

UNIVERSIDADE FEDERAL DE PELOTAS
Centro de Desenvolvimento Tecnológico
Curso de Graduação em Biotecnologia

Trabalho de Conclusão de Curso



**Potencial biotecnológico de lectinas da família Musaceae: uma abordagem de
bioinformática estrutural**

Ana Claudia Oliveira de Freitas

Pelotas, 2021

Ana Claudia Oliveira de Freitas

Potencial biotecnológico de lectinas da família Musaceae: uma abordagem de bioinformática estrutural

Trabalho acadêmico apresentado ao Curso de Graduação em Biotecnologia da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Bacharel em Biotecnologia.

Orientador: Prof. Dr. Luciano da Silva Pinto

Pelotas, 2021

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

F863p Freitas, Ana Claudia Oliveira de

Potencial biotecnológico de lectinas da família
Musaceae : uma abordagem de bioinformática estrutural /
Ana Claudia Oliveira de Freitas ; Luciano da Silva Pinto,
orientador. — Pelotas, 2021.

65 f. : il.

Trabalho de Conclusão de Curso (Bacharelado em
Biotecnologia) — Centro de Desenvolvimento Tecnológico,
Universidade Federal de Pelotas, 2021.

1. Banlec. 2. Clustal. 3. I-Tasser. 4. Docking molecular.
5. Banana. 6. Bioinformática. I. Pinto, Luciano da Silva,
orient. II. Título.

CDD : 634.772

Ana Claudia Oliveira de Freitas

Potencial biotecnológico de lectinas da família Musaceae: uma abordagem de bioinformática estrutural

Trabalho de Conclusão de Curso aprovado, como requisito parcial, para obtenção do Grau de Bacharel em Biotecnologia, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas.

Data da Defesa: 31 de maio de 2021

Banca Examinadora:

Prof. Dr. Luciano da Silva Pinto (Orientador)

Universidade Federal de Pelotas

Profa. Dra Tatiane Santi Gadelha

Universidade Federal da Paraíba

Me. Rafael Woloski

Universidade Federal de Pelotas

Ma. Laura Junqueira de Camargo

Universidade Federal do Rio Grande do Sul

Agradecimentos

Fazer um trabalho de conclusão de curso no meio de uma pandemia não foi uma tarefa fácil. Enfrentar a perda de entes queridos, a minha ansiedade e todos os desafios de aprender novas técnicas foram imensamente desafiantes e ao mesmo tempo desmotivantes. Primeiramente gostaria de agradecer a minha mãe, sem ela eu não seria absolutamente nada e, certamente, não estaria aqui. Obrigada por me sustentar nos momentos mais difíceis e mostrar que não tem outro caminho senão seguir em frente, eu te amo muito.

Em minha trajetória acadêmica e no crescimento como pesquisadora gostaria de agradecer profundamente a Laura Camargo, Juliana Oliveira e Isabel Vighi. Três mulheres completamente diferentes, em fases de vida diferentes que me ensinaram muito sobre persistência, educação, arcar com as consequências, trabalho em grupo e companheirismo. Vocês foram essenciais não só para minha carreira acadêmica, mas também para minha vida com um todo.

Gostaria de agradecer em especial a Alessandra Neis, sem você nada disso seria possível. Obrigada por sempre estar disponível, por sempre mostrar que sou capaz e principalmente me guiar quando mais precisei e menos acreditei em mim. Você é incrível, me ensinou coisas que levarei para o resto da vida e muitos protocolos de bioinformática que tem até meus suspiros de desespero (hahahah). Brincadeiras à parte, você foi o meu pilar acadêmico, obrigada por me ensinar, por ter paciência com as minhas imensas dúvidas, obrigada por me fazer voar sozinha.

O meu professor orientador Luciano da Silva Pinto também merece um parágrafo dedicado a ele. Obrigada por todo o conhecimento passado, mesmo quando eu não entendia a importância de cada puxão de orelha eles tem grande contribuição na profissional que me tornei. Obrigada por sempre estar disponível para me ajudar a entender as mirabolâncias que eu criava fazendo um trabalho EaD, principalmente por me guiar durante todo o processo.

Por fim, não posso deixar de agradecer as minhas amigas, a minha equipe de trabalho no BioPro, vocês animaram meus dias e tudo isso me deixará muita saudade.

Resumo

FREITAS, Ana Claudia. **Potencial biotecnológico de lectinas da família Musaceae: uma abordagem de bioinformática estrutural**. 2021. 65f. Trabalho de Conclusão de Curso – Curso de graduação em Biotecnologia. Universidade Federal de Pelotas, Pelotas.

Lectinas são um grupo de proteínas de origem não imune que se ligam reversivelmente a diversos tipos de carboidratos. A importância do estudo das lectinas reside em seu amplo potencial de aplicação na área da biotecnologia. No entanto, existe uma grande escassez de informações estruturais, genômicas e funcionais acerca de diversas lectinas, principalmente nas espécies da família *Musaceae*. A maior parte das informações estão concentradas no gênero *Musa* e, pouco sabe-se a respeito de lectina de outros gêneros dessa família que podem apresentar grande valor biotecnológico. Assim, o objetivo desse trabalho foi lectinas da família *Musaceae* com potencial biotecnológico através de técnicas de bioinformática, e também prever sua estrutura e função a fim de elucidar um pouco mais o funcionamento dessas lectinas. Para isso foram utilizados softwares de bioinformática como I-TASSER, Clustal, BLAST, Autodock Vina, Plip, GalaxyRefine, PDBSum. Ao identificar e analisar as sequências de lectinas da família *Musaceae* pôde-se observar grande semelhança estrutural com alterações pontuais nos sítios de ligações, com energias de ligação e interação fina a carboidratos semelhantes aos do modelo *Musa acuminata* (Banlec). Assim, podemos concluir que as lectinas da família *Musaceae* apresentam um potencial biotecnológico *in vitro*, *in vivo* e *in silico* ainda inexplorado, apesar de ainda precisar de estudos para melhor elucidar as suas aplicações e funções esta análise *in silico* gerou dados que poderão servir de norte para futuras pesquisas de prospecção dessas lectinas.

Palavras chave: Banlec, CLUSTAL, I-TASSER, Docking molecular, Banana, Bioinformática.

Abstract

FREITAS, Ana Claudia. **Biotechnological potential of lectins from the Musaceae family: Structural and functional analysis** 2021. 65f. Course Conclusion Paper - Undergraduate course in Biotechnology. Federal University of Pelotas, Pelotas

Lectins are a group of non-immune proteins origin that reversibly bind to different types of carbohydrates. The importance of studying lectins lies in their wide potential for application in the biotechnology field. However, there is a great absence of information about, genomics and several lectins, mainly in the *Musaceae* species family. Most of the information is concentrated in the genus *Musa* and not enough is known regarding the lectin of other genres of this family that can have high biotechnological interest. Thereby, the objective of this research is to use bioinformatics tools to identify lectins from the *Musaceae* family with biotechnological potential by predicting their structure and function. For this, bioinformatics software such as I-TASSER, Clustal, BLAST, Autodock Vina, Plip, GalaxyRefine, PDBSum were used. After identifying the sequences of lectins of the *Musaceae* family, it was possible to observe a great structural similarity with specific changes in the binding sites, binding energies and fine interaction with carbohydrates similar to those of the *Musa acuminata* model (Banlec). Therefore, it can be concluded that the *Musaceae lectins* family may present a still unexplored biotechnological potential and further studies are needed to better elucidate their applications and functions.

Keywords: Banlec, CLUSTAL, I-TASSER, Molecular Docking, Banana, Bioinformatics.

Lista de Figuras

Figura 1. Ilustração da diversidade funcional das lectinas. Adaptado de VASTA et. al. (2012). Criado com BioRender.com.	15
Figura 2 Organização visual da família Musaceae dividida em três diferentes gêneros Ensete, Musa e Musella algumas das principais espécies representantes desses gêneros. 19	
Figura 3 Estrutura β -prisma das lectinas vegetais representada pela lectina de banana da Malásia PDB ID: 7KMU Imagem criada utilizando Mol* (D. Sehnal, A.S. Rose, J. Kovca, S.K. Burley, S. Velankar (2018)).....	27
Figura 4 Estrutura tridimensional de <i>M. balbisiana</i> (THU46912.1) e obtida através do software I-TASSER, sobreposta a lectina de <i>M. acuminata</i> referência (PDB ID 3MIT)	45
Figura 5 Alinhamento feito no EMBOSS Needle entre as lectinas <i>M. balbisiana</i> (THU46912.1) e <i>M. acuminata</i> (PDB ID 3MIT).....	46
Figura 6 Sobreposição da <i>M. acuminata</i> referência (PDB ID 3MIT) em azul com os seus ligantes em preto com o seu <i>redocking</i> em vermelho com os ligantes em amarelo.	50
Figura 7 Sobreposição do resultado do <i>docking</i> , utilizando a manose como ligante, para a análise da disposição dos ligantes.	51

Lista de Tabelas

Tabela 1 Seleção das 22 sequências que apresentaram maior similaridade com a sequência <i>M. acuminata</i> (PDB ID 1X1v) o software BLASTp.	38
Tabela 2 Mutações nos locais de ligação ao carboidrato de lectinas da família Musaceae que não possuem estrutura no PDB. As modificações dos resíduos estão assinaladas em negrito e sublinhadas.	40
Tabela 3 Valores de c-score e primeiro PDB hit gerado para as lectinas da família Musaceae através do software I-TASSER.	41
Tabela 4 Sítios de ligações à manose das lectinas da Família Musaceae fornecidos pelo módulo COACH do software I-TASSER.	42
Tabela 5 Valores de RMSD gerados pela sobreposição entre os modelos tridimensionais gerados pelo I-TASSER e a lectina de <i>M. acuminata</i> (PDB ID: 3MIT)	43
Tabela 6 Valores obtidos a partir do gráfico de Ramachandran antes e depois do refinamento das lectinas selecionadas (THU46912.1; RWV92385.1; XP_018673706.1).	47
Tabela 7 Energias de ligação obtidas através de docking molecular para as três lectinas escolhidas para análise (THU46912.1; RWV92385.1; XP_018673706.1) ...	48
Tabela 8 Resíduos de interação com a manose obtidos através do <i>docking</i> molecular para as lectinas selecionadas para a análise (THU46912.1; RWV92385.1; XP_018673706.1).	49

Lista de Abreviaturas e Siglas

Banlec – Lectina de *Musa acuminata*

GdnHCl – Cloridrato de guanidina

SGNPs - Nanopartículas de ouro imobilizadas com glicídios

JRL's – Família das lectinas relacionadas a jacalina

NMR – Ressonância nuclear magnética

mJRLs – Lectinas relacionada a jacalina com reconhecimento a manose

DBL – Lectina da banana del monte

kDa - Kilodaltons

CDR – Domínio de reconhecimento a carboidratos

IgG – Imunoglobulina G

HIV – Vírus da imunodeficiência humana

YFV – Vírus da febre amarela

CHIKV – Vírus da Chikungunya

EV71 – Enterovírus 71

MDBK – linhagem celular de rim bovino

TMV – Vírus do mosaico do tabaco

ERK – Via de sinalização regulada por quinases externas

NFKB – Fator nuclear kappa B

RAS – proteínas oncogênicas derivadas de vírus tumorais que causam sarcoma em ratos

