4.1 (Waters). A sequência dos peptídeos obtidos foram submetidos ao BLASTp, e utilizada para comparação com sequências de proteínas não redundantes do banco de dados

(ALTSCHUL *et al.*, 1990). O alinhamento múltiplo com outras lectinas de espécies similares foi realizado utilizando os programas ClustalOmega (SIEVERS *et al.*, 2011) e ESPript 3.0 (ROBERT; GOUET, 2014).

3.6 Análise filogenética e predição da estrutura secundária

A sequência dos peptídeos obtidos da lectina de *Vataireopsis araroba* (VaL) foi comparada com a sequência de outras proteínas, depositadas no banco de dados UniProt, através da ferramenta de busca de sequências BLASTp (NCBI). Duas lectinas e três precursores foram escolhidos a fim de se realizar a análise filogenética da lectina VaL dentro da subfamília *Papilionoideae*, sendo analisadas similaridades entre as sequências dessas lectinas e precursores de outras tribos. Para isto, foram selecionadas as seguintes lectinas e precursores: lectina de *Vatairea macrocarpa* (UniProt ID: P81371.1), lectina de *Vatairea guianensis* (UniProt ID: P86893.1), precursor da lectina *Cladrastis kentukea* (UniProt ID: Q39527.1), precursor da lectina *Styphnolobium japonicum* (UniProt ID: P93535.1) e precursor da lectina *Robinia pseudoacacia* (UniProt ID: Q41161.1).

Além disso, foi realizado um segundo alinhamento de sequências de aminoácidos de lectinas que possuem sua estrutura tridimensional resolvida e depositada no banco de dados PDB (*Protein Data Bank*), para fins de discussão sobre estruturas terciárias. Para isto, foram selecionadas as seguintes lectinas: lectina de *Vatairea macrocarpa* (PDB ID: 4WV8_A), lectina de *Wisteria floribunda* (PDB ID: 5KXD_A), lectina de *Robinia Pseudoacacia* (PDB ID: 1FNY_A), lectina de *Vicia villosa* (PDB ID: 1N47_A), e lectina de *Glycine max* (PDB ID: 1G9F_A). Os alinhamentos das sequências de proteínas foram realizados utilizando a ferramenta ClustalOmega (SIEVERS *et al.*, 2011) e a árvore filogenética foi construída utilizando o mesmo programa. Considerando os dados disponíveis nos bancos de dados UniProt e PDB, as lectinas mais intimamente relacionadas (com maior identidade) foram alinhadas para análise da conservação dos resíduos envolvidos na estruturação dos sítios de ligação de metais (MBS) e domínio de reconhecimento a carboidratos (CRD).

A predição da estrutura secundária da VaL foi realizada usando o servidor PSIPRED, um sistema automatizado para predição da estrutura secundária (BUCHAN *et al.*, 2013). A estrutura secundária foi aplicada como um dos fatores de validação para a seleção do modelo de homologia.

3.7 Modelagem tridimensional

A predição da estrutura tridimensional da lectina de *Vataireopsis araroba* (VaL) foi realizada pelo programa MODELLER v10.4 e pelo servidor PHYRE2 (KELLEY *et al.*, 2015), que são ferramentas computacionais que utilizam o método de modelagem comparativa por homologia a partir de um certo grau de identidade de proteínas já resolvidas experimentalmente (WEBB; SALI, 2016). Para a modelagem com os metais (Ca^{2+} e Mn^{2+}) foi utilizado o modelo da lectina de *Vatairea macrocarpa* complexada com o antígeno Tn (PDB ID: 4U36), como referência para inclusão manual dos metais na estrutura.

No caso do MODELLER, inicialmente duzentos modelos foram gerados e classificados com base na função de pontuação objetiva do MODELLER (molpdf) e pontuações de energia de proteína otimizada discreta (z-DOPE), para avaliar a qualidade geral da estrutura em que modelos com menor molpdf e z-DOPE foram selecionados. Os modelos gerados pelo MODELLER juntamente com os modelos gerado pelo servidor PHYRE2, foram submetidos à validação de propriedades estereoquímicas, como gráfico de Ramachandran, sobreposições estéricas, parâmetros de desvio C_{β} , rotâmeros e desvios de ângulo de ligação usando o programa PROCHECK (LASKOWSKI *et al.*, 1993). A correlação entre a estrutura primária e terciária foi avaliada pelo servidor Verify3D (LÜTHY; BOWIE; EISENBERG, 1992; BOWIE, LÜTHY; EISENBERG, 1991). Os valores de QMEANDisco também foram avaliados pelas Ferramentas de Avaliação de Estrutura e Modelo de Proteínas incluídas no Swiss-Model (BENKERT; BIASINI; SCHWEDE, 2011; BENKERT; KÜNZLI; SCHWEDE, 2009). A qualidade geral do modelo foi avaliada pelo servidor ERRAT (COLOVOS; YEATES, 1993). O modelo com os melhores valores em todas as validações foi selecionado para representar a estrutura tridimensional da VaL e aplicado nas análises subsequentes de *docking* molecular. As imagens das estruturas moleculares foram produzidas pelo programa PyMOL v.1.7.4.5. (Shrodinger, LLC).

3.8 Docking molecular

De posse da estrutura tridimensional da VaL, resolvida e validada, e considerando também os resultados experimentais obtidos no teste de especificidade de ligação a carboidratos, no estudo desenvolvido por Osterne *et al.* (2022), foram realizados experimentos de *docking* com os seguintes carboidratos: D-galactose, N-acetilgalactosamina (GalNAc), α -lactose e L-ramnose. Para realizar os encaixes, as estruturas dos carboidratos

foram obtidas no PubChem - Centro Nacional de Informações sobre Biotecnologia (The Pubchem Project). Além dos carboidratos, foi também selecionado para este ensaio o antígeno Tn, um importante marcador de câncer, para comparar a interação da VaL e VML com este antígeno.

As simulações da lectina com os carboidratos foram realizadas com o programa GOLD v2022.2, que é um software que utiliza um modo de precisão padrão para determinar as possíveis ligações favoráveis e detectar várias conformações de ligantes flexíveis, mantendo a proteína como uma estrutura rígida. A localização do CRD da VaL foi obtida por superposição com a VML usando, como parâmetros de *docking*, o raio de ligação definido para 6 Å ao redor do CRD. O algoritmo PLANTSPLP foi aplicado para calcular os escores de *docking* (KORB; STÜTZLE; EXNER, 2009), onde valores mais negativos indicam interações mais favoráveis, e as melhores conformações para cada ligante foram selecionadas com base nas interações de hidrogênio, interações hidrofóbicas, interações de Van der Waals, considerando também penalidades de conformação do ligante (valores de torção). Os resultados de *docking* foram validados comparando o perfil de interação da VaL com estes ligantes e os disponíveis em bancos de dados relacionados à VML. O *software* PyMol foi usado para gerar as imagens de interação entre a VaL e os ligantes.

3.9 Avaliação da toxicidade contra náuplios de *Artemia salina*

Os ensaios de toxicidade da lectina VaL contra os náuplios de *Artemia salina* foram realizados com água do mar coletada da Praia do Cumbuco, Ceará, Brasil. Após a coleta da água foi realizado uma filtração comum com papel filtro para remoção de resíduos macroscópicos seguido novamente de filtração em membrana de celulose (0,45 µm, 47 mm) para retenção de partículas microscópicas como microrganismos. Após a filtração, a água foi esterilizada e utilizada no teste de toxicidade. Os cistos de *Artemia salina* foram incubados na água sob iluminação e aeração constante (1 g de cistos por litro de água), sob temperatura de 25 °C. Após 48 h, a aeração foi interrompida e apenas os náuplios de *Artemia* com alta mobilidade foram coletados com pipeta *pasteur* para realização do teste (CAVADA *et al.*, 2019a).

O teste de toxicidade foi realizado em placas de Limbro de 24 poços. A solução da lectina VaL a 200 µg/mL de água do mar foi diluída seriadamente ($\frac{1}{2}$; $\frac{1}{4}$; $\frac{1}{8}$; ...) até a concentração de 3,125 µg/mL. Cada poço da placa continha 2 mL de água do mar. Após a diluição da lectina, foram adicionados 10 náuplios de *Artemia* em cada poço da placa. O

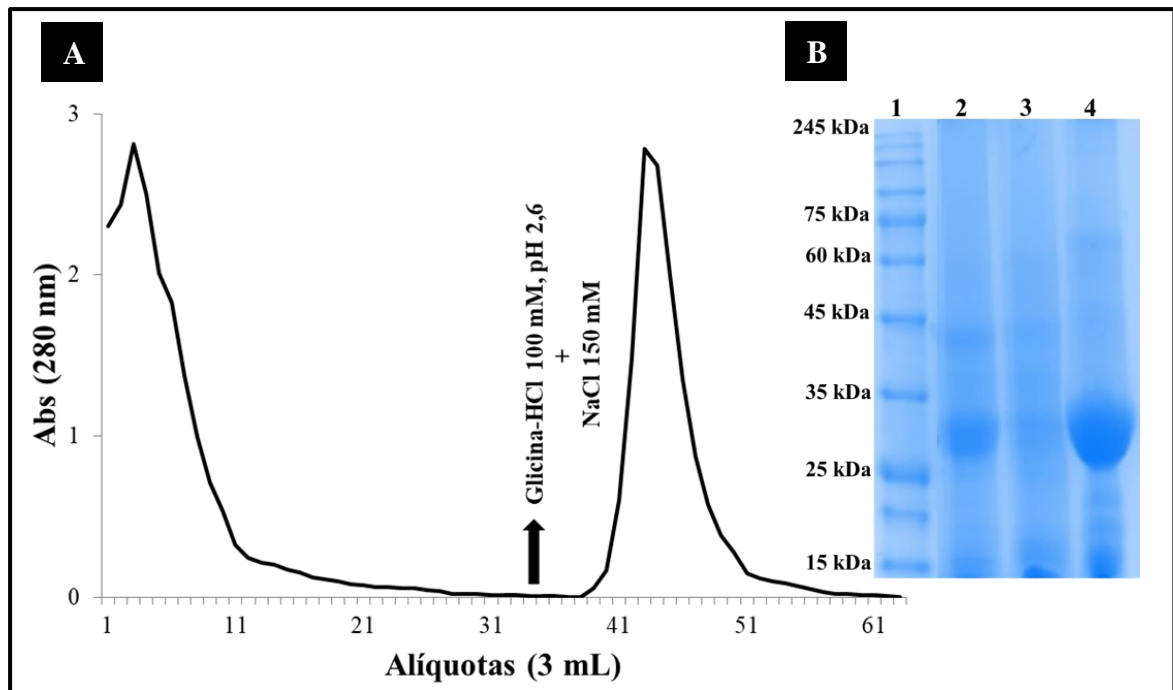
experimento foi realizado em triplicata e os resultados obtidos de cada concentração testada foram comparados com os grupos controles: água do mar, lactose 100 mM + água do mar e lactose 100 mM + VaL. Após 24 e 48 h, foi realizada a contagem dos náuplios mortos e vivos (ARRUDA *et al.*, 2013). Foi utilizado o programa GraphPad Prism (Versão 9.5.0) para plotagem dos dados obtidos mediante a análise de variância (ANOVA).

4 RESULTADOS E DISCUSSÕES

4.1 Purificação da lectina de sementes de *Vataireopsis araroba* (VaL)

A VaL foi purificada por uma única etapa de cromatografia de afinidade em matriz de goma guar reticulada (Figura 7A). Seu perfil cromatográfico foi semelhante à de outras lectinas específicas para *N*-acetilgalactosamina/galactose da tribo *Dalbergieae*, como as lectinas de *Vatairea macrocarpa* (CAVADA *et al.*, 1998) e *Vatairea guianensis* (SILVA *et al.*, 2012). Em seguida, foi feita uma análise de eletroforese (SDS-PAGE), seguindo a metodologia de Laemmli (1970), para observar a pureza da amostra e o peso molecular aparente. A VaL apresentou perfil eletroforético, com massa molecular aparente de aproximadamente 30 kDa (Figura 7B), semelhante ao observado em algumas lectinas específicas para glicose/manose da tribo *Dalbergieae*, com apenas uma única banda, referente à cadeia polipeptídica contendo um único sítio de glicosilação, sem evidência de processamento pós-traducional. Esse perfil eletroforético também foi observado na lectina de *Centrolobium microchaete* (VASCONCELOS *et al.*, 2015).

Figura 7 – Perfil cromatográfico e eletroforético da VaL. (A) O extrato bruto ($V_{amostra} = 6$ mL) foi aplicado em uma matriz de goma de guar (2cm x 15cm) e mantido em contato por 4 horas. A coluna foi previamente equilibrada com NaCl 150 mM, contendo $CaCl_2$ 5 mM e $MnCl_2$ 5 mM. As proteínas não ligantes (fração não retida – P1) foram eluídas com a solução de equilíbrio e a lectina (fração retida – P2) foi eluída com glicina-HCl 100 mM, pH 2.6, contendo NaCl 150 mM. Em cada tubo foram coletadas em alíquotas de 3 mL com fluxo de 1mL/min. As frações foram monitoradas por espectrofotometria a 280 nm. (B) Poço 1: marcador molecular (S2600 TrueColor High Range Protein Marker SINAPSE INC). Poço 2: extrato bruto (30 μ g). Poço 3: fração não retida – P1 (30 μ g). Poço 4: VaL pura (30 μ g).

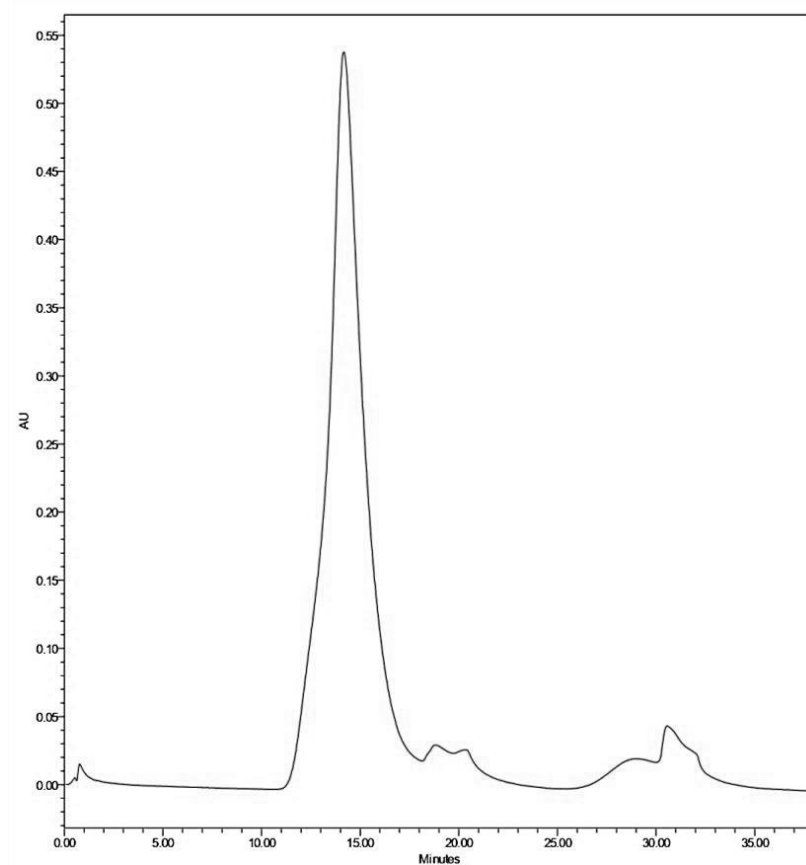


Fonte: Elaborado pela própria autora.

4.2 Cromatografia de exclusão molecular em sistema UPLC

Prosseguindo com a etapa adicional para verificar a pureza da VaL, foi realizado uma análise por cromatografia de exclusão molecular acoplada a um sistema de cromatografia líquida de ultra-performance (UPLC). Com isso, foi observado um perfil semelhante ao obtido no estudo desenvolvido por Osterne *et al.* (2022), sendo constatada a pureza da amostra a partir do cromatograma obtido na análise (Figura 8).

Figura 8 - Cromatograma da VaL, sob condição nativa, em coluna SEC. Foi utilizada uma solução de lectina a 2 mg/mL em NaCl 150 mM. A eluição foi realizada com o tampão Tris-HCl 100 mM, pH 7,6 contendo NaCl 150 mM no modo isocrático. As frações foram coletadas em alíquotas de 1 mL com uma taxa de fluxo de 0,6 mL/min. As frações foram monitoradas por espectrofotometria a 280 nm.



Fonte: Elaborado pela própria autora.

4.3 Estrutura primária parcial da VaL

A lectina de *Vataireopsis araroba* (VaL) foi digerida com as seguintes enzimas proteolíticas: Asp-N, Quimotripsina, Tripsina e Glu-C (Sigma-Aldrich). Os peptídeos obtidos da digestão totalizaram 233 resíduos de aminoácidos, que correspondem a 9 peptídeos obtidos da digestão com a enzima Asp-N, 6 peptídeos obtidos da digestão com a enzima Quimotripsina, 15 peptídeos obtidos da digestão com a enzima Tripsina e 11 peptídeos obtidos da digestão com a enzima Glu-C (Tabela 1).

Em trabalhos anteriores, utilizando espectrômetro de massas do tipo MS/MS, foi obtido uma cobertura de 45,41% da estrutura primária da VaL (OSTERNE *et al.*, 2022). Em novos experimentos de espectrometria de massas, também utilizando um espectrômetro do tipo MS/MS, foram obtidos novos íons e foi possível aumentar a cobertura da sequência completa da VaL para 94,58%, com base nos alinhamentos de sequência com as lectinas de *Vatairea macrocarpa* (UniProt ID: P81371.1) e *Vatairea guianensis* (UniProt ID: P86893.1), que foram usadas como referência, devido essas lectinas apresentarem maior identidade com os peptídeos da VaL. O alinhamento das sequências pode ser visualizado na Figura 9. Os

peptídeos obtidos no estudo serão posteriormente depositados na forma de sequência parcial no banco de dados UNIPROT.

Tabela 1 - Sequência dos peptídeos obtidos por espectrometria de massas em tandem (MS/MS). Os peptídeos identificados com T foram digeridos pela tripsina, os peptídeos identificados com Q foram digeridos por quimotripsina, os peptídeos identificados com G foram digeridos com a endoproteínase-Glu-C e os peptídeos identificados com A foram digeridos pela endoproteínase Asp-N.

PEPTÍDEOS	SEQUÊNCIA	MASSA MOLECULAR (Da)
G1	APPDTQPQKDGGF	1357.68
G2	VKDGKPVHDHSLGRAL	1590.90
G3	KGKLQLTKVKDGG	1313.70
G4	VNSIESMKYVK	1296.62
G5	DVTSKGGKLQL	1088.88
G6	SRDHSVETHDVLV	1568.64
G7	FTKFNPNPK	1091.52
G8	SATSGLSR	892.44
G9	LAPPDTQPQKDGGF	1470.76
G10	FLAPPDTQPQKDGGF	1617.81
G11	FSNTWDP SARHLGLN	1714.85
T1	VAVEFDTF SNTWDP SAR	1186.87
T2	AFFLAPPDTQPQK	1459.81
T3	LPEWVR	798.54
T4	VDHSLGR	783.43
T5	AFFLAPPDTKPKQKDG GFLGLFNDSNK	2824.43
T6	PSNATSYLV SANVDL	1678.98
T7	NPNPKDILQGDALVTSK	1923.08
T8	TANVYLSYQASTK	1444.58
T9	YTAPIHIWDESTDR	1703.84
T10	HIGINVNSIESQK	1438.83
T11	SFATSFSFVVEAPDESK	1848.02
T12	GFSATSGLSR	982.53
T13	RVGFSATSGLSR	1237.69
T14	WAWEDGK	890.65
T15	ALPEWVR	870.52
Q1	SFVVEAPDESKTADGLAFF	2030.08
Q2	VSANVDLKSALPEWVRVGF	2087.15
Q3	TVAVEFDTF	1028.50
Q4	DPSARHIGINVNSIESQKY	2128.14
Q5	HLWDDSTDRVASF	1548.74
Q6	SYEASTK	785.39
A1	SEVVFSFTKFNPNPK	1828.84
A2	AFFLAPPDTQPQK	1459.78
A3 e A4	PVDHSLGRALY	1227.63
A5	DGKPVHDHSLGRALYTAPLHIW	2346.20
A6	DLKSALPEWVRVGF SATSGLSR	2376.18

A7	VLFWSFTSTRR	1399.66
A8	SFTKFNPMPK	1179.59
A9	NKSIQTVAVEF	1234.58

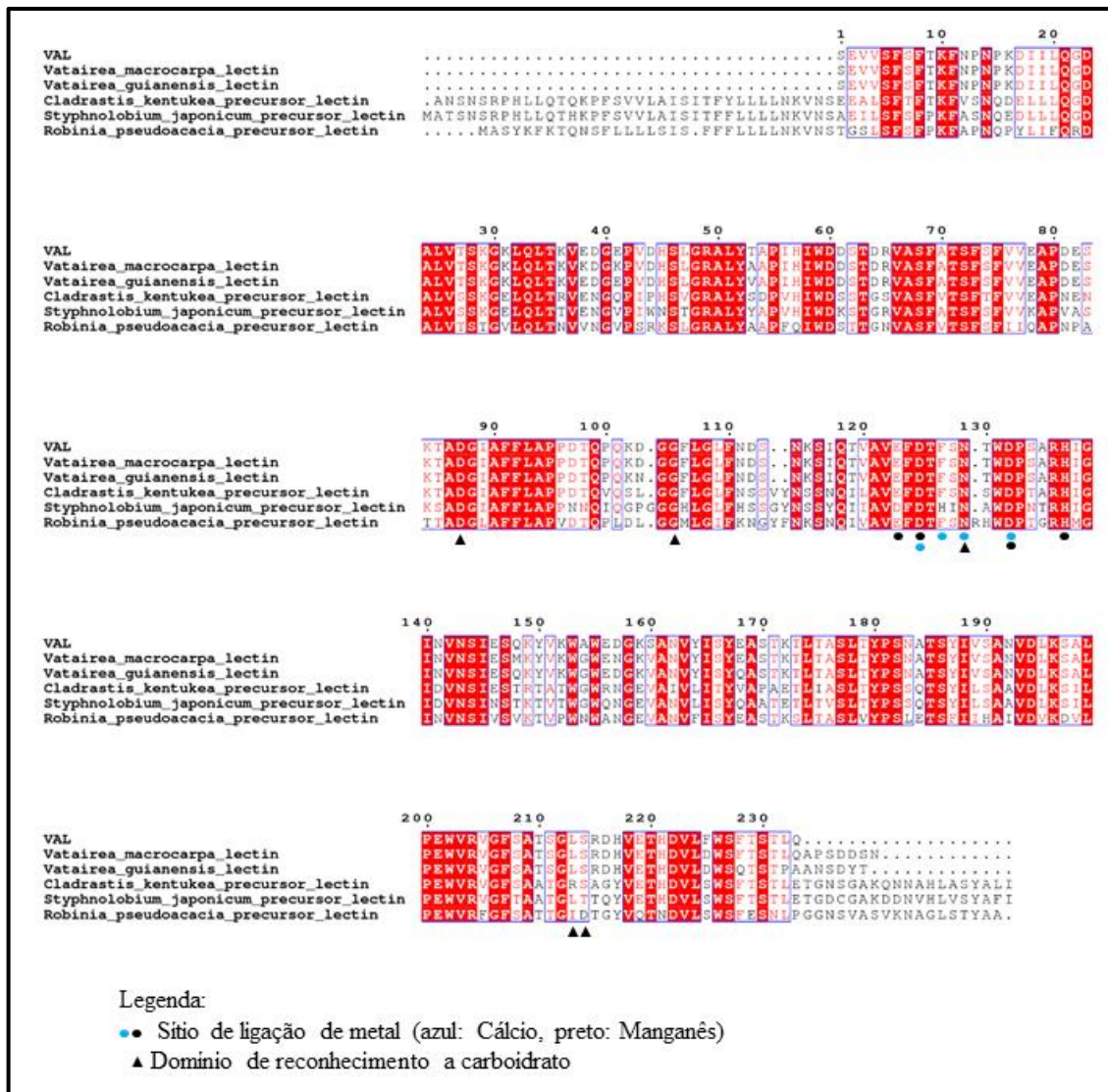
Fonte: Elaborado pela própria autora.

4.4 Alinhamento de sequências e Análise filogenética

De posse da sequência de aminoácidos da lectina de *Vataireopsis araroba* (VaL) foi realizado o alinhamento no banco de dados UniProt (BLASTp). Os resultados do alinhamento indicaram que a VaL é 96,57% idêntica à lectina de *Vatairea macrocarpa* (UniProt ID: P81371.1), 96,97% idêntica à lectina de *Vatairea guianensis* (UniProt ID: P86893.1), 69,23% idêntica ao precursor da lectina *Cladrastis kentukea* (UniProt ID: Q39527.1), 67,37% idêntica ao precursor da lectina *Styphnolobium japonicum* (UniProt ID: P93535.1) e 63,36% idêntica ao precursor da lectina *Robinia pseudoacacia* (UniProt ID: Q41161.1) (Figura 9). O alinhamento mostrou ainda que a lectina *Vatairea macrocarpa* (VML) apresentou um alto grau de identidade e cobertura de sequência (100%) com a lectina de *Vataireopsis araroba* (VaL). Com isso, podemos observar que a VaL apresentou um alto grau de identidade, com cerca de 98,28% de resíduos de aminoácidos conservados, com outras lectinas da tribo *Dalbergieae* específicas para *N*-acetilgalactosamina/galactose (VML e VGL). As demais lectinas desta tribo específicas para manose/glicose não participaram do alinhamento por possuir baixa identidade de sequência com a VaL, como as lectinas de *Platypodium elegans* (PELa) (BENEVIDES *et al.*, 2012), *Pterocarpus rotundifolius* (PRL) e *Pterocarpus angolensis* (PAL) (NASCIMENTO *et al.*, 2020; LORIS *et al.*, 2003).