Padronizar as iniciais deixar todas em maiúsculas

BCL-2 – Família de proteínas reguladoras da apoptose

D - Aspartato

K - Lisina

Q - Glutamina

Gp120 – Glicoproteína 120 da superfície do vírus HIV

Gp41 – Glicoproteína 41 da superfície do vírus HIV

EnV – gene que codifica a proteína do envelope do HIV

RMSD – Distância entre aminoácidos correspondentes

IL - Interleucina

IFN - Interferon

T CD4/T CD8 – Linfócitos T

MALP – Lectina da polpa de *Musa acuminata*

EAC – *Ehrlich ascites carcinoma*

HT29 – Linhagem celular de câncer uterino humano

MCF-7- Linhagem celular de câncer de mama

HeLa – Linhagem celular de câncer cervical

K562 – Linhagem celular de leucemia

Sumário

Sumário

1. INTRODUÇÃO	12
2. REVISÃO BIBLIOGRÁFICA.....	14
2.1 LECTINAS	14
2.2 LECTINAS VEGETAIS:	16
2.3 LECTINAS RELACIONADAS ÀS JACALINAS (JRL's)	18
2.4 FAMÍLIA MUSACEAE:	19
2.5 LECTINAS DA FAMÍLIA MUSACEAE:	20
2.6 ATIVIDADES BIOLÓGICAS DAS LECTINAS DA FAMÍLIA MUSACEAE:.....	22
2.6.1 Atividade antiviral:.....	22
2.6.2 Atividade anticancerígena:.....	24
2.6.3 Isolamento da lectina de <i>Musa</i> :	25
2.7 ESTRUTURA DAS LECTINAS DE <i>MUSA</i> :	26
2.7.1 Lectina de <i>Musa acuminata</i> :	27
2.7.2 Lectina <i>Musa acuminata</i> (Banana del Monte):.....	28
2.7.3 Lectina de <i>Musa paradisiaca</i> :	28
2.8 ESTRATÉGIAS DE BIOINFORMÁTICA APLICADAS AS LECTINAS	29
2.8.1 Estrutura tridimensional e modelagem molecular	29
2.8.2 Refinamento e checagem de qualidade.....	31
2.8.3 <i>Docking</i> molecular	32
3. OBJETIVOS GERAIS:	35
3.1 OBJETIVOS ESPECÍFICOS:	35
4. MATERIAIS E MÉTODOS:	36
4.1 ANÁLISE DAS SEQUÊNCIAS:.....	36
4.2 MODELAGEM ESTRUTURAL E REFINAMENTO DAS SEQUÊNCIAS:	36
4.3 DOCKING MOLECULAR	37
5. RESULTADOS.....	38
5.1 ANÁLISES DAS SEQUÊNCIAS	38
5.2 MODELAGEM ESTRUTURAL E REFINAMENTO DAS SEQUÊNCIAS.....	41
5.3 DOCKING MOLECULAR	46
6. DISCUSSÃO	51
7. CONCLUSÃO.....	56
8. REFERÊNCIAS	57

1. Introdução

Lectinas são um grupo de proteínas de origem não imune que se ligam reversivelmente a diversos tipos de carboidratos (DAMME *et al.*, 1998). A importância do estudo das lectinas reside em seu amplo potencial de aplicação. Nesse contexto, as lectinas extraídas da banana são muito promissoras devido a sua gama de atividades já elucidadas, como: capacidade imunomoduladora, antiproliferativa e antiviral, inclusive sobre o vírus da imunodeficiência humana (HIV) (COVÉS-DATSON *et al.*, 2019; SRINIVAS; SHIVAMADHU; JAYARAMA, 2019).

Os gêneros *Musa* e *Ensete* pertencem a família *Musaceae* a qual engloba todas as espécies de bananas existentes, comestíveis e não comestíveis. Atualmente, a maioria dos estudos das lectinas dessa família se concentram no gênero *Musa*, com grande destaque para a espécie *Musa acuminata*.

A lectina com que mais se destaca nessa família é a Banlec, pois possui diversos estudos abordando atividades antiviral e antiproliferativa, além de possuir uma forma recombinante e mutada chamada de H84T que apresenta uma substituição de uma histidina por treonina tornando a proteína mais imunogênica e com menor mitogenicidade (COVÉS-DATSON *et al.*, 2019, 2020; SWANSON *et al.*, 2015). No entanto, com exceção da *Musa acuminata*, as lectinas dos gêneros *Ensete* e *Musa* não possuem estudos que as caracterizem e explorem as suas atividades biológicas, e pelas características e propriedades apresentadas pela BanLec essas lectinas podem apresentar atividades biológicas de grande interesse para a saúde humana e animal.

Para elucidar e analisar as características das lectinas negligenciadas da família *Musaceae*, a bioinformática se mostra uma ferramenta essencial na análise de dados genômicos e proteômicos, além de promover a redução de custos, tempo de pesquisa ainda disponibilizar uma grande quantidade de dados, que nos permitem avaliar previamente as suas características estruturais e as possíveis atividades biológicas. A maioria das estruturas disponíveis de lectinas da família *Musaceae* são inferidas por homologia, e a carência de informações interfere diretamente nas

análises das funções biológicas das lectinas, uma vez que a função está intimamente ligada à estrutura de uma proteína (ORENGO; TODD; THORNTON, 1999).

A falta de informações estruturais faz com que a modelagem estrutural se torne a maneira mais rápida para obter uma estrutura tridimensional e, assim, ser possível avaliar informações acerca dos domínios de reconhecimento a carboidratos e as interações proteína-carboidrato, que são fundamentais para a compreensão dos efeitos biológicos promovidos por essas proteínas (PINTO-JUNIOR *et al.*, 2017b).

Algumas das ferramentas mais utilizadas para investigar a função de proteínas *in silico* são os métodos de modelagem e *docking* molecular, pois permitem a obtenção da estrutura e análise dos sítios de ligação a carboidratos (CRDs - carbohydrate binding sites). Sabendo que as lectinas de banana possuem maior interação com carboidratos como manose, glicose e frutose é possível investigar se as mutações que ocorrem nos CRDs irão afetar a ligação entre proteína-ligante (CHEUNG; WONG; NG, 2009a; GOLDSTEIN *et al.*, 2001a; SINGH; DEVI; NG, 2014a).

Dessa forma, o objetivo desse trabalho é identificar diferentes lectinas da família *Musaceae* que apresentam potencial biotecnológico e, a partir de ferramentas de bioinformática, analisar a estrutura dessas lectinas e a interação fina com a manose, para que esses resultados possam contribuir em trabalhos futuros que explorem o potencial biotecnológico dessas proteínas.

2. Revisão bibliográfica

2.1 Lectinas

A primeira lectina foi relatada no ano de 1888 por Peter Hermann e Stillmark, como parte de sua tese, Stillmark isolou a ricina de mamona (*Ricinus communis*), descrevendo-a como aglutinante e extremamente tóxica (VAN DAMME, 2014). Em 1907, Lednsteiner apresentou evidências de lectinas não tóxicas e presumiu que tinham atividades análogas aos anticorpos devido a sua especificidade a eritrócitos. No entanto, sem atividades catalíticas nem origem imune descritas, não tinham características para se encaixar em nenhum grupo proteico já existente (VAN DAMME, 2014; XUILI, 2009). Dessa forma, para englobar o novo grupo descoberto, foi criado o termo lectina, derivado do latim *Legere*, que significa selecionar/escolher. O termo passou a representar todas as aglutininas de origem não imune com ligações específicas a glicídios (BOYD; SHAPLEIGH, 1954).

A partir de 1954, as lectinas são descritas como proteínas de origem não imune as quais possuem a capacidade de se ligar de forma reversível a açúcares isolados ou glicoconjugados, como glicoproteínas e glicolipídeos (CAGLIARI; KREMER; PINTO, 2018). Este grupo proteico possui domínios de reconhecimento de carboidratos (CDR - *carbohydrate binding domain*) de ligação específica, adotando o modelo chave-fechadura (KENNEDY *et al.*, 1995). Esses domínios são altamente conservados em cada tipo de lectina e a especificidade ocorre através das interações químicas das ligações (DE OLIVEIRA DIAS *et al.*, 2015). A interação entre lectinas e carboidrato é do tipo não-covalente e envolve ligações de hidrogênio, hidrofóbicas, eletrostáticas e interações de Van der Waals (WEIS; DRICKAMER, 1996).

As ligações com os carboidratos e a diversidade de estruturas moleculares fazem com que essas proteínas pertençam a um grande grupo heterogêneo e apresentem diversas funções em diferentes organismos (DAMME *et al.*, 1998). As lectinas que se encontram no citoplasma podem ser armazenadas em vacúolos; também podem ser excretadas para o meio extracelular; além da possibilidade de serem translocadas para o núcleo celular; reconhecem carboidratos de patógenos; podem se ligar entre si formando tetrâmeros e se ligar na célula que a produziu; e

também se ligarem a carboidratos da matriz celular; podem desencadear cascatas de sinalizações; e promover interação célula-célula e célula-matriz (VASTA *et al.*, 2012).

As atividades biológicas das lectinas são extremamente diversificadas, envolvendo processos, como endocitose, transporte intracelular, indução de apoptose em células tumorais, bloqueio da infecção do HIV e regulação da adesão e migração bacteriana (COVÉS-DATSON *et al.*, 2019; DE OLIVEIRA FIGUEIROA *et al.*, 2017; POIROUX *et al.*, 2017). Além disso, sabe-se que elas desempenham papéis importantes no sistema imunológico, reconhecendo carboidratos encontrados em patógenos ou se ligando àqueles nas células hospedeiras (Figura 1).

A alta diversidade possibilita diversas aplicações biotecnológicas em diversos campos, incluindo saúde humana e animal e também o agronegócio (DE OLIVEIRA DIAS *et al.*, 2015). A maior parte dos estudos com lectinas são de origem vegetal por serem de mais fácil obtenção, assim a maior parte das atividades biológicas relatadas provém de lectinas dessa origem (VAN DAMME, 2014).

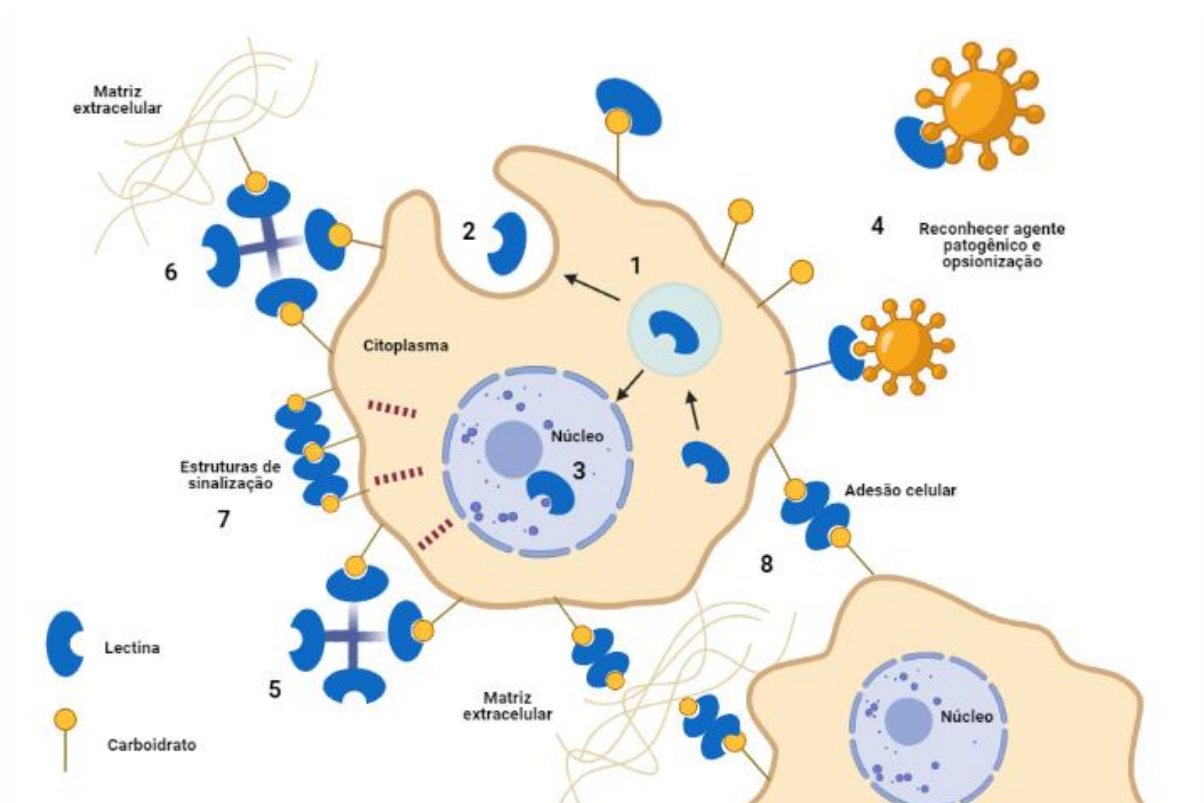


Figura 1 Ilustração da diversidade funcional das lectinas. 1 – As lectinas em vacúolos; 2 – lectinas excretadas para o meio extracelular; 3 – lectinas translocadas para o núcleo celular; 4 - interação carboidratos de patógenos; 5 – formação de tetrâmeros e ligação a célula; 6 – lectina ligada aos

carboidratos da matriz celular; 7 – Estrutura de cascatas de sinalizações; 8 – interação célula-célula e célula-matriz. Adaptado de (VASTA *et al.*, 2012). Criado com BioRender.com.