Semelhante a VML, o domínio de reconhecimento a carboidratos (CRD) e sítios de ligação a metais (MBS) são conservados na VaL. Entre os resíduos conservados estão os aminoácidos envolvidos no sítio de reconhecimento a carboidratos (Asn-129, Gly-105, Asp-87, Leu-213 e Ser-214) e no sítio de ligação a metais (**íon Mn²⁺**: Glu-123, Asp-125, Asp-132, His-137; **íon Ca²⁺**: Asp-125, Phe-127, Asn-129, Asp-132) (CALVETE *et al.*, 1998; SOUSA *et al.*, 2015; SOUSA *et al.*, 2016).

Figura 9 - Alinhamento da sequência parcial da VaL com as sequências obtidas do UniProt de outras lectinas da subfamília *Papilionoideae*. Os aminoácidos destacados em vermelho representam resíduos idênticos. Os resíduos envolvidos com o Domínio de Reconhecimento a Carboidratos (CRD) estão marcados por um triângulo, enquanto resíduos envolvidos com o Sítio de Ligação do Metal (MBS) estão marcados por um círculo preto (Mn²⁺) e azul (Ca²⁺).

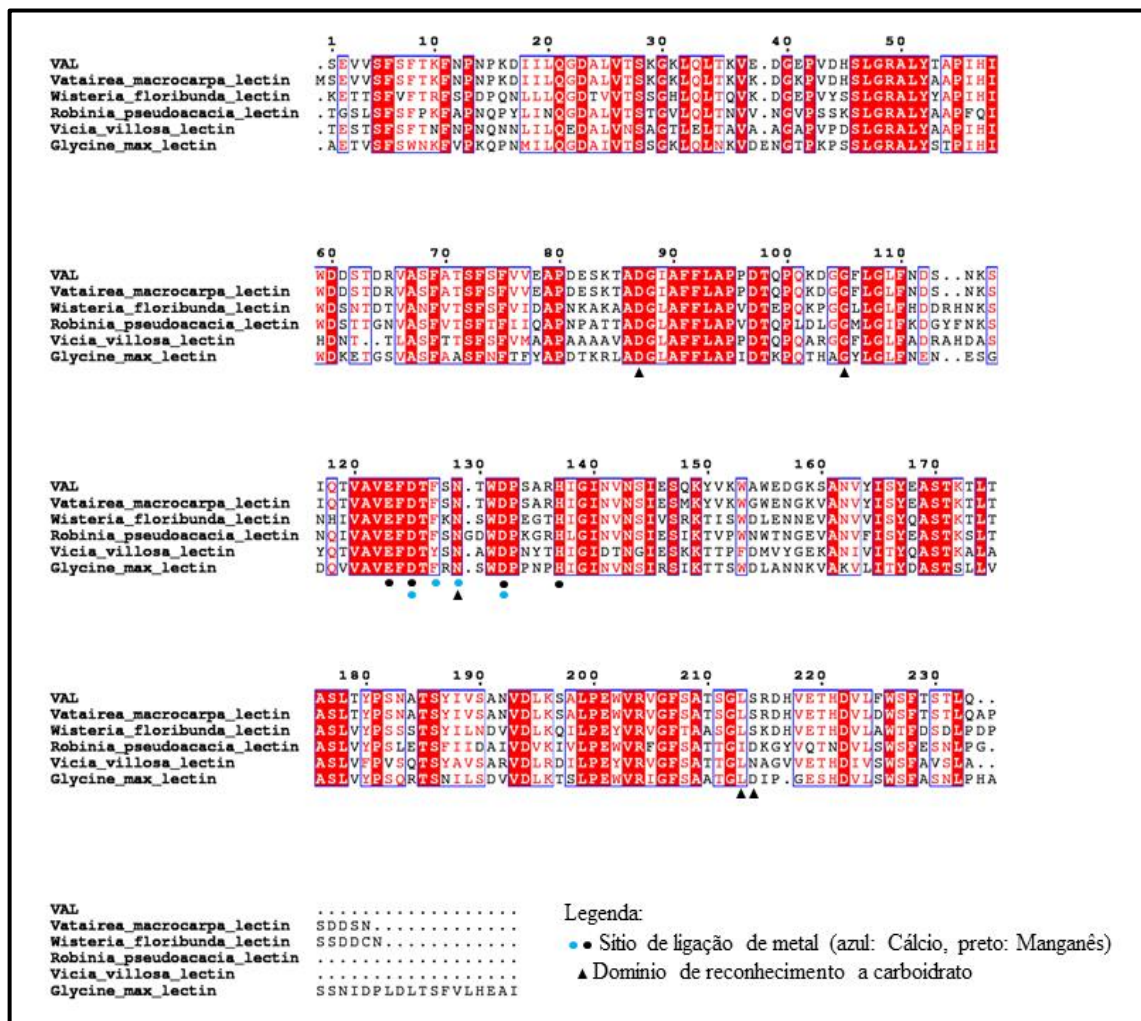


Fonte: Elaborado pela própria autora.

Dando continuidade ao estudo estrutural, foi realizado um segundo alinhamento de seqüências de aminoácidos de lectinas que possuem sua estrutura tridimensional resolvida e depositada no banco de dados PDB (*Protein Data Bank*). Os resultados do novo alinhamento indicaram que a estrutura primária da VaL é 96,57% idêntica à lectina de *Vatairea macrocarpa* (PDB ID: 4WV8_A), 66,95% idêntica à lectina de *Wisteria floribunda* (PDB ID: 5KXD_A), 64,22% idêntica à lectina de *Robinia Pseudoacacia* (PDB ID: 1FNY_A), 61,11% idêntica à lectina de *Vicia villosa* (PDB ID: 1N47_A) e 60,52% idêntica à lectina de *Glycine max* (PDB ID: 1G9F_A) (Figura 10). Corroborando com o resultado do primeiro alinhamento, podemos observar que a VaL apresentou um alto grau de identidade, em relação à estrutura primária, e um alto grau de cobertura (100%) com a lectina de *Vatairea macrocarpa*. Os resíduos de aminoácidos envolvidos no domínio de reconhecimento a

carboidratos (Asn-129, Gly-105, Asp-87, Leu-213 e Ser-214) e nos sítios de ligação a metais (íon Mn^{2+} : Glu-123, Asp-125, Asp-132, His-137; íon Ca^{2+} : Asp-125, Phe-127, Asn-129, Asp-132) (CALVETE *et al.*, 1998; SOUSA *et al.*, 2015; SOUSA *et al.*, 2016) se mostraram conservados e são iguais aos observados no alinhamento anterior. Este segundo alinhamento confirmou a utilização da estrutura tridimensional da VML como modelo de referência nos processos de modelagem molecular.

Figura 10 - Alinhamento da sequência parcial da VaL com sequências de aminoácidos de outras lectinas que possuem sua estrutura tridimensional resolvida e depositada no banco de dados PDB. Os aminoácidos destacados em vermelho representam resíduos semelhantes. Os resíduos envolvidos no Domínio de Reconhecimento a Carboidratos (CRD) estão marcados por um triângulo, enquanto resíduos envolvidos com o Sítio de Ligação do Metal (MBS) estão marcados por um círculo preto (Mn^{2+}) e azul (Ca^{2+}).

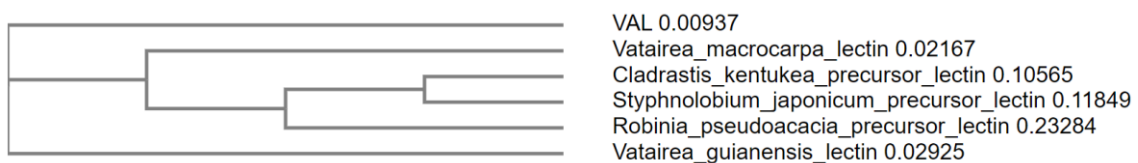


Fonte: Elaborado pela própria autora.

Na árvore filogenética construída a partir da sequência de aminoácidos de lectinas da subfamília *Papilionoideae* (Figura 11), é possível notar que a VaL é uma das lectinas mais primitivas da tribo *Dalbergieae* juntamente com a lectina de *Vatairea guianensis*, não compartilhando o mesmo clado com as demais lectinas na escala evolutiva, como o precursor da lectina *Cladrastis kentukea*, precursor da lectina *Styphnolobium japonicum*, e o precursor da lectina *Robinia pseudoacacia* (tribo *Robinieae*). Com isso, essas espécies pertencentes a tribos, como *Robinieae*, são mais recentes na escala evolutiva em comparação com a tribo *Dalbergieae*.

Já na árvore filogenética construída a partir de sequências de aminoácidos de lectinas que possuem sua estrutura tridimensional resolvida, é possível observar a proximidade evolutiva já esperada entre as lectinas de *Vataireopsis araroba* e *Vatairea macrocarpa*, devido à alta similaridade de sequência e de estruturas terciárias entre essas duas lectinas (Figura 12). Desse modo, é importante considerar que além de estudos estruturais e de aplicações referentes às lectinas, também se deve buscar compreender a origem evolutiva dessas proteínas, auxiliando no entendimento sobre qual papel elas podem ter desempenhado na adaptação das plantas ao ambiente terrestre e como se diferenciam entre si desde então (SUVARNA; SHARMA, 2020; VAN HOLLE; VAN DAMME, 2019).

Figura 11 - Árvore filogenética construída a partir de sequências de aminoácidos, obtidos do UniProt, de lectinas da subfamília *Papilionoideae*, demonstrando a distância evolucionária da VaL em comparação com outras lectinas da subfamília.



Fonte: Elaborado pela própria autora.

Figura 12 - Árvore filogenética construída a partir de sequências de aminoácidos de lectinas que possuem sua estrutura tridimensional resolvida e depositada no banco de dados PDB, demonstrando a distância evolucionária da VaL em comparação com outras lectinas.



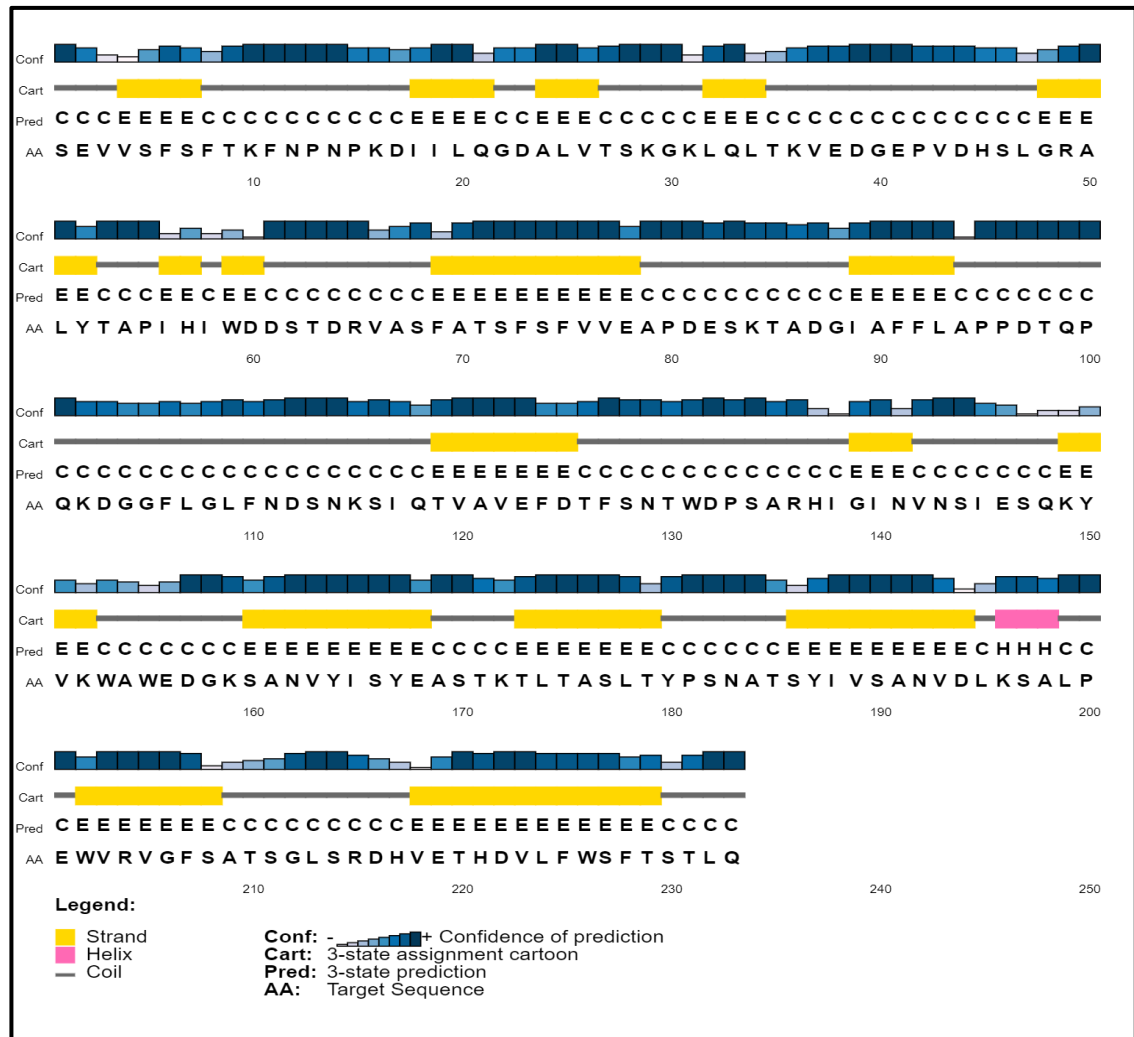
Fonte: Elaborado pela própria autora.