2.2 Lectinas vegetais:

Lectinas são identificadas em muitos grupos de plantas, tanto em monocotiledôneas quanto em dicotiledôneas, porém, a grande maioria é encontrada nas famílias *Leguminosae* e *Euphorbiaceae*. As proteínas se encontram amplamente distribuídas nas folhas, frutos (banana e tomate), caule, bulbo, sementes (trigo, soja, feijão), rizoma, floema. Além disso, possuem também, diferentes localizações celulares e propriedades moleculares (DAMME *et al.*, 1998; PEUMANS; VAN DAMME, 1994).

Desde a descoberta da lectina de planta por volta de 130 anos atrás, muitos relatos sobre sua estrutura, características e funções vêm sendo descrita na literatura. Uma das lectinas mais estudadas e pioneiras nas aplicações biotecnológicas foi a Concanavalina A, purificada a partir de sementes de feijão-de-porco (*Jack bean*) *Canavalia ensiformis* (CAVADA *et al.*, 2019). Ao passar dos anos os avanços nas áreas da bioquímica e biologia molecular contribuíram intrinsecamente para as análises estruturais, sequenciamento da proteína, clonagem baseada no cDNA, ampliando, assim, a fronteira do conhecimento e alcançando os pré-requisitos necessários para desvendar a estrutura tridimensional e seu impacto na ligação específica a carboidratos (PEUMANS; VAN DAMME, 1994; VAN DAMME, 2011).

Nas últimas duas décadas, a pesquisa teve o intuito de focar nas análises funcionais, com o objetivo de decifrar a importância fisiológica das lectinas na planta. A compreensão das atividades fisiológicas também permitiu o desenvolvimento de ferramentas biotecnológicas para aplicações biomédicas e agrônômicas (HIRABAYASHI, 2004; VAN DAMME, 2014). Lectinas de origem vegetal são muito estudadas devido a sua alta especificidade e aplicações variadas pois, apresentam regiões ricas em aminoácidos hidrofóbicos que permitem as suas ligações com as porções glicídicas (DE OLIVEIRA DIAS *et al.*, 2015).

A ligação específica de uma lectina a uma estrutura de carboidratos pode ser explorada em cromatografia de afinidade, onde estas se encontram imobilizadas na

coluna, permitindo assim selecionar proteínas glicosiladas de tecidos, extratos celulares bem como possibilita a separação e análise estrutural de glicoproteínas e oligossacarídeos. Além disso, a tecnologia de lectina *microarrays*, com lectinas presente nos chips, foi desenvolvida em uma ferramenta para análise de alto rendimento, por exemplo, análises de tumores com alterações na glicosilação (VAN DAMME, 2011, 2014).

A fim de prover um melhor entendimento das lectinas vegetais, diversas propostas foram sugeridas para classificação das mesmas. Porém, com o constante aumento de conhecimento acerca de sua estrutura e atividade, as definições para os grupos foram sendo adaptadas e se tornando cada vez mais amplas que as propostas anteriores (DIXON, 1981; GOLDSTEIN *et al.*, 1980).

Com uma análise abrangente das sequências, em 1998, um sistema foi criado e dividiu as proteínas em 7 famílias diferentes. Essa classificação foi baseada nas relações entre a sequência de aminoácidos e históricos evolutivos das lectinas, sendo organizadas em: lectinas de leguminosas, lectinas de monocotiledôneas ligantes a manose, lectinas ligantes à quitina contendo domínios heveínicos, proteínas inativadoras de ribossomos tipo 2, proteínas relacionadas a jacalina, lectinas de *Amaranthaceae* e lectinas do floema de *Cucurbitaceae* (DAMME *et al.*, 1998b).

Este sistema por um período se provou útil, porém, diversas atualizações foram realizadas a medida que mais informações se tornaram disponíveis. A partir de análises genômicas e transcriptômicas das sequências disponíveis, foi criado um modelo de organização em 12 famílias baseado na sequência dos polipeptídios e da estrutura de seus domínios de reconhecimento ao carboidrato (CDRs) (VAN DAMME, 2014).]

Para facilitar o estudo e entendimento das lectinas vegetais Lanoo e Van Damme detalharam 7 das principais famílias de lectinas, as separando em solúveis e associadas a membrana (LANNON; VAN DAMME, 2014). As lectinas ligadas à membrana incluem as quinases receptoras de lectinas (LecRK), já as proteínas solúveis incluem as amarantinas e calreticulina/calnexina, lectinas relacionadas com lectinas de euonymus (EUL), lectinas relacionadas às jacalinas (JRLs), lectinas relacionadas a nictaba e lectinas de ricina-B (LANNON; VAN DAMME, 2014).

2.3 Lectinas relacionadas às jacalinas (JRL's)

As lectinas relacionadas às Jacalina (JRL's) estão entre uma das famílias de lectinas mais estudadas. Elas apresentam um ou mais domínios que são homólogos ao da proteína jacalina, uma lectina extraída das sementes de jaca (*Jack fruit*). As JRL's podem ter afinidade a manose ou a galactose, possuindo assim uma subdivisão na família (LANNON; VAN DAMME, 2014). Estruturalmente, possuem quatro monômeros compostos de grandes cadeias α e pequenas cadeias β contendo um local único de ligação de açúcar. Esse conjunto proteico está relacionado nas plantas a resistência a doenças, vias de estresse abiótico, dano a insetos e estresses múltiplos (SONG *et al.*, 2014).

A família pode receber grande destaque pelas suas inúmeras atividades biológicas, pois exibe funções como: ligação específica ao tumor associado ao antígeno T, liga-se a regiões específicas do HIV causando sua inibição, estimula de forma potente e seletiva diferentes funções das células T, e ainda, apresentam uma capacidade de reconhecer especificamente a imunoglobulina A1 (IgA1) de soro humano. Analisando o que foi posto, muitas aplicações biomédicas surgem diante dessas funções, como por exemplos a detecção de tumores específicos auxiliando no diagnóstico e prognóstico de pacientes com cânceres, ser utilizada em cromatografias de afinidade de IgA1 e por fim, atuar em uma imunoestimulação seletiva (RAVAL *et al.*, 2004).

Mesmo com grande potencial biotecnológico, muitas lectinas classificadas como JRLs são pouco estudadas e podem apresentar um potencial ainda não explorado, como as lectinas da família *Musaceae*. Muito se sabe acerca da Banlec uma lectina extraída da *Musa acuminata*, pertencente ao gênero *Musa*, a qual apresenta diversas atividades antivirais e até mesmo anticancerígenas (COVÉS-DATSON *et al.*, 2019; SRINIVAS; SHIVAMADHU; JAYARAMA, 2019). Porém, sobre outras lectinas da mesma família ou, até do mesmo gênero, poucos relatos são encontrados na literatura.

2. Família Musaceae:

A família Musaceae é uma pequena família de plantas que abriga três diferentes gêneros: *Ensete*, *Musa* e *Musella* que são típicos de florestas tropicais e subtropicais. O gênero *Ensete* e *Musella* possui oito espécies distribuídos pela Ásia e África tropical. Por fim, o gênero *Musa* compreende a maioria das espécies e sua faixa de distribuição coincide com os *hotspots* asiáticos que apresentam cerca de 60 a 70 espécimes diferentes (JANSSENS *et al.*, 2016).

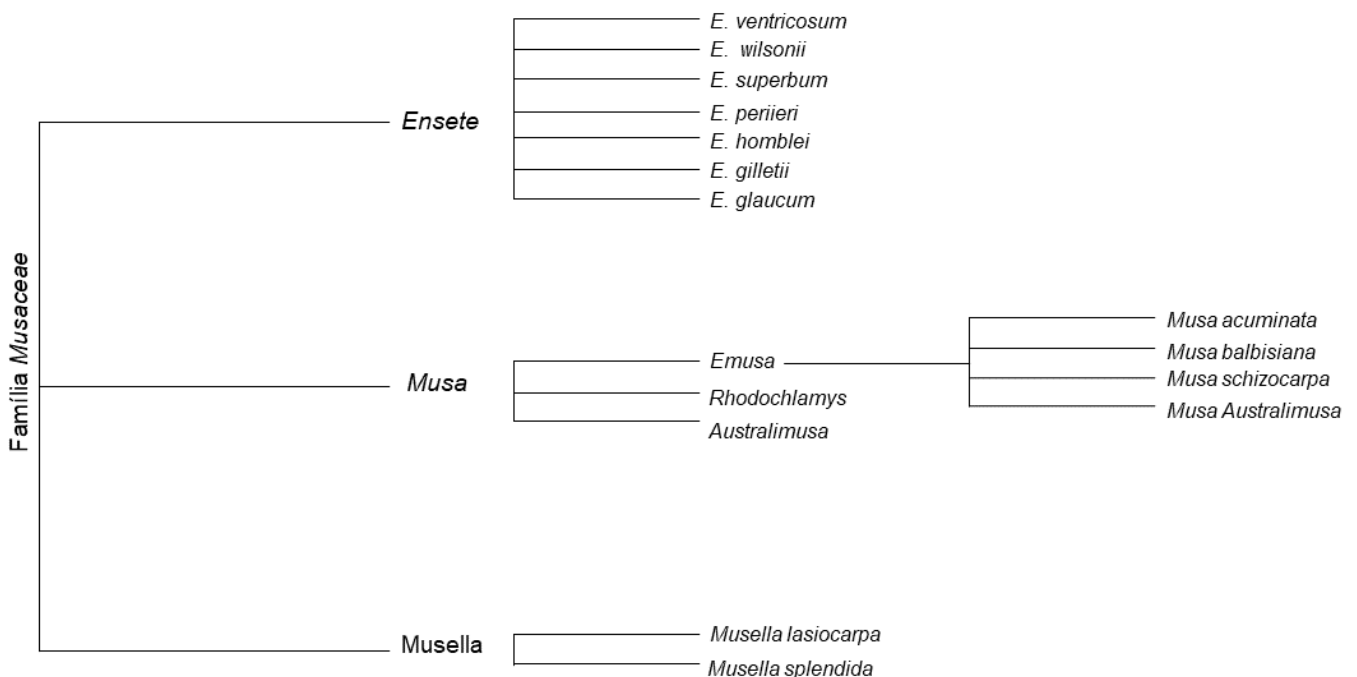


Figura 2 Organização visual da família *Musaceae* dividida em três diferentes gêneros *Ensete*, *Musa* e *Musella* algumas das principais espécies representantes desses gêneros.

A espécie que mais representa a família *Musaceae* são as bananas, uma das principais culturas de frutas tropicais, sendo distribuída em mais de 120 países com uma larga produção de 102 milhões de toneladas por ano. Sendo uma das primeiras culturas a serem domesticadas, há cerca de 7.000 anos, e possuem seus pontos de origem no Sudeste da asiático e no Pacífico Ocidental.

Atualmente é um alimento básico para diversas populações humanas e animais, sendo assim um importante recurso alimentar, apresentam um enorme papel socioeconômico e ecológico (PAUL *et al.*, 2017). Tradicionalmente, a banana tem sido usada para diversos fins medicinais. Por exemplo, para o tratamento de diarreia e úlceras gástricas. Com o passar dos anos, pesquisas vêm sendo desenvolvidas com

esse pseudofruto, avaliando sua atividade antibacteriana, antioxidante, antiviral e anticancerígena (D'HONT *et al.*, 2012; SINGH; DEVI; NG, 2014).

A maioria das bananas as quais são consideradas comestíveis pertencem a secção *eumusa* do gênero *Musa* (CHEESMAN, 1947). Como dito anteriormente o gênero *Musa* pertence à família Musaceae e suas espécies selvagens são divididas em: *eumusa*, *rhodochlamys*, *australimusa*. Ainda, dentro da secção *eumusa*, existem diferentes espécies que são divididas de acordo com seus antecedentes genéticos, *Musa acuminata*, *Musa balbisiana*, *Musa schizocarpa* e algumas espécies de *M. australiana*. A maioria das espécies comestíveis cultivadas foram originadas da hibridização entre as espécies selvagens de *M. acuminata* e *M. balbisiana*, que gerou diversos genótipos de bananas atualmente cultivadas (DAVEY *et al.*, 2013).

O gênero *Ensete* é composto por sete espécies de origem africana e asiáticas, *E. gillettii*, *E. glaucum*, *E. homblei*, *E. perrieri*, *E. superbum*, *E. ventricosum* e *E. wilsonii*. São chamadas de "falsas bananeiras" devido suas semelhanças morfológicas, mesmo que não pertençam ao gênero *Musa* (BIRMETA; NYBOM; BEKELE, 2004). O gênero *Ensete* representa 65% da produção agrícola do Sul da Etiópia, seu rizoma e o pseudocaule são altamente ricos em carboidratos e utilizados na alimentação para um quarto da população que habita o Sul e o Sudoeste do país. Dentre as espécies do gênero, a *Ensete ventricosum* (Welw.) *cheesman* é a mais importante socioeconomicamente em alguns países da África e Ásia (PLOETZ RANDY *et al.*, 2007).

2.5 Lectinas da família Musaceae:

As bananeiras apresentam propriedades medicinais em toda a sua estrutura, e a lectina de banana comestível já demonstrou capacidade de inibir a atividade da transcriptase reversa do HIV-1, suprimir a proliferação de células cancerígenas e também estimular a atividade de macrófagos (SINGH; DEVI; NG, 2014a). A lectina da banana foi descoberta durante uma pesquisa acerca da resposta de anticorpos humanos a vários alimentos, a partir dessa pesquisa, foi possível observar uma marcante ligação de IgG4 ao extrato de banana (*Musa paradisiaca*) (KOSHTE *et al.*, 1990).

Após a purificação do extrato foi detectada a presença de uma lectina e foi nomeada como BanLec-I. Foi relatado que a BanLec é uma proteína homodimérica que se liga a oligossacarídeos de manose e glicose, funcionando como um potente mitógeno de células T. Em posteriores estudos estruturais de Banlec *Musa acuminata*, que é muito semelhante a *M. paradisiaca*, confirmaram que a sua estrutura é composta por duas subunidades idênticas de 15 kDa, demonstrando especificidade em relação à manose e glicose pertencendo a família das JRLs. (KOSHTE *et al.*, 1992; MO *et al.*, 2001; PEUMANS *et al.*, 2000a).

Pouco se sabe a respeito das lectinas do gênero *Ensete*, não sendo encontrada nenhuma informação sobre a atividade biológica e a caracterização de nenhuma lectina desse grupo. Existem trabalhos que exploram apenas extratos totais de bananas da espécie *Ensete superbum* (DUTTA *et al.*, 1968; PANDA *et al.*, 2020a). Como são pesquisas com os extratos totais e não isolados, não se pode afirmar que a atividade antiviral desenvolvida é causada pelas lectinas, mesmo sabendo que elas possuem a atividade bem relatada como no caso da *M.acuminata* (COVÉS-DATSON *et al.*, 2019, 2020; DE CAMARGO *et al.*, 2020; LIU; LI; ZHANG, 2014; SRINIVAS; SHIVAMADHU; JAYARAMA, 2019; WONG; NG, 2006a).

Bananas da espécie *Ensete superbum* são comumente usadas por tribos indígenas da Índia para o tratamento do sarampo e para a catapora. No entanto, a única evidência científica sobre a ação antiviral da espécie é relatada em um artigo de 1968, em que o extrato final, obtido através da extração alcoólica das sementes da banana, foi utilizado em testes *in vitro* e *in vivo* contra uma cepa de varíola (DUTTA *et al.*, 1968). Os resultados dos testes *in vitro* e *in vivo* apresentou diminuição de pústulas e da mortalidade nos animais (DUTTA *et al.*, 1968).

De forma contrária, um estudo mais recente explorou a atividade antiviral do extrato total *E. superbum* contra o vírus da febre amarela (YFV), o vírus da Chikungunya (CHIKV) e o enterovírus 71 (EV71). Não foi detectada atividade antiviral para *E. superbum*, mesmo que outras cultivares isoladas no estudo apresentaram eficiência contra os vírus citados (PANDA *et al.*, 2020a).

2.6 Atividades biológicas das lectinas da família Musaceae:

2.6.1 Atividade antiviral:

Os vírus envelopados possuem glicoproteínas de superfície que mediam a ligação e a fusão com a membrana celular alvo. Essas proteínas constituem o primeiro contato com o hospedeiro, se tornando o alvo mais exposto que o sistema imunológico pode atacar (MISHRA *et al.*, 2019). Isto posto, diversas técnicas são estudadas para suprir as demandas de tratamento e prevenção para tais doenças. Uma das maneiras de tratar doenças virais atualmente em estudo é o uso de proteínas capazes de atingir os glicanos presentes na superfície das glicoproteínas do envelope, ou seja, lectinas (ACHARYA *et al.*, 2015; BALZARINI, 2007).

As lectinas vegetais possuem a capacidade de ligar glicanos presentes nas glicoproteínas do envelope viral, impedindo a penetração nas células hospedeiras. Além desses mecanismos, essas proteínas têm a capacidade de reticular os glicanos dos vírus, através de processos baseado na criação de ligações covalentes que impedem as interações com co-receptores, devido alterações causadas na estrutura tridimensional dessas moléculas após a interação da lectina. Analisando esses mecanismos de ação, é possível observar que a atividade antiviral das lectinas está ligada à sua especificidade de ligação ao carboidrato. (BARTON *et al.*, 2014).

Um estudo recente, mostra que a lectina de banana nativa (BanLec), extraída da polpa do fruto da *Musa acuminata*, ou seja, Banlec nativa, tem uma alta afinidade pelos glicanos que apresentam manose abundante. A atividade antiviral e virucida da BanLec foi analisada frente a dois vírus de interesse veterinário BVDV-1 e o BoHV-1, utilizando células de rim bovino (MDBK) linhagem suscetível ao vírus. O tratamento com a Banlec resultou em altos níveis de inibição contra o BVDV-1 (99,98%) e BoHV-1 (99,68%), sem afetar a viabilidade celular (DE CAMARGO *et al.*, 2020).

HOPPER *et al.*, (2017) demonstraram a capacidade da Banlec nativa de neutralizar o vírus do HIV através do reconhecimento de subconjunto de glicanos, com alto teor de manose, por meio de interações que abrangem os dois locais de ligação presentes em cada monômero BanLec os quais, anteriormente, eram considerados domínios de reconhecimento a carboidratos separados.

Porém, as lectinas extraídas da fruta *in natura*, apresentam uma grande desvantagem em possíveis aplicações clínicas, devido aos efeitos colaterais que podem ser gerados por sua atividade mitogênica. Se uma lectina com atividade mitogênica for usada topicamente em um microbicida anti-HIV, pode levar a inflamação, um aumento na transmissão e na replicação do HIV devido sua capacidade de ativar e recrutar células T (SWANSON *et al.*, 2015).

Para resolver esse grande impasse foi realizado a projeção da lectina isolada da *Musa Acuminata* (Banlec), de modo a eliminar sua mitogenicidade e manter sua potente atividade antiviral (SINGH; DEVI; NG, 2014a). Assim, desenvolveu-se uma versão recombinante que mantém a atividade antiviral e exclui a capacidade mitogênica.

Essa alteração foi possível a partir da substituição de um único aminoácido, histidina na posição 84 por treonina. H84T como foi nomeada, continua mantendo a sua ligação à manose e atividade anti-HIV (NAKAMURA-TSURUTA *et al.*, 2008; SWANSON *et al.*, 2015). A versão recombinante da Banlec H84T é altamente eficaz contra as versões epidêmicas da influenza aviária *in vitro* e, também, contra os casos mais letais causados através da infecção pelo vírus influenza *in vivo*.

A H84T é considerada um possível medicamento de amplo espectro pois, inibe a replicação de todas as cepas resistentes e sensíveis ao oseltamivir e de todas as cepas de influenza A e B, se destacando como um promissor medicamento empírico para a doença, que será extremamente necessária no caso de um novo vírus da gripe de potencial pandêmico ou aumento da resistência aos medicamentos que inibidores da neuraminidase (COVÉS-DATSON *et al.*, 2020). A H84T, foi utilizada, em estudo, contra ebola vírus (EBOV) e foi demonstrada a capacidade de inibir a replicação do vírus em culturas de células. A lectina atua principalmente na inibição da entrada da partícula viral nas células e, também, no processo de transcrição das proteínas virais (COVÉS-DATSON *et al.*, 2019).

Na área de biotecnologia vegetal, a ação de uma versão engenheirada da lectina de *Musa paradisiaca* denominada BanLec-1, quando inserida no genoma do tabaco, através da técnica de transgenia, resulta na diminuição do tamanho foliar e também mais resistência à infecção pelo vírus do mosaico do tabaco (TMV). A ação

antiviral ocorre devido a capacidade da aBanLec-1 de se ligar diretamente à proteína do capsídeo do vírus e inibir a infecção. (LIU; LI; ZHANG, 2014).

2.6.2 Atividade anticancerígena:

Inicialmente as lectinas eram utilizadas como uma ferramenta simples para diferenciar tumores benignos de malignos e, também, para identificar o grau de glicosilação associada à metástase. Dentro do contexto do câncer, o balanço entre apoptose e autofagia podem levar conjuntamente ao final da patologia. Atualmente, algumas lectinas possuem suas atividades antitumorais bem elucidadas direcionando a morte celular programada (FU *et al.*, 2011).

As lectinas possuem um reconhecimento seletivo das moléculas que apresentam açúcares livres, algumas são capazes de controlar e induzir, através da regulação de cascatas de sinalização como ERK, RAS e NF-kB (vias de sinalização), a divisão celular, proliferação, citotoxicidade e, até mesmo, a apoptose e conseqüentemente bloquear a ação de agentes cancerígenos nos tecidos alvos e, dessa forma, impedir o desenvolvimento do tumor (SRINIVAS; SHIVAMADHU; JAYARAMA, 2019).