Com base nas análises das duas árvores filogenéticas, é possível notar a diferença de disponibilidade de dados de estruturas primária e terciária de lectinas, principalmente um menor volume de dados de estruturas tridimensionais referentes às lectinas de tribos mais primitivas, como *Dalbergieae*. Reiterando a importância da realização de estudos estruturais das lectinas desta tribo.

4.5 Predição da estrutura secundária

A predição da estrutura secundária usando o *software* PSIPRED (BUCHAN et al., 2013) indicou que a VaL segue o padrão observado nas lectinas de *Dalbergieae* com grande predominância de folhas beta em sua estrutura secundária, apresentando em sua composição 41,2% de folhas- β , 1,28% de α -hélice e 54,93% de *loops* (Figura 13). Este resultado corrobora com os dados que foram relatados no tópico 4.6, onde é possível ver uma prevalência de folhas- β na estrutura geral do monômero VaL.

Figura 13 – Predição da estrutura secundária da lectina de *Vataireopsis araroba* (VaL). As fitas amarelas representam os resíduos de aminoácidos envolvidos na estrutura das folhas- β , a fita rosa representa os resíduos de aminoácidos envolvidos na estrutura da conformação α -hélice e os tracejados cinza representam os resíduos de aminoácidos envolvidos na estrutura de *loops*.



Fonte: Elaborado pela própria autora.

A glicosilação é uma das modificações pós-traducionais mais comuns, promovida por um sistema altamente regulado de enzimas e transportadores que adicionam covalentemente cadeias de açúcares (glicanos) a proteínas, produzindo um grupo diversificado de glicoproteínas (CALVETE *et al.*, 1998). No estudo desenvolvido por Osterne *et al.* (2022) foi observado que a lectina de *Vataireopsis araroba* (VaL) apresentou um teor de açúcar de 0,6% por miligrama de proteína. Com isso, para a predição dos sítios de glicosilação da VaL, foi usada sua sequência no programa NetNglyc, sendo observado a posição de sítios de *N*-glicosilação nos resíduos de asparagina nas posições 111, 114, 183. Várias lectinas de *Dalbergiae* se apresentaram como glicoproteínas em estudos de caracterização, como a VGL (SILVA *et al.*, 2012) e a VML (CAVADA *et al.*, 1998) que possuem estruturas similares e composição de carboidratos semelhantes, com *N*-glicanos nos resíduos de asparagina nas posições 111 e 183. Outras lectinas, como a CML, contém aproximadamente 1,5% de

carboidrato (VASCONCELOS *et al.*, 2015) semelhantemente a CTL que apresentou aproximadamente 1,7% de carboidrato em sua composição (ALMEIDA *et al.*, 2014a).

4.6 Modelagem tridimensional

A predição da estrutura tridimensional da lectina de *Vataireopsis araroba* (VaL) foi realizada pelo programa MODELLER v9.16 e pelo servidor PHYRE2 (KELLEY *et al.*, 2015), utilizando o método de modelagem comparativa por homologia. A estrutura da lectina de *Vatairea macrocarpa* complexada com o antígeno Tn (PDB ID: 4U36) foi utilizada como modelo, devido ela apresentar um alto grau de identidade, em relação à estrutura primária, e um alto grau de cobertura (100%) com a VaL.

Os modelos gerados pelo MODELLER juntamente com os modelos gerado pelo servidor PHYRE2, foram submetidos à validação de propriedades estereoquímicas, como gráfico de Ramachandran, sobreposições estéricas, parâmetros de desvio C_{β} , rotâmeros e desvios de ângulo de ligação usando o programa PROCHECK (Tabela 2). De acordo com a análise dos parâmetros de validação, foi escolhido o modelo VAL.12. O desvio quadrático médio das posições atômicas (RMSD) referente a sobreposição da VaL com a VML foi de 0,111 Å. Esse valor corresponde a 91,41% de sobreposição dos átomos.

Tabela 2 – Valores referentes aos parâmetros de validação para a predição da estrutura terciária da lectina de *Vataireopsis araroba*. Os dados em vermelho correspondem às informações do modelo escolhido para modelagem tridimensional.

Modelos	ERRAT	Ramachandran (regiões permitidas)	Ramachandran (regiões parcialmente permitidas)	Ramachandran (regiões generosamente permitidas)	Ramachandran (regiões não permitidas)	QMEANDisco
VAL.12	84,4444	90.4%	8.7%	0.5%	0.5%	0.90
VAL.13	76,4444	90.4%	9.1%	0.0%	0.5%	0.89
VAL.14	83,5553	89.9%	9.1%	0.5%	0.5%	0.90
VAL.15	84	89.9%	8.7%	1.0%	0.5%	0.90
Phyre 2	54,6667	90.9%	8.7%	0.5%	0.0%	0.88

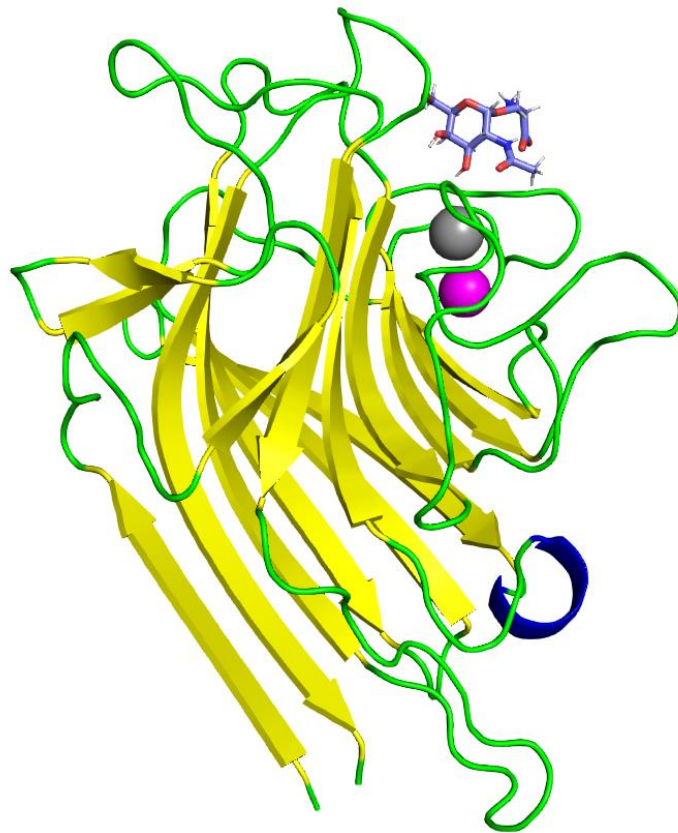
Fonte: Elaborado pela própria autora.

A VaL apresenta estrutura tridimensional com motivo do tipo *jellyroll* ou β -sanduíche, conformação amplamente conservada em lectinas de *Leguminosae*. Lectinas com motivo *jellyroll* são caracterizadas pela presença de duas folhas beta sobrepostas, como em

um sanduíche. As folhas betas são constituídas por fitas anti-paralelas, com seis fitas-beta longas e planas formando uma folha beta e sete fitas encurvadas compondo a outra folha. Cada uma destas estruturas são conectadas por *loops*, que geralmente são as regiões que apresentam grande parte das alterações de sequência (LORIS *et al.*, 1998).

A figura 14 mostra o protômero do tipo *jellyroll* da VaL, constituído por 233 resíduos de aminoácidos, um sítio de ligação a metais, com um íon cálcio e um íon manganês, e um sítio de reconhecimento a carboidratos com um ligante Tn. Esse padrão estrutural foi encontrado também em outras lectinas da tribo *Dalbergieae* como as lectinas de *Arachis hypogaea* (PNA) (BANERJEE *et al.*, 1996), *Pterocarpus angolensis* (PAL) (LORIS *et al.*, 2003), *Platypodium elegans* (PELa) (CAVADA *et al.*, 2018) e VML (SOUSA *et al.*, 2015).

Figura 14 – Representação do monômero de VaL complexado com o antígeno Tn. Os íons cálcio e manganês se encontram representados como as esferas cinza e rosa respectivamente, e o ligante Tn está representado em conformação de bastão de cor azul, posicionado no CRD da lectina.

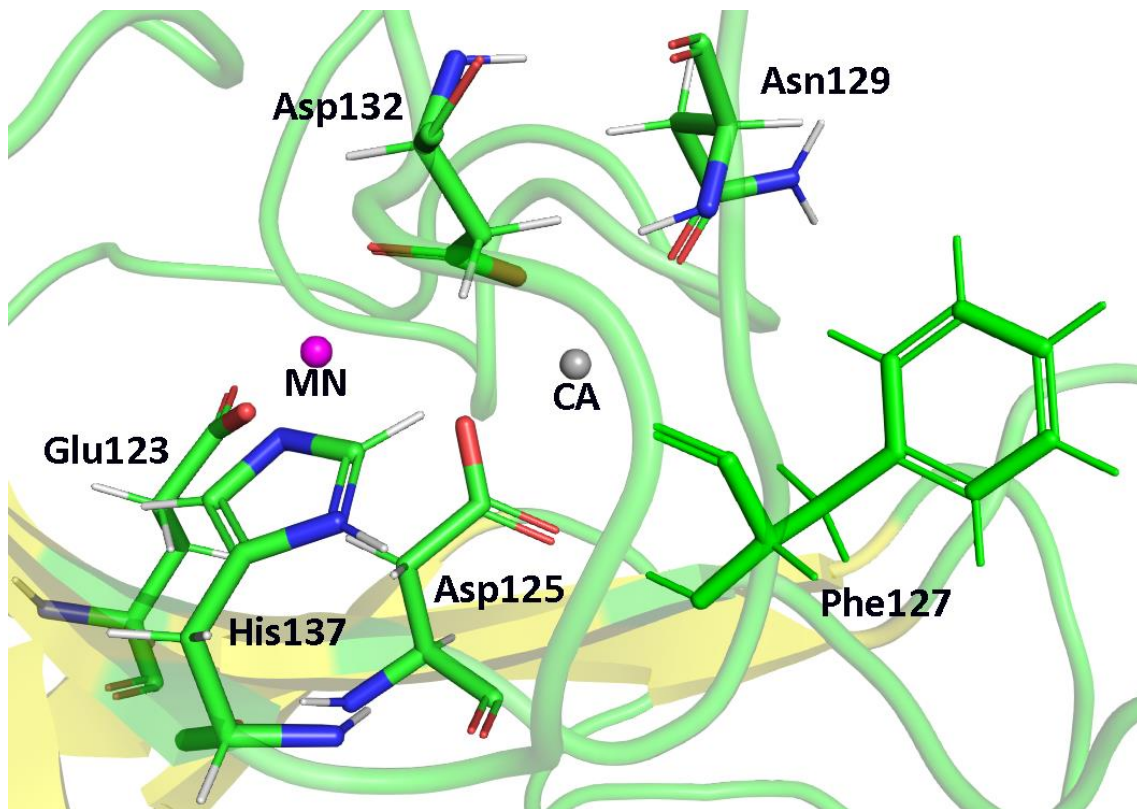


Fonte: Elaborado pela própria autora. A representação gráfica foi produzida pelo programa PyMOL v.1.7.4.5.

A presença dos íons cálcio e manganês foi confirmada na estrutura terciária de VaL, caracterizando essa lectina como uma metaloproteína. Os resíduos participantes do sítio

de ligação a metais (MBS) são bastante conservados, em comparação com a VML (SOUSA *et al.*, 2015; SOUSA *et al.*, 2016), com o íon cálcio sendo coordenado pelos resíduos Asp-125, Asn-129, Asp-132, Phe-127 e o íon manganês sendo coordenado pelos os resíduos Glu-123, Asp-125, Asp-132, His-137 (Figura 15). O MBS se encontra próximo ao CRD, pois a presença dos íons é fundamental para o estabelecimento da conformação apropriada e ativa do CRD (OSTERNE *et al.*, 2017).

Figura 15 - Representação tridimensional do MBS da VaL. Os resíduos envolvidos no sítio estão representados em conformação de bastão de cor verde. As esferas cinza e rosa representam os íons cálcio e manganês, respectivamente.



Fonte: Elaborado pela própria autora. A representação gráfica foi produzida pelo programa PyMOL v.1.7.4.5.

4.7 Docking molecular

Para entender melhor os tipos de interações que os carboidratos estabelecem com o CRD da lectina de *Vataireopsis araroba* (VaL), e procurando também estudar a possível geometria que esses ligantes adquirem durante essa ligação, alguns dos carboidratos com resultados interessantes obtidos do teste de especificidade de ligação da VaL, no estudo

desenvolvido por Osterne *et al.* (2022), foram escolhidos para as análises de *docking* molecular.