Para induzir a apoptose as lectinas interagem com receptores na membrana plasmática, são internalizadas e direcionadas para a mitocôndria, gerando ROS que ativa a via de morte celular do citocromo c. Conjuntamente, a ROS realiza a regulação negativa de moléculas antiapoptóticas como a BCL-2. Ademais, as lectinas vegetais quando internalizadas nas células, podem promover aglutinação e agregação de células tumorigênicas, impedindo o desenvolvimento e dispersão dessas células pelo corpo do paciente (BHUTIA *et al.*, 2019). Também, possuem capacidade de se ligar a superfície de células do sistema imune desencadeando a liberação de citocinas com atividades tumoricida em células cancerosas (BHUTIA; MALLICK; MAITI, 2009; SOUZA *et al.*, 2013). Ainda, as lectinas podem se ligar aos ribossomos, apresentando uma toxicidade ao inibir a síntese de proteínas e impedir o desenvolvimento exacerbado das células cancerígenas e levando a apoptose das mesmas (DE MEJÍA; PRISECARU, 2005a; HAMID *et al.*, 2013).

A lectina de Banlec, extraída da poupa madura *Musa acuminata* apresentou propriedade imunomoduladora quando administrada oralmente em camundongos, essa atividade pode potencialmente afetar a resposta do hospedeiro a infecções e doenças imunológicas e ao câncer. Isso ocorre, pois, a lectina revelou o aumento de interleucina 10 (IL-10), interleucina 7 (IL-17), interferon (INF) e linfócito T CD4⁺ e redução de interferon (IFN), interleucina 6 (IL-6) e linfócito T CD8⁺ e nos animais modelos (SANSONE *et al.*, 2016). Basicamente, as interleucinas e o interferon são sinais químicos que auxiliam o sistema imune ao combate ao câncer (WHITTINGTON; FAULDS, 1993).

Outra variação da lectina de banana chamada de MALP (Lectina de *Musa acuminata*), extraída do floema da planta, demonstra um efeito citotóxico promissor contra linhagens de cânceres humanos. A partir desse estudo deu seguimento para o modelo *in vivo* em camundongos *Ehrlich Ascites Carcinoma* (EAC) a fim de analisar sua atividade antitumoral. Através de análises histológicas, foi possível observar que o tratamento com a lectina de banana apresentou atividade antiproliferativa nas linhagens de câncer de colo de útero humano (HT29), câncer de mama (MCF-7), câncer cervical (HeLa) e leucemia (K562), com inibição do crescimento do tumor. O mesmo estudo também verificou que a MALP prolongou a vida útil dos camundongos portadores de EAC sem apresentar efeitos colaterais nos controles. Por fim, foi possível avaliar que a lectina é capaz de induzir a apoptose nas células de carcinoma Ehrlich Ascites (SRINIVAS; SHIVAMADHU; JAYARAMA, 2019).

2.6.3 Isolamento da lectina de *Musa*:

O processo de purificação e isolamento da proteína é uma etapa essencial, pois interfere diretamente no enovelamento e conformação da molécula, e assim, pode alterar a estrutura tridimensional e gerar grande impacto nas atividades biológicas da lectina. Isso ocorre, pois, a interação carboidrato-proteína está intimamente ligada à estrutura tridimensional (3D) dos sítios de ligação (HEGYI; GERSTEIN, 1999).

Originalmente para o isolamento dessas proteínas eram realizadas por métodos de cromatografia de afinidade combinada com ácido acético para extração da lectina da polpa de bananas maduras. No entanto, esse método possui desvantagens como baixo rendimento e a degradação da lectina (WEARNE; WINTER; GOLDSTEIN, 2013). A fim de aprimorar a extração, surgiu um novo processo para a

purificação que consiste na utilização de nanopartículas de ouro imobilizadas com açúcar (SGNPs). Essa nova técnica aumentou o rendimento em quase 100x e que proporciona um produto final de alta pureza. (SACHIKO NAKAMURA-TSURUTA, 2009).

Diversos protocolos visam aprimorar o isolamento e purificação dessa molécula, pois há grande interesse na obtenção de uma estrutura na conformação ideal produzida por um processo pouco oneroso (R. RICHARD BUEGESS; MURRAY P. DEUTSCHER, 2009). Muitos compostos químicos são estudados e, acredita-se que o cloridrato de guanidina (GdnHCl) seja um ideal desnaturante químico para reação de desdobração de proteínas.

A sequência de eventos de desdobração da estrutura dimérica da lectina de banana (Banlec), induzida pelo cloridrato de guanidina (GdnHCl), gerou dois estados de desdobração distintos. Sendo, uma *coil* aleatória usual e o outro representa uma nova espécie, com eluição, comportamento e estrutura semelhantes ao monômero dissociado, mas sem mostrar estrutura secundária regular. Revelando a sutileza das características estruturais dos estados desdobrados do Banlec no GdnHCl (GHOSH; MANDAL, 2014).

2.7 Estrutura das lectinas de *Musa*:

O estudo das estruturas de lectinas é extremamente importante pois auxilia o entendimento da função proteica e suas possíveis aplicações, ainda, é importante ressaltar que a estrutura íntegra de uma lectina é essencial para sua atividade biológica correta. A estrutura pode ser obtida experimentalmente através das metodologias de cristalografia de Raios-X ou análise NMR. As estruturas β -prisma estão presentes na maioria das lectinas vegetais, se tornando a conformação mais comum dessas proteínas. A estrutura β -prisma compreende três folhas β repetidas na qual as três são reunidas em torno de um eixo pseudo-tríplo, resultando em uma estrutura de prisma triangular (FUJIMOTO; TATENO; HIRABAYASHI, 2014).

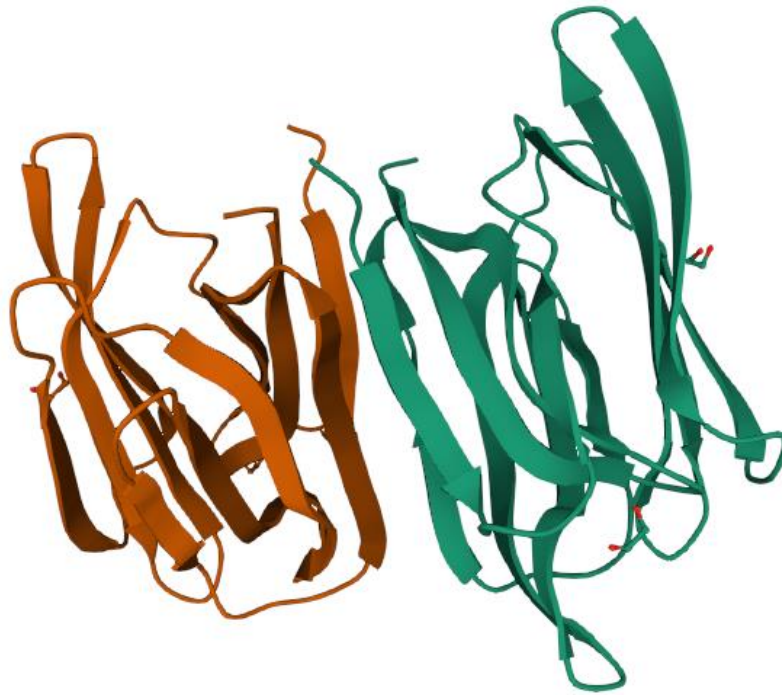


Figura 3 Estrutura β -prisma das lectinas vegetais sendo representada pela lectina de banana da Malásia PDB ID: 7KMU Imagem criada utilizando Mol* (D. Sehnal, A.S. Rose, J. Kovca, S.K. Burley, S. Velankar (2018).

A lectina da semente de jaca, a jacalina, foi a primeira da família das JRLs a ser estudada por cristalografia Raios-X, e sua estrutura revelou uma nova dobra nomeada de dobra β -prisma-I presente em cada uma das subunidades, formada por chaves gregas em ambas as cadeias, estrutura característica da família (SINGH; DEVI; NG, 2014a).

É possível observar que as regiões de folha beta são altamente conservadas principalmente na primeira e terceira folha, porém os loops são extremamente variáveis o que é consistente no papel de diversificar as funcionalidades das lectinas da família (SINGH; DEVI; NG, 2014b). Os sítios de ligações a carboidratos são encontrados nos loops sendo de extrema importância para a atividade da proteína (SHARMA ALOK; VIJAYAN MAMANNAMANA, 2011).

2.7.1 Lectina de *Musa acuminata* - Banlec

A Banlec é uma proteína dimérica composta por duas subunidades de 15 kDa e contém uma sequência de 141 aminoácidos. A clonagem molecular e a da estrutura proteica, revelaram que a Banlec possui similaridade estrutural com as outras lectinas da família das JRLs (MEAGHER *et al.*, 2005).

Geralmente a interação entre subunidades proteicas ocorre por interação hidrofóbica entre as moléculas, porém, no caso da Banlec a interação ocorre através de ligações de hidrogênio juntamente com pontes de água. Essa organização estrutural faz com que a lectina de banana seja extremamente estável, somente sofrendo desnaturação em altas concentrações de agentes caotrópicos (GUPTA; SINHA; SUROLIA, 2008).

Os sítios de ligação a carboidratos foram analisados e foi demonstrado que a estrutura da Banlec possui um segundo local de ligação a açúcar, observado pela primeira vez nessa estrutura. O segundo sítio de ligação apresenta resíduos comuns a outras lectinas dessa família, evidenciando um novo grupo de lectinas relacionadas as jacalinas específicas para manose (mJRLs). A partir disso, podemos dizer que a lectina possui dois sítios de ligação a manose por subunidade da lectina de banana (PEUMANS *et al.*, 2000). Além da ligação com a manose, outros estudos também comprovaram que a BanLec tem afinidade por glicose (GOLDSTEIN, 2002a; GOLDSTEIN *et al.*, 2001).

2.7.2 Lectina *Musa acuminata* (Banana Del Monte):

A estrutura da lectina extraída da polpa madura das bananas Del Monte (DBL) é semelhante às lectinas de outras espécies de *Musa*, exceto pela exclusão do resíduo de glicina N-terminal na lectina de banana Del Monte. A DBL possui a massa molecular (30 kDa) semelhante aos (27–30 kDa) de lectinas de outras espécies de bananas. Seu maior diferencial se dá pela capacidade de reconhecimento e ligação aos açúcares, frutose e glicosamina (CHEUNG; WONG; NG, 2009; KOSHTE *et al.*, 1990; PEUMANS *et al.*, 2000a; WONG; NG, 2006b). A perda do aminoácido N-terminal, em comparação com outras lectinas de banana, não afeta negativamente suas atividades biológicas, significando que este aminoácido não é essencial para as suas atividades. Isto evidencia que a lectina de banana se assemelha a outras lectinas monocotiledôneas na sequência N-terminal (CHEUNG; WONG; NG, 2009b).

2.7.3 Lectina de *Musa paradisiaca*:

A *Musa paradisiaca* é oriunda do cruzamento direto entre as variantes da *M. acuminata* e *M. balbisiana*. Como em outras lectinas, a lectina extraída da *M. paradisiaca* (Banana da terra) possui uma cadeia polipeptídica com 141 aminoácidos,

sendo as regiões 14-17 e 129-133 envolvidas na ligação de metil- α -D-manosídeo com interações semelhantes a outras lectinas descritas. Os loops são correspondentes ao segundo local de ligação 57-61 e 34-38. Surpreendentemente, as geometrias dos dois locais de ligação e a natureza das interações com o açúcar, são quase idênticas. Ao contrário das outras, a lectina de banana da terra tem dois sítios de ligação primários idênticos, um situado na chave grega I e o outro na chave grega II (D D SINGH *et al.*, 2015).

2.8 Estratégias de bioinformática aplicadas as lectinas

2.8.1 Estrutura tridimensional e modelagem molecular

A determinação da estrutura tridimensional de proteínas é de extrema importância para biotecnologia, pois permite a compreensão das funções biológicas e dos mecanismos moleculares das doenças (ADIYAMAN; MCGUFFIN, 2019). Para isso existem métodos experimentais bem estabelecidos para determinar a conformação 3D da proteína com alta precisão, a exemplo da cristalografia de Raios-X, ressonância nuclear magnética (NMR) e a microscopia crio eletrônica, porém não conseguem corresponder ao ritmo de novos dados genéticos, devido seus processos serem onerosos e trabalhosos.