Para estas análises, a estrutura tridimensional da VaL, resolvida e validada pelo método de modelagem comparativa por homologia, teve o seu sítio de reconhecimento a carboidratos fixado e diferentes açúcares (como a D-galactose, N-acetilgalactosamina, α -lactose e L-ramnose) presentes no banco de dados PubChem foram selecionados e posicionados no CRD para a realização das interações. Além dos carboidratos, foi também selecionado para este ensaio o antígeno Tn (GalNAc-O-Ser/Thr). As disposições dos ligantes selecionados estão demonstrados na Figura 16, bem como os valores de escores obtidos se encontram na Tabela 3.

Não por acaso, o ligante Tn foi selecionado por ser um importante biomarcador de câncer, podendo ser utilizado como uma potencial ferramenta no diagnóstico, tendo em vista sua expressão precoce em células transformadas, tanto na carcinogênese humana quanto na animal (BABINO *et al.*, 2000; ITZKOWITZ *et al.*, 1992). Sabe-se que alterações dos padrões de glicosilação na superfície da membrana são características consistentes das células cancerígenas. Em tumores, essa glicosilação é consequência da síntese incompleta de cadeias de carboidratos, permitindo maior expressão de porções precursoras de carboidratos, como o antígeno Tn (CAO *et al.*, 2008). Na literatura algumas lectinas de leguminosas com afinidade para o antígeno Tn, como VML (SOUSA *et al.*, 2015), já são utilizadas como marcadores histológicos, com muitas aplicações potenciais na terapia do câncer. Então, considerando a possibilidade de usar a VaL como uma nova ferramenta para a pesquisa do câncer, foram analisadas as características estruturais deste complexo, VaL+Tn, por modelagem tridimensional e simulações de *docking* molecular.

Tabela 3 – Valores de escores, encaixes, interações de hidrogênio, interações hidrofóbicas e torções obtidas no *docking* de VaL com diferentes carboidratos. Os valores da concentração inibitória mínima (MIC) foram obtidas no estudo desenvolvido por Osterne *et al.* (2022).

Carboidratos	Escores	Encaixes	Interações de hidrogênio	Interações hidrofóbicas	Torções	MIC
L-Ramnose	-29,87	-5,10	-4,63	-20,68	0	50mM
D-Galactose	-30,79	-1,16	-7,72	-21,34	0	25mM
α-Lactose	-32,59	-4,52	-5,91	-22,49	0,77	1,6mM
N-acetil-D-galactosamina	-37,06	-4,90	-8,57	-26,04	0,42	0,4mM

Antígeno Tn	-45,01	-4,52	-8,64	-32,57	1,08	-
--------------------	--------	-------	-------	--------	------	---

Fonte: Elaborado pela própria autora.

Em geral, os valores de escores obtidos no *docking* correlacionaram bem com os dados que se encontram na literatura sobre a afinidade da VaL com diferentes açúcares, como a D-galactose, N-acetilgalactosamina, α -lactose e L-ramnose (Osterne *et al.*, 2022). Como demonstrado na tabela acima, a VaL possui interações mais favoráveis com o antígeno Tn, N-acetilgalactosamina e α -lactose com escores de -45,01, -37,06 e -32,59, respectivamente. Quanto mais negativo for o valor do escore atribuído ao complexo, mais favorável é a interação entre a lectina e o ligante testado.

No caso do complexo VaL+L-ramnose (escore -29,87) (Figura 16A) (Tabela 3), podemos observar que este complexo apresentou interações menos favoráveis devido às interações de hidrogênio envolvendo apenas os resíduos Asn-129, Gly-105, Asp-87, Leu-213. Com isso é possível notar a ausência das interações de hidrogênio com o resíduo de Ser-214, diminuindo a afinidade da VaL com o ligante (L-ramnose) e desfavorecendo sua estabilização no CRD da lectina.

No que se refere ao complexo VaL+D-galactose (escore -30,79) (Figura 16B) (Tabela 3), as interações foram razoavelmente mais favoráveis, em comparação com o complexo VaL+L-ramnose, devido a presença das interações de hidrogênio envolvendo os resíduos Asn-129, Gly-105, Asp-87, Leu-213 e também o resíduo de Ser-214. Estabelecendo um escore mais negativo de interações de hidrogênio (-7,72) e um encaixe mais favorável (-1,16) do ligante (D-galactose) no CRD da lectina, em comparação com o complexo VaL+L-ramnose.

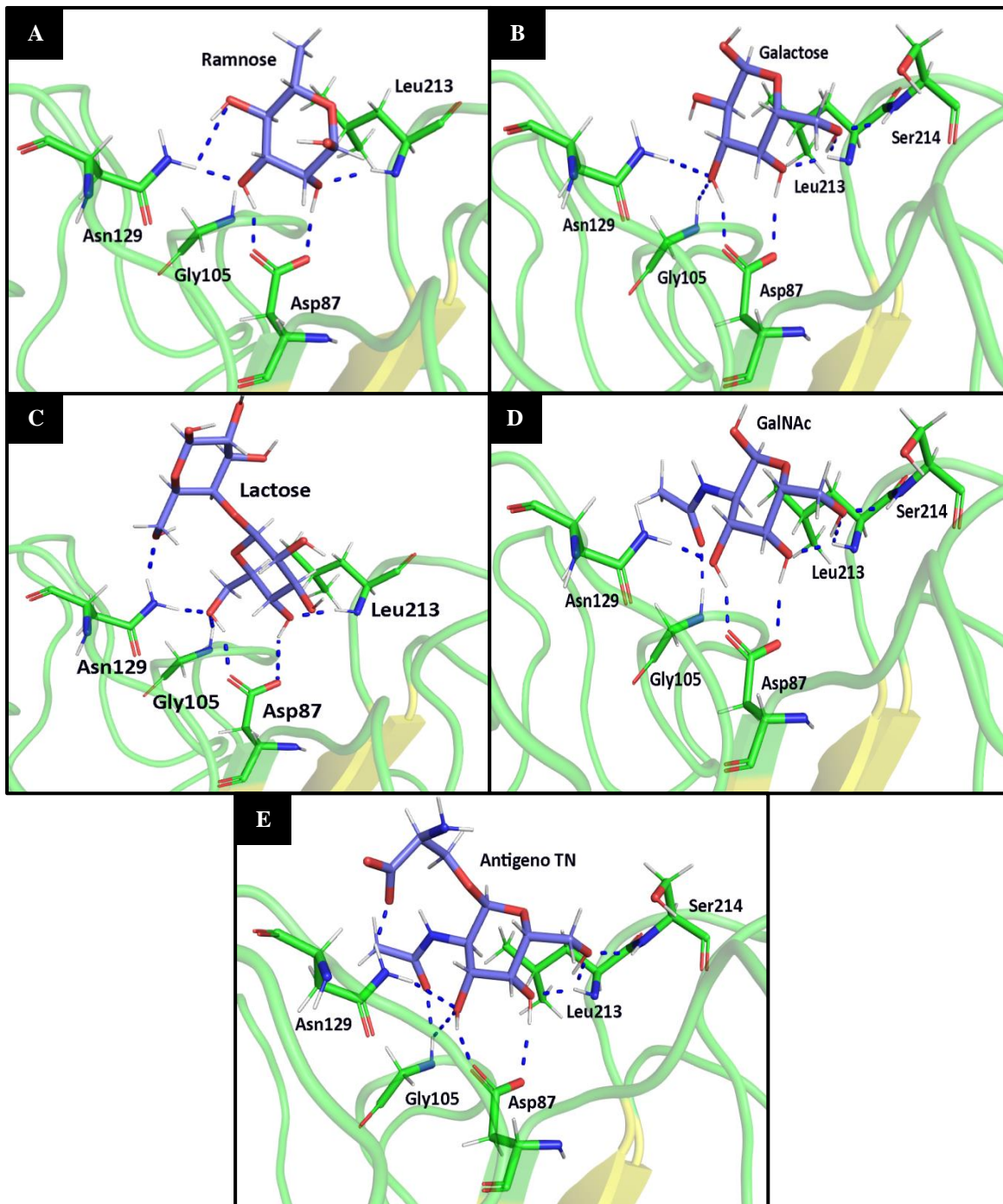
No caso do complexo VaL+ α -lactose (escore -32,59) (Figura 16C) (Tabela 3), as interações foram ainda mais favoráveis, em comparação com o complexo VaL+D-galactose, devido às interações de hidrogênio envolvendo os resíduos Asn-129, Gly-105, Asp-87, Leu-213, e neste caso há também as fortes interações não-polares (-22,49) que favorecem o encaixe (-4,52) do açúcar ligante (α -lactose) no CRD da lectina. O escore mais negativo para essas interações é provavelmente devido a α -lactose ser uma molécula com maior número de átomos para realizar estas interações.

Ademais, o complexo VaL+GalNAc (escore -37,06) (Figura 16D) (Tabela 3) apresentou interações bastante favoráveis, melhor do que os complexos já citados anteriormente. Isso é decorrente das interações de hidrogênio envolvendo os resíduos Asn-129, Gly-105, Asp-87, Leu-213 e também o resíduo de Ser-214, além das fortes interações

hidrofóbicas (-26,04) devido a presença do grupo acetil no ligante GalNAc, que favorece ainda mais o encaixe (-4,90) do açúcar ligante no CRD.

E por fim, de forma semelhante a como ocorre com o complexo VaL+GalNAc, o complexo VaL+Tn (escore -45,01) (Figura 16E) (Tabela 3) também apresenta interações de hidrogênio envolvendo os resíduos Asn-129, Gly-105, Asp-87, Leu-213 e Ser-214, porém neste caso, o CRD estabelece com o ligante fortes interações hidrofóbicas (-32,57) devido a presença do grupo acetil no ligante Tn, além da presença de um resíduo de serina ligado diretamente no açúcar, favorecendo ainda mais a estabilização do ligante no CRD.

Figura 16 – Representação do CRD da VaL ligado aos carboidratos *L*-ramnose (**A**), *D*-galactose (**B**), α -lactose (**C**), *N*-acetil-*D*-galactosamina (GalNAc) (**D**) e o antígeno Tn (**E**).

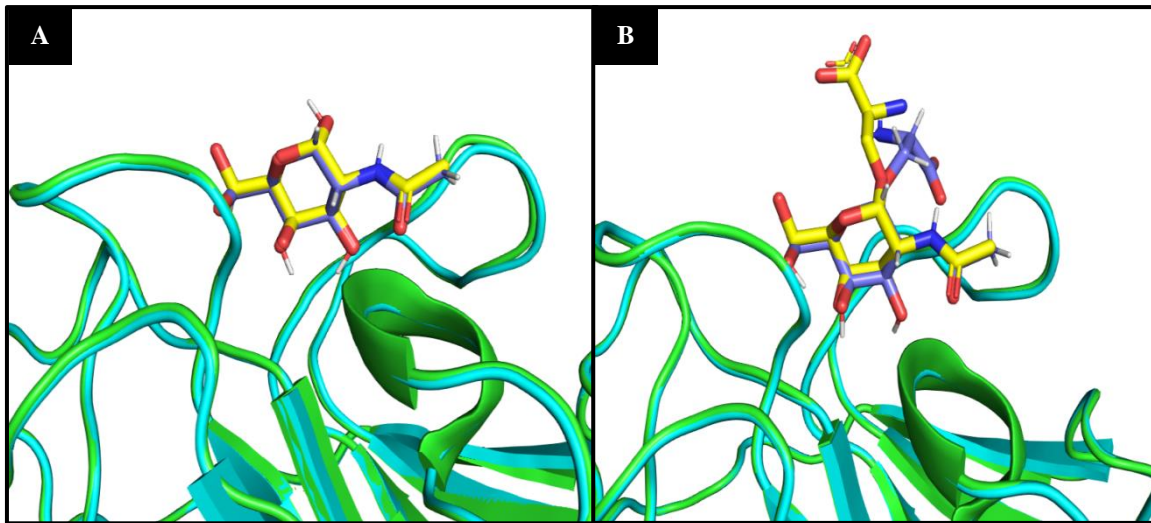


Fonte: Elaborado pela própria autora. As representações gráficas foram produzidas pelo programa PyMOL v.1.7.4.5.

Para validação das estruturas tridimensionais da VaL com os ligantes, obtidas nas simulações de *docking* molecular, foram realizadas sobreposições da estrutura da VaL com a estrutura tridimensional da VML (PDB ID: 4U36) resolvida de forma experimental (Figura 17). Através do alinhamento é possível observar que a estrutura tridimensional β -sanduíche é visivelmente bem conservada entre essas lectinas e que as mudanças estruturais possíveis de

serem visualizadas se encontram apenas na disposição parcial do ligante Tn, observado na Figura 17B. Com isso, estas análises corroboram com a hipótese de que a lectina de *Vataireopsis araroba* (VaL) é uma ferramenta promissora para a pesquisa do câncer, sendo capaz de se ligar ao antígeno Tn.

Figura 17 – Sobreposição das estruturas tridimensionais da VaL (VaL em azul, ligante em roxo) com a estrutura tridimensional da VML (VML em verde, ligante em amarelo). (A) Validação do complexo VaL+GalNAc, (B) Validação do complexo VaL+Tn.

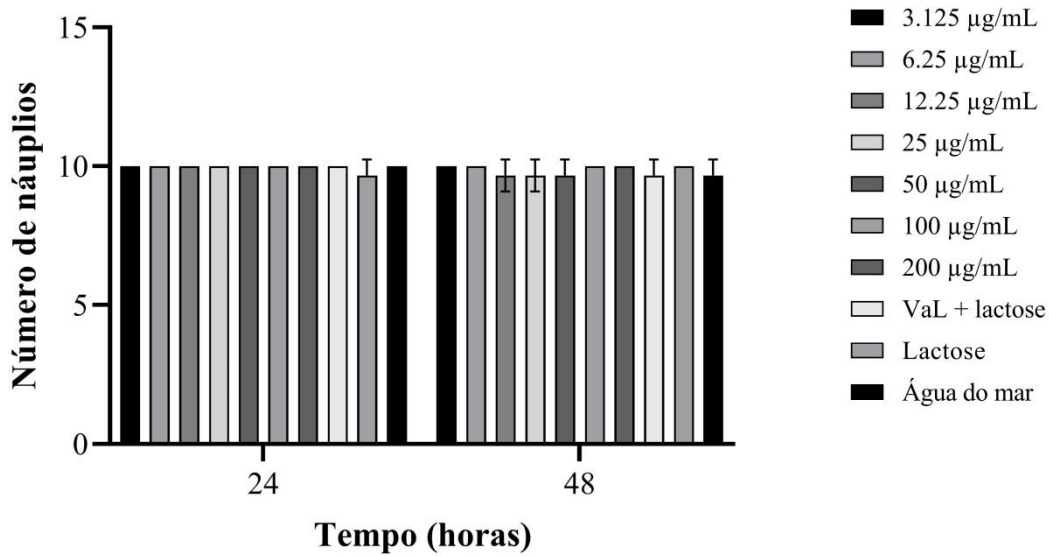


Fonte: Elaborado pela própria autora. As representações gráficas foram produzidas pelo programa PyMOL v.1.7.4.5.

4.8 Toxicidade contra náuplios de *Artemia salina*

Os ensaios com náuplios de *Artemia salina* foram utilizados para avaliar a toxicidade da lectina de *Vataireopsis araroba* (VaL). Conforme descrito no tópico 3.9, as concentrações utilizadas nesse teste variaram entre 3,125 e 200 µg/mL. Foi observado que não houve morte expressiva dos náuplios, comparativamente aos grupos controle, indicando uma não toxicidade da lectina nessas condições (Gráfico 1).

Gráfico 1 – Toxicidade contra náuplios de *Artemia salina*, após incubação por 24 horas e 48 horas. As barras representam a média da triplicata do número de náuplios vivos em cada concentração testada.



Fonte: Elaborado pela própria autora.

Os resultados obtidos no teste foram analisados com base em cálculos estatísticos, utilizando análise de variância, onde os valores médios de náuplios vivos de cada concentração foram comparados quantitativamente com os valores médios de náuplios vivos de cada controle. Assim, foi observado que não houve diferença estatística significativa indicando que as concentrações testadas da VaL não foram tóxicas para a *Artemia sp* no intervalo de tempo de 24 e 48 horas.

Em contrapartida aos resultados obtidos no presente estudo, outras lectinas da subtribo *Diocleinae* exibiram atividade citotóxica em vários ensaios frente à *Artemia sp.*, como as lectina de *Canavalia brasiliensis* (ConBr), *Canavalia ensiformis* (ConA) e *C. boliviana* (ConBol), apresentando toxicidade com CL₅₀ (Concentrações Letais) de 54,38 µg/mL, 376,48 µg/mL e 218,13 µg/mL, respectivamente (ARRUDA *et al.*, 2013). Com isso, é possível observar que algumas lectinas exibem atividade citotóxica contra náuplios de *Artemia sp.*

5 CONCLUSÃO

Com base nos resultados apresentados neste trabalho, podemos concluir que a lectina presente nas sementes de *Vataireopsis araroba* (VaL) teve sua estrutura primária parcial sequenciada por espectrometria de massas em tandem (MS/MS). A VaL apresentou um alto grau de identidade, em relação a sua estrutura primária, e um alto grau de cobertura com outras lectinas da tribo *Dalbergieae* específicas para *N*-acetilgalactosamina/galactose. A predição da estrutura secundária da VaL seguiu o modelo observado nas lectinas da tribo *Dalbergieae*, com grande predominância de folhas beta. Sua estrutura terciária apresentou o mesmo padrão conformacional de outras lectinas de *Leguminosae*, sendo composta pelo motivo *jellyroll*. Os sítios de ligação a metais e de reconhecimento a carboidratos apresentaram composição de aminoácidos idênticos aos de outras lectinas da tribo *Dalbergieae*. E por fim, nos experimentos de *docking* com diferentes carboidratos, foi possível observar que a VaL apresentou alta afinidade com o antígeno Tn, um importante biomarcador de câncer, podendo-se considerar a possibilidade de usar a VaL como uma nova ferramenta para a pesquisa do câncer.

Como perspectiva para as próximas etapas deste trabalho podemos citar a realização de estudos cristalográficos para melhor caracterizar a estrutura tridimensional da VaL, experimentos de dinâmica molecular para avaliar por simulação computacional a ligação da VaL com o antígeno Tn, assim considerando a variável de tempo e a flexibilidade do complexo como um todo. Além disso, também podem ser inclusos a realização de ensaios biológicos para reforçar os resultados obtidos em ensaios *in silico*, tornando a VaL um potencial alvo terapêutico no combate contra o câncer.

REFERÊNCIAS

- ALENCAR, N.M.N.; ASSREUY, A.M.S.; ALENCAR, V.B.M.; MELO, S.C.; RAMOS, M.V.; CAVADA, B.S.; CUNHA, F.Q.; RIBEIRO, R.A. The galactose-binding lectin from *Vatairea macrocarpa* seeds induces in vivo neutrophil migration by indirect mechanism. **The International Journal of Biochemistry & Cell Biology**, v. 35, n. 12, p. 1674-1681, 2003.
- ALENCAR, N.M.N.; CAVALCANTE, C.F.; VASCONCELOS, M.P.; LEITE, K.B.; ARAGÃO, K.S.; ASSREUY, A.M.S.; NOGUEIRA, N.A.P.; CAVADA, B.S.; VALE, M.R. Anti-inflammatory and antimicrobial effect of lectin from *Lonchocarpus sericeus* seeds in an experimental rat model of infectious peritonitis. **Journal of pharmacy and pharmacology**, v. 57, n. 7, p. 919-922, 2005.
- ALMEIDA, ALYSSON CHAVES. **Caraterização estrutural e biológica de uma lectina de sementes de *Centrolobium tomentosum* GUILL. EX BENTH.** Tese (Doutorado em Bioquímica) - Centro de Ciências, Universidade Federal do Ceará, Fortaleza, 2016.
- ALMEIDA, A.C.; OSTERNE, V.J.S.; SANTIAGO, M.Q.; PINTO-JUNIOR, V.R.; SILVA-FILHO, J.C.; LOSSIO, C.F.; NASCIMENTO, F.L.F.; ALMEIDA, R.P.H.; TEIXEIRA, C.S.; LEAL, R.B.; DELATORRE, P.; ROCHA, B.A.M.; ASSREUY, A.M.S.; NASCIMENTO, K.S.; CAVADA, B.S. Structural analysis of *Centrolobium tomentosum* seed lectin with inflammatory activity. **Archives of biochemistry and biophysics**, v. 596, p. 73-83, 2016.
- ALMEIDA, A.C.; SILVA, H.C.; PEREIRA-JUNIOR, F.N.; CAJAZEIRAS, J.B.; DELATORRE, P.; NAGANO, C.S.; NASCIMENTO, K.S.; CAVADA, B.S. Purification and partial characterization of a new mannose/glucose-specific lectin from *Centrolobium tomentosum* Guill. ex Benth seeds exhibiting low toxicity on *Artemia* sp. **International Journal of Indigenous Medicinal Plants**, v. 47, p. 1567–1577, 2014a.
- ALTSCHUL, S.F.; GISH, W.; MILLER, W.; MYERS, E.W.; LIPMAN, D.J. Basic local alignment search tool. **Journal of molecular biology**, v. 215, n. 3, p. 403-410, 1990.
- ALVES FILHO, J.G.; NASCIMENTO, A.S.F.; GONDIM, A.C.S.; PEREIRA, R.H.; CUNHA, R.M.S.; NAGANO, C.S.; TEIXEIRA, E.H.; NASCIMENTO, K.S.; CAVADA, B.S. Isoform characterisation, heterologous expression and functional analysis of two lectins from *Vatairea macrocarpa*. **Protein and Peptide Letters**, v. 20, n. 11, p. 1204-1210, 2013.
- ARARIPE, D.A.; PINTO-JUNIOR, V.R.; NECO, A.H.B.; SANTIAGO, M.Q.; OSTERNE, V.J.S.; PIRES, A.F.; LOSSIO, C.F.; MARTINS, M.G.Q.; CORREIA, J.L.A.; BENEVIDES, R.G.; LEAL, R.B.; ASSREUY, A.M.S.; NASCIMENTO, K.S.; CAVADA, B.S. Partial characterization and immobilization in CNBr-activated Sepharose of a native lectin from *Platypodium elegans* seeds (PELa) and comparative study of edematogenic effect with the recombinant form. **International journal of biological macromolecules**, v. 102, p. 323-330, 2017.
- ARRUDA, F. V. S.; MELO, A. A.; VASCONCELOS, M. A.; CARNEIRO, R. F.; BARROSO-NETO, I. L.; SILVA, S. R.; SAMPAIO, A. H.; CAVADA, B. S. Toxicity and binding profile of lectins from the Genus *canavalia* on brine shrimp. **BioMed Research International**, v. 2013, 2013.

AVICHEZER, D.; ARNON, R. Differential reactivities of the *Arachis hypogaea* (peanut) and *Vicia villosa* B4 lectins with human ovarian carcinoma cells, grown either in vitro or in vivo xenograft model. **FEBS letters**, v. 395, n. 2-3, p. 103-108, 1996.

AZANI, N.; BABINEAU, M.; BAILEY, C.D.; BANKS, H.; BARBOSA, A.R.; PINTO, R.B.; BOATWRIGHT, J.S.; BORGES, L.M.; BROWN, G.K.; BRUNEAU, A.; CANDIDO, E.; CARDOSO, D.; CHUNG, K.-F.; CLARK, R.P.; CONCEIÇÃO, A. S.; CRISP, M.; CUBAS, P.; DELGADO-SALINAS, A.; DEXTER, K.G.; DOYLE, J.J.; DUMINIL, J.; EGAN, A.N.; ESTRELLA, M.; FALCÃO, M.J.; FILATOV, D.A.; FORTUNA-PEREZ, A.P.; FORTUNATO, R.H.; GAGNON, E.; GASSON, P.; RANDO, J.G.; AZEVEDO TOZZI, A.M.G.; GUNN, B.; HARRIS, D.; HASTON, E.; HAWKINS, J.A.; HERENDEEN, P.S.; HUGHES, C.E.; IGANCI, J.R.V.; JAVADI, F.; KANU, S.A.; KAZEMPOUR-OSALOO, S.; KITE, G.C.; KLITGAARD, B.B.; KOCHANOVSKI, F.J.; KOENEN, E.J.M.; KOVAR, L.; LAVIN, M.; ROUX, M.; LEWIS, G.P.; LIMA, H.C.; LÓPEZROBERTS, M.C.; MACKINDER, B.; MAIA, V.H.; MALÉCOT, V.; MANSANO, V.F.; MARAZZI, B.; MATTAPHA, S.; MILLER, J.T.; MITSUYUKI, C.; MOURA, T.; MURPHY, D.J.; NAGESWARA-RAO, M.; NEVADO, B.; NEVES, D.; OJEDA, D.I.; PENNINGTON, R.T.; PRADO, D.E.; PRENNER, G.; QUEIROZ, L.P.; RAMOS, G.; FILARDI, F.L.R.; RIBEIRO, P.G.; RICO-ARCE, M. L.; SANDERSON, M.J.; TOS-SILVA, J. S.; SÃO-MATEUS, W.M.B.; SILVA, M.J.S.; SIMON, M.F.; SINOU, C.; SNAK, C.; SOUZA, É.R.; SPRENT, J.; STEELE, K.P.; STEIER, J.E.; STEEVES, R.; STIRTON, C.H.; TAGANE, S.; TORKE, B.M.; TOYAMA, H. CRUZ, D.T.; VATANPARAST, M.; WIERINGA, J.J.; WINK, M.; WOJCIECHOWSKI, M.F.; YAHARA, T.; YI, T.; ZIMMERMAN, E. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny: the legume phylogeny working group (LPWG). **Taxon**, v. 66, n. 1, p. 44-77, 2017.