Nesse cenário, a modelagem *in silico* é muito mais rápida e econômica e ajuda a preencher a lacuna entre as sequências conhecidas e as estruturas disponíveis (NASCIMENTO *et al.*, 2018). Sequências de proteínas, como lectinas, que não possuem estruturas modeladas nos bancos de dados podem ter sua estrutura predita através de softwares. Os *softwares* realizam essa tarefa através de diferentes metodologias, as predições foram categorizadas em comparativas e *ab initio*. Os métodos comparativos englobam duas diferentes metodologias, a homologia e o *threading* (LESK, 2014).

A modelagem por homologia é baseada no conceito de evolução molecular, partindo do princípio de similaridade da estrutura primária da proteína alvo e de proteínas-molde homólogas com estruturas e sequência conhecidas. Para ser considerado um modelo com qualidade boa suficiente para inferir a sua função é necessário apresentar ao menos 40% de resíduos idênticos a um homólogo de

estrutura conhecida, assim, quanto maior a porcentagem maior à similaridade entre a proteína desconhecida e o modelo (ALTSCHUL *et al.*, 1990; HILLISCH; PINEDA; HILGENFELD, 2004; ORENCO; TODD; THORNTON, 1999; SANTOS FILHO; BICCA DE ALENCASTRO, 2003).

Alguns *softwares* úteis para a modelagem de proteínas são o SWISS-MODEL (<https://www.expasy.org/resources/swiss-model-workspace>) e MODELLER (<https://salilab.org/modeller/>) (BIENERT *et al.*, 2017; GUEx; PEITSCH; SCHWEDE, 2009; PEITSCH, 1997; WATERHOUSE *et al.*, 2018). O *threading* é uma metodologia baseada no reconhecimento de padrões de enovelamento. Para a procura dos padrões de enovelamento é utilizado a biblioteca PDB e também dobramentos hipotéticos. A ideia básica é construir muitos modelos aproximados da proteína de interesse com base em cada uma das dobras conhecidas usando alinhamentos de sequências de proteínas conhecidas e desconhecidas (LESK, 2014).

A modelagem por homologia se concentra em um conjunto de alinhamentos de sequências, tendo como objetivo um modelo muito detalhado; já o *threading* explora muitos alinhamentos possíveis e lida com modelos aproximados; e podem às vezes revelar relações mais distantes evolutivamente do que métodos puramente baseados em sequência (LESK, 2014). Um dos softwares mais utilizados com essa metodologia é o I-TASSER (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>), sendo um dos melhores avaliados nas avaliações da CASP (*Critical Assessment of protein Structure Prediction*) (ZHANG, 2008).

A modelagem *ab initio*, também chamada de denovo, baseia-se nas propriedades físico-químicas conhecidas de cada aminoácido para a construção de funções de energia (LESK, 2014). Os métodos *ab initio* mostram-se importantes quando a similaridade ou homologia é baixa, restringindo o uso e a qualidade das outras abordagens. O método utiliza algoritmos e conhecimento teórico sobre a proteômica, bioquímica avançada e linguagem computacional. Isto posto, podemos concluir que é uma modelagem complexa e demorada, e ainda é geralmente limitada pelo tamanho da proteína, não sendo indicado para moléculas maiores que 200 aminoácidos (BAKER; SALI, 2001; BONNEAU; BAKER, 2001).

Para a predição *ab initio*, foram desenvolvidos softwares online considerados semi *ab initio*, pois utilizam estatísticas sobre sequências pequenas de aminoácidos para a predição, como softwares online que possibilitam a predição *ab initio*, como

QUARK (<https://zhanglab.ccmb.med.umich.edu/QUARK/>) e Rosetta (<https://web.archive.org/web/20150819163428/http://www.rosetta.org/>).

Para obtenção da estrutura de lectinas de banana a maioria dos estudos utiliza o método de cristalografia de raios-X, que é um processo extremamente oneroso e laborioso (MEAGHER *et al.*, 2005; PEUMANS *et al.*, 2000b; SHARMA; VIJAYAN, 2011b; SINGH *et al.*, 2005). Existe uma grande lacuna no conhecimento de novas estruturas de lectinas de banana, que poderia ser resolvida através da utilização da modelagem computacional.

2.8.2 Refinamento e checagem de qualidade

A precisão dos modelos tridimensionais previstos é um fator crítico para estudos mais detalhados, como design de drogas, *docking* de proteínas, previsão da função de proteínas, estudos de mutagênese, triagem de dados e análises de sítios de ligação. A predição da estrutura tridimensional não é isenta de falhas, podendo apresentar erros locais e globais, incluindo, ligações irregulares, ângulos e comprimentos incomuns nos modelos previstos. Esses erros limitam o uso dos modelos para os estudos posteriores. A necessidade de aumentar a precisão da modelagem de estruturas terciárias e a correção de possíveis erros levaram ao desenvolvimento de métodos para o refinamento de modelos (ADIYAMAN; MCGUFFIN, 2019).

As abordagens de refinamento da estrutura da proteína usam amostragem conformacional para otimizar um campo de força atomístico, e trazer a estrutura inicial mais próxima de seu estado nativo (BHATTACHARYA, 2019). Normalmente o refinamento dos modelos 3D são divididos em dois estágios principais, amostragem e pontuação (ADIYAMAN; MCGUFFIN, 2019). Para um refinamento bem-sucedido, em primeiro lugar, as abordagens de amostragem devem ser capazes de gerar modelos 3D alternativos que estejam mais próximos da estrutura nativa do que o modelo inicial e, em segundo lugar, os modelos 3D gerados devem ser pontuados com precisão por ferramentas de checagem de qualidade, a fim de facilitar identificação daqueles que estão mais próximos da estrutura nativa (FEIG, 2017)

Na checagem de qualidade, os métodos de pontuação de energia variam, dependendo da escolha da referência usada para analisar as interações atômicas com base em estruturas conhecidas (LESK, 2014). A pontuação mais baixa produzida por

estes métodos geralmente se correlaciona com a pontuação mais baixa obtida para o RMSD (distância média entre átomos correspondentes) quanto menor, maior a resolução ou maior similaridade entre estruturas comparadas, pois a distância entre os átomos é menor (ADIYAMAN; MCGUFFIN, 2019; FEIG, 2017).

Outras duas métricas importantes e mais acuradas que o RMSD para a medição de similaridade entre estruturas são os *scores* GDT que demonstra a pontuação global de proteínas com a mesma sequência mas diferentes estruturas terciária, TM que é uma métrica para avaliar a similaridade topológica das estruturas da proteína e conta todos os pares de resíduos, sendo mais sensível a topologia global que a variações locais (LI *et al.*, 2016; XU; ZHANG, 2010; ZEMLA, 2003).

Existem diversos softwares para cada propriedade físico-química apresentada pelas proteínas sendo alguns, o Qmean, Modfold6, Galaxyrefine para mínimas energéticas globais e locais (BENKERT; KÜNZLI; SCHWEDE, 2009; MAGHRABI; MCGUFFIN, 2017); Modloop e HeliQuest para regiões de difícil predição (FISER; DO; ŠALI, 2000; GAUTIER *et al.*, 2008); RAMPAGE, PROCHECK, PDBSum Generate para a checagem de qualidade das posições adotadas pelos aminoácidos (LASKOWSKI *et al.*, 1993, 1997; LOVELL *et al.*, 2003).

2.8.3 Docking molecular

O *docking* molecular é uma ferramenta importante na biologia molecular estrutural e no design de medicamentos auxiliado por computador. O objetivo do *docking* é compreender e prever reconhecimento molecular, tanto estruturalmente, encontrando modos de ligação prováveis, e energeticamente, prevendo afinidade de ligação. O encaixe molecular é realizado entre uma pequena molécula e uma macromolécula alvo. Isso é frequentemente referido como proteína-ligante e em alguns casos, proteína-proteína (MORRIS; LIM-WILBY, 2008). Essa ferramenta tem uma ampla variedade de usos e aplicações na descoberta de drogas, incluindo estudos de estrutura-atividade, fornecendo hipóteses de ligação para facilitar as previsões para estudos de mutagenese quando as mutações estão nos sítios de interação ao ligante, auxiliando a cristalografia de raios-X no ajuste de substratos e inibidores para densidade de elétrons e estudos de mecanismo químico (PAGADALA; SYED; TUSZYNSKI, 2017).

A maioria dos *softwares* de *docking* molecular em uso é responsável por gerar o o espaço conformacional que o ligante (ligante flexível) pode interagir e promove várias tentativas de modelar um receptor de proteína flexível. Cada interação entre a proteína e o ligante é chamado de pose. Os softwares geram um grande número de poses entre o ligante e a proteína em potencial, algumas podem ser rejeitadas imediatamente devido a conflitos com a proteína. As demais são avaliadas usando alguma função de pontuação, que assume uma pose como entrada e retorna um número indicando a probabilidade de que a pose representa uma interação de ligação favorável e classifica um ligante em relação a outro. A maioria das funções de pontuação são campos de força da mecânica molecular com base na física que estimam a energia da postura dentro do local de ligação (AJAY; MURCKO, 1995; MENG *et al.*, 2012).

Todos os métodos de encaixe requerem uma pontuação para classificar as várias possíveis ligações e de um método de pesquisa para explorar as variáveis de estado. As pontuações podem ser empíricas, baseadas no campo de força ou no conhecimento, enquanto os métodos de pesquisa são divididos em duas categorias principais, sistemática e estocástica. (MORRIS; LIM-WILBY, 2008).

A abordagem sistemática explora todos os graus de liberdade do ligante de maneira inúmeras vezes. Porém, essas análises geram do um custo computacional muito alto, dessa forma, para ligantes com muitos graus de liberdade e em análises que dependem de métodos rápidos, essa abordagem torna-se impensável. (BROOIJMANS; KUNTZ, 2003). A respeito dos métodos estocásticos eles variam todos os graus de liberdade do ligante aleatoriamente em cada etapa, gerando uma imensa variedade de soluções. Essas soluções são avaliadas por critérios de probabilidade que decide quais delas serão descartadas e mantidas. A principal desvantagem dessa abordagem é que não garante uma convergência a mínima global, sendo necessário múltiplas execuções para maximizar a probabilidade de encontrar um resultado ótimo (JONES *et al.*, 1997a).

No estudo das lectinas as aplicações do *docking* são muito utilizadas para analisar as possíveis interações das proteínas com glicídios, ajudando a entender as atividades biológicas desempenhadas pelas lectinas (PINTO-JUNIOR *et al.*, 2017b).

A ferramenta de *docking* também pode ser utilizada para comparações entre lectinas do mesmo gênero, para melhor elucidar a sua atividade e interações com ligantes (NASCIMENTO *et al.*, 2018). Em ambos os casos o *docking* molecular foi

realizado com o software CLC Drug Discovery Workbench software, v. 3.0 que possibilitou a análise do perfil interações proteína-ligante com diferentes carboidratos (NASCIMENTO *et al.*, 2018; PINTO-JUNIOR *et al.*, 2017b).

Os softwares mais utilizados são AutoDock Vina, GOLD e MOE-Dock pois apresentaram as melhores predições e pontuações das poses (CORBEIL; WILLIAMS; LABUTE, 2012; JONES *et al.*, 1997b; TROTT; OLSON, 2009; WANG *et al.*, 2016). Também, são muito utilizados nos estudos os seguintes programas: AutoDock (MORRIS *et al.*, 1998) DOCK (EWING; KUNTZ, 1997) e FlexX (RAREY *et al.*, 1996).