BABINO, A.; OPPEZZO, P.; BIANCO, S.; BARRIOS, E.; BEROIS, N.; NAVARRETE, H.; OSINAGA, E. Tn antigen is a pre-cancerous biomarker in breast tissue and serum in n-nitrosomethylurea-induced rat mammary carcinogenesis. **International journal of cancer**, v. 86, n. 6, p. 753-759, 2000.

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. **Journal of molecular biology**, v. 259, n. 2, p. 281-296, 1996.

BENEVIDES, R.G.; GANNE, G.; SIMÕES, R.C.; SCHUBERT, V.; NIEMIETZ, M.; UNVERZAGT, C.; CHAZALET, V.; BRETON, C.; VARROT, A.; CAVADA, B.S.; IMBERTY, A. A lectin from *Platypodium elegans* with unusual specificity and affinity for asymmetric complex N-glycans. **Journal of Biological Chemistry**, v. 287, n. 31, p. 26352-26364, 2012.

BENKERT, P.; BIASINI, M.; SCHWEDE, T. Toward the estimation of the absolute quality of individual protein structure models. **Bioinformatics**, v. 27, n. 3, p. 343-350, 2011.

BENKERT, P.; KÜNZLI, M.; SCHWEDE, T. QMEAN server for protein model quality estimation. **Nucleic acids research**, v. 37, n. suppl_2, p. W510-W514, 2009.

BUCHAN, D.W.A.; MINNECI, F.; NUGENT, T.C.O.; BRYSON, T. C.; JONES, D. T. Scalable web services for the PSIPRED protein analysis Workbench. **Nucleic acids research**, v. 41, n. W1, p. W349-W357, 2013.

BONNARDEL, F.; PEREZ, S.; LISACEK, F.; IMBERTY, A. Structural Database for Lectins and the UniLectin Web Platform. **Methods Mol. Biol.**, v. 2132, p. 1–14, 2020.

BOWIE, J. U.; LÜTHY, R.; EISENBERG, D. A method to identify protein sequences that fold into a known three-dimensional structure. **Science**, v. 253, n. 5016, p. 164-170, 1991.

CAO, Y.; MERLING, A.; KARSTEN, U.; GOLETZ, S.; PUNZEL, M.; KRAFT, R.; SCHWARTZ-ALBIEZ, R. Expression of CD175 (Tn), CD175s (sialosyl-Tn) and CD176 (Thomsen-Friedenreich antigen) on malignant human hematopoietic cells. **International journal of cancer**, v. 123, n. 1, p. 89-99, 2008.

CARDOSO, D.; PENNINGTON, R.T.; QUEIROZ, L.P.; BOATWRIGHT, J.S.; VAN WYK, B.E.; WOJCIECHOWSKI, M.F.; LAVIN, M. Reconstructing the deep-branching relationships of the papilionoid legumes. **South African Journal of Botany**, v. 89, p. 58-75, 2013.

CARDOSO, D.; QUEIROZ, L. P.; LIMA, H. C.; SUGANUMA, E.; BERG, C. V. D.; LAVIN, M. A molecular phylogeny of the vataireoid legumes underscores floral evolvability that is general to many early-branching papilionoid lineages. **American Journal of Botany**, v. 100, n. 2, p. 403-421, 2013.

CALVETE, J.J.; SANTOS, C.F.; MANN, K.; GRANGEIRO, T.B.; NIMTZ, M.; URBANKE, C.; CAVADA, B. S. Amino acid sequence, glycan structure, and proteolytic processing of the lectin of *Vatairea macrocarpa* seeds. **FEBS letters**, v. 425, n. 2, p. 286-292, 1998.

CARRONI, M.; SAIBIL, H.R. Cryo Electron Microscopy to Determine the Structure of Macromolecular Complexes. **Methods**, v. 95, p. 78-85, 2016.

CAVADA, B. S., OSTERNE, V. J. S., LOSSIO, C. F., PINTO-JUNIOR, V. R., OLIVEIRA, M. V., SILVA, M. T. L., LEAL, R.B., NASCIMENTO, K.S. One century of ConA and 40 years of ConBr research: A structural review. **International Journal of Biological Macromolecules**, v. 134, p. 901-911, 2019.

CAVADA, B. S.; BARI, A. U.; PINTO-JUNIOR, V. R.; LOSSIO, C. F.; SILVA, M. T. L.; Luiz SOUZA, A. G.; OLIVEIRA, M. V.; SOUZA-FILHO, C. H. D.; CORREIA, S. E. G.; VITAL, A. P. M. S.; LIMA, L. D.; OSTERNE, V. J. S.; NASCIMENTO, K. S. Purification and partial characterization of a new lectin from *Parkia panurensis* Benth. ex HC Hopkins seeds (Leguminosae family; Mimosoideae subfamily) and evaluation of its biological effects. **International journal of biological macromolecules**, v. 145, p. 845-855, 2019.

CAVADA, B.S.; ARARIPE, D.A.; SILVA, I.B.; PINTO-JUNIOR, V.R.; OSTERNE, V.J.S.; NECO, A.H.B.; LARANJEIRA, E.P.P.; LOSSIO, C.F.; CORREIA, J.L.A.; PIRES, A.F.; ASSREUY, A.M.S.; NASCIMENTO, K.S. Structural studies and nociceptive activity of a native lectin from *Platypodium elegans* seeds (nPELa). **International journal of biological macromolecules**, v. 107, p. 236-246, 2018.

CAVADA, B.S.; OSTERNE, V.J.S.; OLIVEIRA, M.V.; PINTO-JUNIOR, V.R.; SILVA, M.T.L.; BARI, A.U.; LIMA, L.D.; LOSSIO, C.F.; NASCIMENTO, K.S. Reviewing Mimosoideae lectins: a group of under explored legume lectins. **International Journal of Biological Macromolecules**, v. 154, p. 159-165, 2020.

CAVADA, B.S.; PINTO-JUNIOR, V.R.; OSTERNE, V.J.S.; OLIVEIRA, M.V.; LOSSIO, C.F.; SILVA, M.T.L.; BARI, A.U.; LIMA, L.D.; SOUZA-FILHO, C.H.D.; NASCIMENTO, K.S. Comprehensive review on Caelsalpinioideae lectins: From purification to biological activities. **International Journal of Biological Macromolecules**, v. 162, p. 333-348, 2020.

CAVADA, B.S.; SANTOS, C.F.; GRANGEIRO, T.B.; NUNES, E.P.; SALES, P.V.; RAMOS, R.L.; SOUSA, F.A.; CRISOSTOMO, C.V.; CALVETE, J.J. Purification and characterization of a lectin from seeds of *Vatairea macrocarpa* Duke. **Phytochemistry**, v. 49, n. 3, p. 675-680, 1998.

CIPOLLA, L.; RESCIGNO, M.; LEONE, A.; PERI, F.; LA FERLA, B.; NICOTRA, F. Novel Tn antigen containing neoglycopeptides: synthesis and evaluation as antitumor vaccines. **Bioorganic & medicinal chemistry**, v. 10, n. 5, p. 1639-1646, 2002.

COLOVOS, C.; YEATES, T. O. Verification of protein structures: patterns of nonbonded atomic interactions. **Protein science**, v. 2, n. 9, p. 1511-1519, 1993.

DE CONINCK, T.; VAN DAMME, E. J. M. Plant lectins: handymen at the cell surface. **The Cell Surface**, p. 100091, 2022.

DESAI, P.R. Immunoreactive T and Tn antigens in malignancy: role in carcinoma diagnosis, prognosis, and immunotherapy. **Transfusion medicine reviews**, v. 14, n. 4, p. 312-325, 2000.

ELFSTRAND, M. Über blutkörperchenagglutinierende Eiweisse. **Görberdorfer Veröffentlichungen a. Band I**, p. 1-159, 1898.

ESTRELLA, M.; FOREST, F.; KLITGÅRD, B.; LEWIS, G.P.; MACKINDER, B.A. QUEIROZ, L.P.; WIERINGA, J.J.; BRUNEAU, A. A new phylogeny-based tribal classification of subfamily Detarioideae, an early branching clade of florally diverse tropical arborescent legumes. **Scientific Reports**, v. 8, n. 1, p. 1-14, 2018.

FREITAS PIRES, A.; BEZERRA, M.M.; AMORIM, R.M.F.; NASCIMENTO, F.L.F.; MARINHO, M.M.; MOURA, R.M.; SILVA, M.T.L.; CORREIA, J.L.A.; CAVADA, B.S.; ASSREUY, A.M.S.; NASCIMENTO, K.S. Lectin purified from *Lonchocarpus campestris* seeds inhibits inflammatory nociception. **International journal of biological macromolecules**, v. 125, p. 53-60, 2019.

GINALSKI, K. Comparative Modeling for Protein Structure Prediction. **Current opinion in structural biology**, v. 16, p. 172-177, 2006.

GOUET, P.; COURCELLE, E.; STUART, D. I.; MÉTOZ, F. ESPript: Analysis of multiple sequence alignments in PostScript. **Bioinformatics**, v. 15, n. 4, p. 305-308, 1999.

HAKOMORI S. Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens. **Advances in cancer research**, v. 52, p. 257-331, 1989.

ITZKOWITZ, S.; BLOOM, E.; LAU, T.; KIM, Y. Mucin associated Tn and sialosyl-Tn antigen expression in colorectal polyps. **Gut**, v. 33, n. 4, p. 518-523, 1992.

JUAN, L. L.; RECIO, V. G.; LÓPEZ, P. J.; JUAN, T. G.; CORDOBA-DIAZ, M.; CORDOBA-DIAZ, D. Pharmaceutical applications of lectins. **Journal of Drug Delivery Science and Technology**, v. 42, p. 126-133, 2017.

KELLEY, L. A.; MEZULIS, S., YATES, C. M., WASS, M. N., STERNBERG, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. **Nature protocols**, v. 10, n. 6, p. 845-858, 2015.

KERKHOF, P.C.M.V.D.; VALK, P.G.M.V. D.; SWINKELS, O.Q.J.; KUCHARKOVA, M.; RIE, M.A.D.; VRIES, H.J.C.; DAMSTRA, R.; ORANJE, A.P. W.V.D; NEER, P. V.; LIJNEN, R.L.P.; KUNKELER, A.C.M.; HEES, C. V.; HAERTLEIN, N.G.J.; HOL, C.W. A comparison of twice-daily calcipotriol ointment with once-daily short-contact dithranol cream therapy: a randomized controlled trial of supervised treatment of psoriasis vulgaris in a day-care setting. **British Journal of Dermatology**, v. 155, n. 4, p. 800-807, 2006.

KLEBE, G. Experimental Methods of Structure Determination. In *Drug Design: Methodology, Concepts, and Mode-of-Action*, Heidelberg, 2013.

KORB, O.; STÜTZLE, T.; EXNER, T.E. Empirical scoring functions for advanced protein-ligand docking with PLANTS. **Journal of chemical information and modeling**, v. 49, n. 1, p. 84-96, 2009.

LANNOO, N.; VAN DAMME, E. J. M. Nucleocytoplasmic plant lectins. **Biochimica et Biophysica Acta (BBA)-General Subjects**, v. 1800, n. 2, p. 190-201, 2010.

LAGARDA-DIAZ, I.; GUZMAN-PARTIDA, A. M.; VAZQUEZ-MORENO, L. Legume Lectins: Proteins with Diverse Applications. **International Journal of Molecular Sciences**, v. 18, n. 6, p. 1242, 2017.

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

LASKOWSKI, R.A.; MACARTHUR, M.W.; MOSS, D.S.; THORNTON, J. M. PROCHECK: a program to check the stereochemical quality of protein structures, **Journal of applied crystallography**, v. 26, n. 2, p. 283-291, 1993.

LAUWEREYS, M.; FORIERS, A.; SHAROR, N.; STROSBERG, A.D. Sequence studies of peanut agglutinin. **FEBS letters**, v. 181, n. 2, p. 241-244, 1985.

LAW, I.J.; HAYLETT, T.; STRIJDOM, B.W. Differences in properties of peanut seed lectin and purified galactose- and mannose-binding lectins from nodules of peanut. **Planta**, v. 176, n. 1, p. 19-27, 1988.

LEWIS, G.P.; SCHRIRE, B.; MACKINDER, B.; LOCK, M. Legumes of the World. Royal Botanic Gardens, **Kew**, 2005.

LI, X.-T.; HE, M.-L.; ZHOU, Z.-Y.; JIANG, Y.; CHENG, L. The antitumor activity of PNA modified vinblastine cationic liposomes on Lewis lung tumor cells: In vitro and in vivo evaluation. **International journal of pharmaceutics**, v. 487, n. 1-2, p. 223-233, 2015.

LIMA, H.C. Revisão taxonômica do gênero *Vataireopsis* Ducke (Leg. Fab.). **Rodriguésia**, p. 21-40, 1980.

LIN, P.; NG, T.B.; Preparation and biological properties of a melibiose binding lectin from *Bauhinia variegata* seeds. **Journal of Agricultural and Food Chemistry**, v. 56, n. 22, p. 10481-10486, 2008.

LORIS, R.; HAMELRYCK, T.; BOUCKAERT, J.; WYNS, L. Legume lectin structure. **Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology**, v. 1383, n. 1, p. 9–36, 1998.

LORIS, R.; IMBERTY, A.; BEECKMANS, S.; VAN DRIESSCHE, E.; READ, J.S.; BOUCKAERT, J.; GREVE, H.; BUTS, L.; WYNS, L. Crystal structure of *Pterocarpus angolensis* lectin in complex with glucose, sucrose, and turanose. **Journal of Biological Chemistry**, v. 278, n. 18, p. 16297-16303, 2003.

LOTAN, R.; SKUTELSKY, E.; DANON, D.; SHARON, N. The purification, composition, and specificity of the anti-T lectin from peanut (*Arachis hypogaea*). **Journal of Biological Chemistry**, v. 250, n. 21, p. 8518-8523, 1975.

LÜTHY, R.; BOWIE, J.U.; EISENBERG, D. Assessment of protein models with threedimensional profiles. **Nature**, v. 356, n. 6364, p. 83-85, 1992.

MARQUES, G.F.O.; OSTERNE, V.J.S.; ALMEIDA, L.M.; OLIVEIRA, M.V.; BRIZENO, L.A.C.; PINTO-JUNIOR, V.R.; SANTIAGO, M.Q.; NECO, A.H.B.; MOTA, M.R.L.; SOUZA, L.A.G.; NASCIMENTO, K.S.; PIRES, A.F.; CAVADA, B.S.; ASSREUY, A.M.S. Contribution of the carbohydrate-binding ability of *Vatairea guianensis* lectin to induce edematogenic activity. **Biochimie**, v. 140, p. 58-65, 2017.

MARQUES, G.F.O.; PIRES, A.F.; OSTERNE, V.J.S.; PINTO-JUNIOR, V.R.; SILVA, I.B.; MARTINS, M.G.Q.; OLIVEIRA, M.V.; GOMES, A.M.; SOUZA, L.A.G.; PAVÃO, M.S.G.; CAVADA, B.S.; ASSREUY, A.M.S.; NASCIMENTO, K.S. *Vatairea guianensis* lectin stimulates changes in gene expression and release of TNF- α from rat peritoneal macrophages via glycoconjugate binding. **Journal of Molecular Recognition**, v. 34, n. 10, p. e2922, 2021.

MARTÍNEZ-ALARCÓN, D.; BLANCO-LABRA, A.; GARCÍA-GASCA, T. Expression of lectins in heterologous systems. **International journal of molecular sciences**, v. 19, n. 2, p. 616, 2018.

MENG, X. Y.; ZHANG, H. X.; MEZEI, M.; CUI, M. Molecular docking: a powerful approach for structure-based drug discovery. **Current computer-aided drug design**, v. 7, n. 2, p. 146–57, 2011.

NASCIMENTO, K.S.; ARARIPE, D.A.; PINTO-JUNIOR, V.R.; OSTERNE, V.J.S.; MARTINS, F.W.V.; NECO, A.H.B.; FARIAS, G.A.; CAVADA, B.S. Homology modeling, molecular docking, and dynamics of two α -methyl-D-mannoside-specific lectins from *Arachis* genus. **Journal of molecular modeling**, v. 24, n. 9, p. 1-10, 2018.

NASCIMENTO, K.S.; SILVA, M.T.L.; OLIVEIRA, M.V.; LOSSIO, C.F.; PINTO-JUNIOR, V.R.; OSTERNE, V.J.S.; CAVADA, B.S. Dalbergieae lectins: a review of lectins from species of a primitive Papilionoideae (leguminous) tribe. **International Journal of Biological Macromolecules**, v. 144, p. 509-526, 2020.

NECO, A. H. B.; PINTO-JUNIOR, V.R.; ARARIPE, D.A.; SANTIAGO, M.Q.; OSTERNE, V.J.S.; LOSSIO, C.F.; NOBRE, C.A.S.; OLIVEIRA, M.V.; SILVA, M.T.L.; MARTINS, M.G.Q.; CAJAZEIRAS, J.B.; MARQUES, G.F.O.; COSTA, D.R.; NASCIMENTO, K.S.; ASSREUY, A.M.S.; CAVADA, B.S. Structural analysis, molecular docking and molecular dynamics of an edematogenic lectin from *Centrolobium microchaete* seeds. **International journal of biological macromolecules**, v. 117, p. 124–133, 2018.

OSTERNE, V. J. S.; SILVA-FILHO, J. C.; SANTIAGO, M. Q.; PINTO-JUNIOR, V. R.; ALMEIDA, A. C.; BARRETO, A. A. G. C.; CAVADA, B. S. Structural characterization of a lectin from *Canavalia virosa* seeds with inflammatory and cytotoxic activities. **International journal of biological macromolecules**, v. 94, p. 271-282, 2017.

OSTERNE, V.J.S.; OLIVEIRA, M.V.; SCHUTTER, K.; SERNA, S.; REICHARDT, N.C.; SMAGGHE, G.; CAVADA, B.S.; VAN DAMME, E.J.M.; NASCIMENTO, K. S. A galactoside-specific Dalbergieae legume lectin from seeds of *Vataireopsis araroba* (Aguiar) Ducke. **Glycoconjugate Journal**, p. 1-11, 2022.

PAGADALA, N. S.; SYED, K.; TUSZYNSKI, J. Software for molecular docking: a review. **Biophysical reviews**, v. 9, n. 2, p. 91-102, 2017.

PATEL, B.; SINGH, V.; PATEL, D. Structural Bioinformatics. In *Essentials of Bioinformatics*. **Springer**, v. I, p. 169–199, 2019

PATHAK, M.; SINGH, B.; SHARMA, A.; AGRAWAL, P.; PASHA, S.B.; DAS, H.R.; Molecular cloning, expression, and cytokinin (6-benzylaminopurine) antagonist activity of peanut (*Arachis hypogaea*) lectin SL-I. **Plant molecular biology**, v. 62, n. 4, p. 529-545, 2006.

PEUMANS, W.J., VAN DAMME, E.J.M.: Lectins as plant defense proteins. **Plant Physiol.** v. 109, n. 2, p. 347–352, 1995.

PIRES, A.F.; RODRIGUES, N.V.F.C.; SOARES, P.M.G.; RIBEIRO, R. A.; ARAGÃO, K.S.; MARINHO, M.M.; SILVA, M.T.L.; CAVADA, B.S.; ASSREUY, A.M.S. A novel N-acetylglucosamine lectin of *Lonchocarpus araripensis* attenuates acute cellular inflammation in mice. **Inflammation Research**, v. 65, n. 1, p. 43-52, 2016.

POLHILL, R. M.; RAVEN, P. H.; Royal botanic gardens. *Advances in legume systematics*. **Kew (Ed.)**, 1981.

REYES-MONTAÑO, E.A.; VEGA-CASTRO, N.A. Plant Lectins with Insecticidal and Insectistatic Activities. **Insecticides - Agriculture and Toxicology**, p. 17-42, 2018.

ROBERT, X.; GOUET, P. Deciphering key features in protein structures with the new ENDscript server. **Nucleic acids research**, v. 42, n. W1, p. W320-W324, 2014.

RYDER, S.D.; SMITH, J.A.; RHODES, J.M. Peanut lectin: a mitogen for normal human colonic epithelium and human HT29 colorectal cancer cells. **JNCI: Journal of the National Cancer Institute**, v. 84, n. 18, p. 1410-1416, 1992.

SCHWEDE, T. Protein Modeling: What Happened to the “Protein Structure Gap”? **Structure**, v. 21, n. 9, p. 1531-1540, 2013.

SHEVCHENKO, A.; TOMAS, H.; HAVLIS, J.; OLSEN, J.V.; MANN, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. **Nature protocols**, v. 1, n. 6, p. 2856-2860, 2006.

SIEVERS, F.; WILM, A.; DINEEN, D.; GIBSON, T.J.; KARPLUS, K.; LI, W.; LOPES, R.; MCWILLIAM, H.; REMMERT, M.; SÖDING, J.; THOMPSON, J.D.; HIGGINS, D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular systems biology**, v. 7, p. 539, 2011.

SILVA, M.L.S. Lectin-based biosensors as analytical tools for clinical oncology. **Cancer Letters**, v. 436, p. 63-74, 2018.

SILVA, H.C.; NAGANO, C.S.; SOUZA, L.A.G.; NASCIMENTO, K.S.; ISÍDRO, DELATORRE, R.; ROCHA, P. B.A.M.; SAMPAIO, A.H.; ASSREUY, A.M.S.; PIRES, A.F.; DAMASCENO, L.E.A.; MARQUES-DOMINGOS, G.F.O.; CAVADA, B.S. Purification and primary structure determination of a galactose-specific lectin from *Vatairea guianensis* Aublet seeds that exhibits vasorelaxant effect. **Process Biochemistry**, v. 47, n. 12, p. 2347-2355, 2012.

SOUSA, B.L.; FILHO, J.C. S.; KUMAR, P.; GRAEWERT, M.A.; PEREIRA, R.I.; CUNHA, R.M.S.; NASCIMENTO, K.S.; BEZERRA, G.A.; DELATORRE, P.; DJINOVIC-CARUGO, K.; NAGANO, C.S.; GRUBER, K.; CAVADA, B.S. Structural characterization of a *Vatairea macrocarpa* lectin in complex with a tumor-associated antigen: a new tool for cancer research. **The international journal of biochemistry & cell biology**, v. 72, p. 27-39, 2016.

SOUSA, B.L.; FILHO, J.C.S.; KUMAR, P.; PEREIRA, R.I.; ŁYSKOWSKI, A.; ROCHA, B.A.M.; DELATORRE, P.; BEZERRA, G.A.; NAGANO, C.S.; GRUBER, K.; CAVADA, B.S. High-resolution structure of a new Tn antigen-binding lectin from *Vatairea macrocarpa* and a comparative analysis of Tn-binding legume lectins. **The International Journal of Biochemistry & Cell Biology**, v. 59, p. 103-110, 2015.

SPRINGER, G.F. Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. **Journal of molecular medicine**, v. 75, p. 594-602, 1997.

STILLMARK, H. **Über Ricin ein giftiges Ferment aus den Samen von Ricinus communis L. und einige anderen Euphorbiaceen**. 1888. Tese (Doutorado) - Pró-Reitoria de Pesquisa e Pós-Graduação, Universidade de Dorpat, Estônia, 1888.

SUVARNA, G.; SHARMA, B. B. An analogy of Canavalia lectins by molecular phylogenetic characterization. **Journal of Proteins and Proteomics**, v. 11, n. 3, p. 193–203, 2020

TSANEVA, M.; VAN DAMME, E.J.M. 130 years of Plant Lectin Research. **Glycoconjugate journal**, v. 37, n. 5, p. 533–551, 2020.

VAN DAMME, E. J. M. 35 years in plant lectin research: a journey from basic science to applications in agriculture and medicine. **Glycoconjugate Journal**, v. 39, n. 1, p. 83-97, 2022.

VAN DAMME, E.J.M.; LANNOO, N.; PEUMANS, W.J.: Plant lectins. **Advances in Botanical Research**, v. 48, p. 107-209, 2008.

VAN DAMME, E.J.M., PEUMANS, W.J., BARRE, A., ROUGÉ, P. Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. **Critical Reviews in Plant Sciences**, v. 17, n. 6, p. 575-692, 1998.

VAN HOLLE, S.; VAN DAMME, E. J. M. Messages from the past: New insights in plant lectin evolution. **Frontiers in Plant Science**, v. 10, n. January, p. 1–14, 2019.

VASCONCELOS, M. A.; ALVES, A. C.; CARNEIRO, R. F.; DIAS, A. H. S.; MARTINS, F. W. V.; CAJAZEIRAS, J. B.; NAGANO, C. S.; TEIXEIRA, E. H.; NASCIMENTO, K. S.; CAVADA, B. S. Purification and primary structure of a novel mannose-specific lectin from *Centrolobium microchaete* Mart seeds. **International journal of biological macromolecules**, v. 81, p. 600-607, 2015.

VERLI, H. **Bioinformática: da biologia à flexibilidade molecular**. 2014

WALLACE, A. C.; LASKOWSKI, R. A.; THORNTON, J. M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. **Protein engineering, design and selection**, v. 8, n. 2, p. 127-134, 1995.

WEBB, B.; SALI, A. Comparative protein structure modeling using MODELLER. **Current protocols in bioinformatics**, v. 54, n. 1, p. 1-5.6. 37, 2016.

YOUNG, N.M.; JOHNSTON, R.A.; WATSON, D.C. The amino acid sequence of peanut agglutinin. **European journal of biochemistry**, v. 196, n. 3, p. 631-637, 1991.

YURIEV, E.; HOLIEN, J.; RAMSLAND, P. A. Improvements, trends, and new ideas in molecular docking: 2012-2013 in review. **Journal of Molecular Recognition**, v. 28, n. 10, p. 581–604, 2015.

ZHAO, Y.Y.; TAKAHASHI, M.; GU, J.G.; MIYOSHI, E.; MATSUMOTO, A.; KITAZUME, S. Functional roles of N-glycans in cell signaling and cell adhesion in cancer. **Cancer science**, v. 99, n. 7, p. 1304-1310, 2008.