3. Objetivos gerais:

Identificar e analisar lectinas da família Musaceae com potencial biotecnológico, através de estratégias de bioinformática, comparando a modelagem estrutural e o perfil de ligação a carboidrato das lectinas com a lectina de *M. acuminata* (PDB ID: 1X1V).

3.1 Objetivos específicos:

- Encontrar e analisar lectinas da família Musaceae com potencial biotecnológico, similares a lectina *M. acuminata* (PDB ID 1X1V);
- Utilizar técnicas de bioinformática para obter a estrutura tridimensional para análises da similaridade com a lectina *M. acuminata* (PDB ID: 1X1V);
- Elucidar a interação ligação aos carboidratos, através da técnica de *docking* molecular, de lectinas do gênero *Musa* e *Ensete*.

4. Materiais e métodos:

4.1 Análise das sequências:

A ferramenta de BLASTp (*Basic Local Alignment Search Tool*) foi utilizada para encontrar as lectinas mais semelhantes da família Musaceae, sendo realizada a partir do alinhamento local, as lectinas mais similares a sequência da lectina de *Musa acuminata* (PDB ID: 1X1V).

Os seguintes parâmetros foram utilizados, a busca foi feita no *database* proteínas não redundantes (nr) tendo como limitante, organismos pertencentes à família Musaceae (ALTSCHUL *et al.*, 1990). Seleccionamos as 22 sequências que apresentaram a maior cobertura e com maior similaridade. Logo após, estas sequências foram alinhadas de forma global utilizando a ferramenta CLUSTAL a fim de avaliar as modificações específicas nos sítios de ligação ao carboidrato (HIGGINS, 1994).

Os dados acerca dos sítios de ligação, das lectinas referência *M.acuminata* (PDB ID 3MIT) e *M. acuminata* (PDB ID 1X1V), foram obtidos no UNIPROT (APWEILER, 2004), de onde seleccionamos 15 lectinas que não possuíam a sua estrutura depositada no banco de dados PDB: THU46913.1; THU46913.1; THU46908.1; THU46912.1; XP_018673706.1; XP_018673705.1; XP_018673708.1; XP_018673707.1; XP_009417126.1; RWW53215.1; RWW21725.1; RWW53214.1; RWV92385.1; RWV92385.1; ADW77219.1; ABS86034.1. Todas as sequências foram analisadas pelo software SignalP para detectar a presença de peptídeo sinal, pois essa porção causa problemas nos resultados da modelagem.

4.2 Modelagem estrutural e refinamento das sequências:

A fim de visualizar a estrutura e analisar o impacto nos sítios de ligação a carboidratos, as sequências que não apresentavam estrutura no PDB foram submetidas a uma modelagem estrutural no *software* I-TASSER (ZHANG, 2008), o qual utiliza uma metodologia de *meta-threading* que consiste na procura em bancos de dados por proteínas com modelos de dobras semelhantes.

A qualidade da modelagem é medida pelo *c-score*, além disso, o *software* ainda fornece dados sobre as ligações a carboidratos e os sítios envolvidos, assim como

outras atividades biológicas. As moléculas foram analisadas visualmente pelo software PyMol (PYMOL MOLECULAR GRAPHICS SYSTEM, SCHOREDINGER, versão 4.20) e foram sobrepostas a *Musa acuminata* (PDB ID: 3MIT) que possui identidade de 100% com a *Musa acuminata* (PDB ID: 1X1V). A escolha da *Musa acuminata* (PDB ID: 3MIT) deve-se ao fato da lectina já estar ligada a manose, que será o carboidrato analisado no docking, e será inferida por cristalografia o que irá aumentar a confiabilidade das predições do *docking*.

Três lectinas foram escolhidas para as análises dos sítios de ligação *M. balbisiana* (THU46912.1), *E. ventricosum* (RWV92385.1) e *M. acuminata subsp. malaccensis* (XP_018673706.1). Foram selecionadas somente três devido à semelhança estrutural entre os resultados do BLASTp e, também, pelas diversas isoformas das mesmas espécies encontradas. Consequentemente, foi escolhida uma representante de cada espécie encontrada. Com a finalidade de refinar e checar a qualidade das moléculas para as futuras análises foram utilizados os seguintes softwares: Galaxyrefine, PDBsum e ProteinSol (GOLDSTEIN *et al.*, 2001; HEO; PARK; SEOK, 2013; LASKOWSKI *et al.*, 1993, 1997; MAGHRABI; MCGUFFIN, 2017).

4.3 Docking molecular

Para analisar o perfil de interação com a manose, três lectinas foram escolhidas (THU46912.1; RWV92385.1; XP_018673706.1), a escolha levou em conta a falta de estudo com as espécies *M. balbisiana* e *E. ventricosum* em conjunto com a diferença nos sítios ativos e a similaridade dos mesmos na espécie de *M. acuminata subsp. malaccensis* pois, essa lectina pode apresentar atividades semelhantes as já descritas para a espécie.

O programa de *docking* empregado foi Autodock Vina e Autodock Tools (MORRIS *et al.*, 2009; TROTT; OLSON, 2009). Através da linha de comando foi obtida a energia de ligação (kcal/mol), calculada pelo software, entre lectina e o carboidrato, assim como a conformações do ligante. O *docking* foi realizado entre as lectinas selecionadas a partir das estruturas obtidas pelo I-TASSER após o refinamento. O ligante utilizado foi a manose e a sua estrutura 3D foi obtida através do RCSB PDB (<http://ligand-expo.rcsb.org/>). O software Plip (*Protein Ligand Interaction Profiler*) forneceu as informações sobre as ligações de hidrogênio com os aminoácidos envolvidos nos sítios de ligação a manose (Salentin *et al.*, 2015).

5. Resultados

5.1 Análises das sequências

Para encontrar novas lectinas potenciais, é essencial buscar por sequências similares já depositadas. Por isso, o alinhamento local é uma ferramenta útil para buscar sequências similares dentro de um vasto banco de dados. A fim de verificar a similaridade entre a sequência de *Musa acuminata* (PDB ID: 1X1V), foi realizada através de BLAST proteína-proteína (BLASTp) tendo como limitante da busca a da família Musaceae a qual a lectina faz parte.

Foram encontradas sequências similares às depositadas referentes a proteínas de *M. acuminata*, *M. balbisiana*, e também de *Ensete ventricosum*. Quando comparadas a sequência de referência (PDB ID: 1X1V) utilizada como *input* para BLASTp, aquelas que apresentaram maior similaridade foram das espécies *M. balbisiana* com 99% (THU46907.1) e *M. acuminata* com 100% (ABS86034.1) enquanto para as do gênero *Ensete ventricosum*, a maior similaridade encontrada é de 89% RWW53215.1 (Tabela 1). Também, excluímos as sequências de *M. acuminata* que já possuíam depósito no PDB, pois para explorar novas lectinas com potencial biotecnológico, optamos por modelar lectinas que ainda não estavam depositadas nos bancos de dados, a fim de analisar as interações com os carboidratos e alterações nos sítios de ligação.

Tabela 1 Seleção das 22 sequências que apresentaram maior similaridade com a sequência *M. acuminata* (PDB ID 1X1v) o software BLASTp.

Número de acesso	Organismo de origem	Similaridade
THU46913.1	<i>Musa balbisiana</i>	96%
THU46907.1	<i>Musa balbisiana</i>	99%
THU46908.1	<i>Musa balbisiana</i>	95%
THU46912.1	<i>Musa balbisiana</i>	91%
XP_018673706.1	<i>Musa acuminata</i> subsp. <i>malaccensis</i>	96%
XP_018673705.1	<i>Musa acuminata</i> subsp. <i>malaccensis</i>	96%
XP_018673708.1	<i>Musa acuminata</i> subsp. <i>malaccensis</i>	95%

XP_018673707.1	<i>Musa acuminata subsp. malaccensis</i>	95%
XP_009417126.1	<i>Musa acuminata subsp. malaccensis</i>	91%
1X1V_A	<i>Musa acuminata</i>	100%
5EXG_A	<i>Musa acuminata</i>	93%
2BMY_A	<i>Musa acuminata</i>	93%
4PIF_A	<i>Musa acuminata</i>	93%
4PIU_A	<i>Musa acuminata</i>	92%
4PIT_A	<i>Musa acuminata</i>	92%
RWW53215.1	<i>Ensete ventricosum</i>	89%
RWW21725.1	<i>Ensete ventricosum</i>	89%
RWW53214.1	<i>Ensete ventricosum</i>	88%
RWV92385.1	<i>Ensete ventricosum</i>	87%
ADW77219.1	<i>Musa acuminata</i> AAA group	92%
ABS86034.1	<i>Musa acuminata</i> AAA group	100%

Logo após o alinhamento global no CLUSTAL, foram observadas modificações pontuais nos sítios de ligação ao carboidrato (33-38; 60-63; 105-106; 130-133) recuperadas do banco de dados UniProt da sequência de referência *M. acuminata* (PDB ID 1X1V). Variações nesses sítios estão potencialmente ligadas a alterações na afinidade, podendo alterar todo o perfil de ligação ao açúcar e a atividade biológica (SWANSON *et al.*, 2015). A maioria das mutações ocorreram no sítio 130-133.

Quando observadas mais acuradamente, essas alterações se fizeram mais presentes no aminoácido da posição 130 das sequências de lectinas analisadas (Tabela 2). Na sequência de referência o sítio 130-133 é composto pelos aminoácidos DFID, o aspartato (D), localizado na posição 130, possui carga negativa, em algumas sequências ele foi trocado para uma lisina (K) ou glutamina (Q), que são aminoácidos com característica carga positiva e apolar, respectivamente (Tabela 2).

Tabela 2 Mutações nos locais de ligação ao carboidrato de lectinas da família Musaceae que não possuem estrutura no PDB. As modificações dos resíduos estão assinaladas em negrito e sublinhadas.

Organismo de origem	ID	Locais de ligação ao carboidrato			
		30-38	60-63	105-106	130-133
<i>Musa acuminata</i>	1X1V	SGDVVD	GTPH	GN	DFID
<i>Musa balbisiana</i>	THU46913.1	SGDVVD	GTPH	GN	<u>KFLD</u>
<i>Musa balbisiana</i>	THU46907.1	SGDVVD	GTPH	GN	<u>DFLD</u>
<i>Musa balbisiana</i>	THU46908.1	SGDVVD	GTPH	GN	<u>KFLD</u>
<i>Musa balbisiana</i>	THU46912.1	SGDVVD	GTPH	GN	<u>QFLD</u>
<i>Musa acuminata</i> subsp. <i>malaccensis</i>	XP_018673706.1	SGDVVD	GTPH	GN	<u>DFLD</u>
<i>Musa acuminata</i> subsp. <i>malaccensis</i>	XP_018673705.1	SGDVVD	GTPH	GN	<u>DFLD</u>
<i>Musa acuminata</i> subsp. <i>malaccensis</i>	XP_018673708.1	SGDVVD	GTPH	GN	<u>QFLD</u>
<i>Musa acuminata</i> subsp. <i>malaccensis</i>	XP_018673707.1	SGDVVD	GTPH	GN	<u>KFLD</u>
<i>Musa Acuminata</i> subsp. <i>malaccensis</i>	XP_009417126.1	SGDVVD	GTPH	GN	<u>KFLD</u>
<i>Ensete</i> <i>ventricosum</i>	RWW53215.1	SGDVVD	<u>GN</u> PH	GN	<u>QFLD</u>
<i>Ensete</i> <i>ventricosum</i>	RWW21725.1	SGDVVD	<u>GN</u> PH	GN	<u>QFID</u>
<i>Ensete</i> <i>ventricosum</i>	RWW53214.1	SGDVVD	<u>GN</u> PH	GN	<u>QFLD</u>
<i>Ensete</i> <i>ventricosum</i>	RWV92385.1	SG <u>G</u> VVD	<u>GN</u> PH	GN	<u>QFLD</u>

<i>Musa acuminata</i> AAA group	ADW77219.1	SGDVVD	GTPH	GN	<u>QFLD</u>
<i>Musa acuminata</i> AAA group	ABS86034.1	SGDVVD	GTPH	GN	DFID

Após analisar os resultados é possível observar que existem 3 grupos distintos de proteínas. Do gênero *Musa* foram encontradas lectinas de duas espécies *M. acuminata malacencis* e *M. balbisiana* ambas com diferentes isoformas da mesma lectinas. Do gênero *Ensete* foi encontrada a lectina da espécie *E. ventricosum* com diferentes isoformas.

5.2 Modelagem estrutural e refinamento das sequências

Com o intuito de observar as modificações nos sítios de carboidrato para posteriormente analisar as interações com o ligante, as moléculas foram modeladas no I-TASSER, pois apenas a sequência primária de uma proteína não nos permite explorar essas informações (WRIGHT; DYSON, 1999). As moléculas obtidas no I-TASSER apresentaram um *c-score* variando entre -0,97 até 1,29 (Tabela 3). O *c-score* é uma pontuação de confiança adotada para medir a qualidade dos modelos previstos pelo I-TASSER, os valores variam entre -5 e 2, sendo que os valores mais altos apresentam modelos de melhor confiança. As lectinas da espécie *M. balbisiana* apresentaram os únicos valores negativos de *c-score*, sendo assim, em uma análise geral a espécie que apresentou as menores pontuações.

Tabela 3 Valores de *c-score* e primeiro PDB hit gerado para as lectinas da família Musaceae através do software I-TASSER.

ID	Nome	<i>c-score</i>
THU46913.1	<i>Musa balbisiana</i>	0,92
THU46907.1	<i>Musa balbisiana</i>	-0,93
THU46908.1	<i>Musa balbisiana</i>	-0,18
THU46912.1	<i>Musa balbisiana</i>	-0,97
XP_018673706.1	<i>Musa acuminata</i> subsp. malaccensis	0,86

XP_018673705.1	<i>Musa acuminata</i> subsp. malaccensis	1,27
XP_018673708.1	<i>Musa acuminata</i> subsp. malaccensis	1,27
XP_018673707.1	<i>Musa acuminata</i> subsp. malaccensis	1,29
XP_009417126.1	<i>Musa acuminata</i> subsp. malaccensis	1,28
RWW53215.1	<i>Ensete ventricosum</i>	0,89
RWW21725.1	<i>Ensete ventricosum</i>	0,90
RWW53214.1	<i>Ensete ventricosum</i>	1,18
RWV92385.1	<i>Ensete ventricosum</i>	0,86
ADW77219.1	<i>Musa acuminata</i> AAA group	1,28
ABS86034.1	<i>Musa acuminata</i> AAA group	1,23

Além do *c-score*, o I-TASSER fornece a previsão de possíveis sítios de ligação e o provável ligante. Desta maneira podemos visualizar uma predição acerca das ligações entre as lectinas selecionadas a carboidratos, as únicas que não apresentaram, na predição, ligação a manose foram *M. acuminata subsp. malaccensis* (XP_018673705.1) e *M. acuminata* AAA group (ABS86034.1). A maioria dos sítios de ligação foram condizentes com os dados obtidos no Uniprot e as divergências que ocorreram foram relativamente próximas aos sítios utilizados como referência.

Tabela 4 Sítios de ligações à manose das lectinas da Família Musaceae fornecidos pelo módulo COACH do software I-TASSER.

ID	Nome	Sítios de ligação à manose
THU46913.1	<i>Musa balbisiana</i>	36,37,66,106,107,108,109
THU46907.1	<i>Musa balbisiana</i>	64,65,94,134,135,136,137
THU46908.1	<i>Musa balbisiana</i>	50,51,80,120,121,122,123
THU46912.1	<i>Musa balbisiana</i>	70,71,142,185,186,187,189; 89,90,119,159,160,161,162
XP_018673706.1	<i>Musa acuminata</i> subsp. malaccensis	36,37,66,106,107,108,109

XP_018673705.1	<i>Musa acuminata</i> subsp. malaccensis	Não apresentou sítio de ligação
XP_018673708.1	<i>Musa acuminata</i> subsp. malaccensis	33,34,63,103,104,105,106
XP_018673707.1	<i>Musa acuminata</i> subsp. malaccensis	33,34,63,103,104,105,106
XP_009417126.1	<i>Musa acuminata</i> subsp. malaccensis	33,34,63,103,104,105,106
RWW53215.1	<i>Ensete ventricosum</i>	36,37,66,106,107,108,109; 17,18,89,132,133,134,136
RWW21725.1	<i>Ensete ventricosum</i>	36,37,66,106,107,108,109; 17,18,89,132,133,134,136
RWW53214.1	<i>Ensete ventricosum</i>	35,36,65,105,106,107,108; 17,18,89,132,133,134,136
RWV92385.1	<i>Ensete ventricosum</i>	36,37,66,106,107,108,109
ADW77219.1	<i>Musa acuminata</i> AAA group	33,34,63,103,104,105,106
ABS86034.1	<i>Musa acuminata</i> AAA group	Não apresentou sítio de ligação

No software Pymol foi realizada a sobreposição das estruturas geradas pelo I-TASSER utilizando a molécula de *M. acuminata* resolvida (PDB ID: 3MIT) como referência pois sua estrutura está ligada à manose. O intuito foi analisar a pontuação RMSD, para observar as diferenças estruturais entre os modelos gerados e a referência (Tabela 5).

Tabela 5 Valores de RMSD gerados pela sobreposição entre os modelos tridimensionais gerados pelo I-TASSER e a lectina de *M. acuminata* (PDB ID: 3MIT)

ID	Organismo de origem	RMSD
THU46907.1	<i>Musa balbisiana</i>	0,583
THU46913.1:	<i>Musa balbisiana</i>	0,586
THU46908.1	<i>Musa balbisiana</i>	0,694
THU46912.1	<i>Musa balbisiana</i>	5.955
XP_018673706.1	<i>Musa acuminata</i> subsp. malaccensis	0,581
XP_018673705.1	<i>Musa acuminata</i> subsp. malaccensis	0,453
XP_018673708.1	<i>Musa acuminata</i> subsp. malaccensis	0,477
XP_018673708.1	<i>Musa acuminata</i> subsp. malaccensis	0,477

XP_018673707.1	<i>Musa acuminata subsp. malaccensis</i>	0,598
XP_009417126.1	<i>Musa acuminata subsp. malaccensis</i>	0,581
RWW53215.1	<i>Ensete ventricosum</i>	0,435
RWW21725.1	<i>Ensete ventricosum</i>	0,527
RWW53214.1	<i>Ensete ventricosum</i>	0,583
RWV92385.1	<i>Ensete ventricosum</i>	0,527
ADW77219.1	<i>Musa acuminata AAA group</i>	0,522
ABS86034.1	<i>Musa acuminata AAA group</i>	0,914

A *M. balbisiana* (THU46912.1), foi a lectina que apresentou o maior valor de RMSD (5,955) em relação à *M. acuminata* (3MIT) diferindo de todos os demais modelos gerados que obtiveram valores de RMSD inferiores a 1 (Tabela 5). Quanto maior valor para o RMSD maior diferença estrutural, já que é uma métrica que avalia o desvio médio dos átomos de duas estruturas tridimensionais.

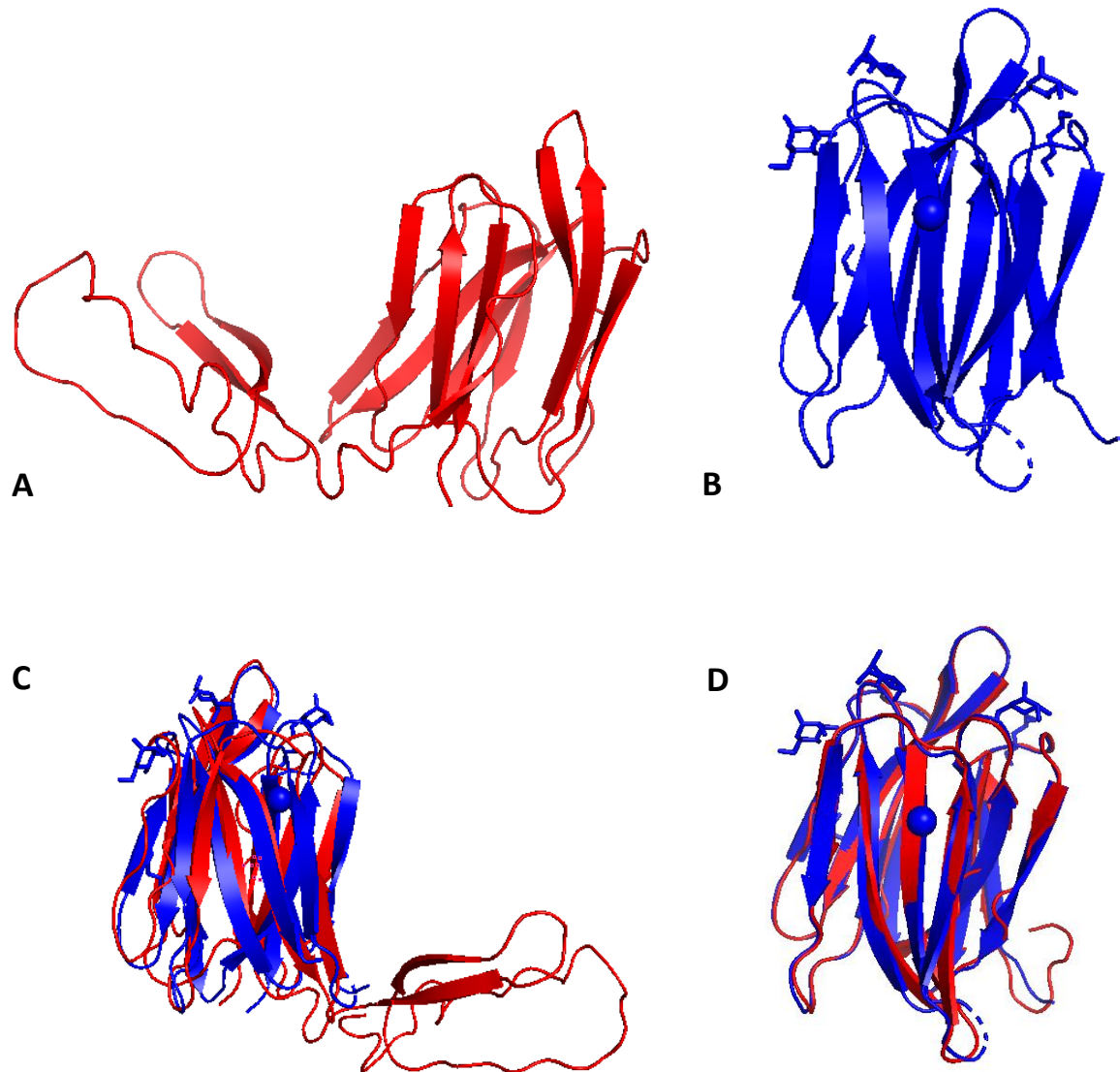


Figura 4 Estrutura tridimensional de *M. balbisiana* (THU46912.1) e obtida através do software I-TASSER, sobreposta a lectina de *M. acuminata* referência (PDB ID 3MIT) . *M. balbisiana* (THU46912.1) (A); *M. acuminata* (PDB ID 3MIT) (B); Sobreposição das lectinas de *M. balbisiana* (THU46912.1) e *M. acuminata* (PDB ID 3MIT) (C); Sobreposição da lectina *M. balbisiana* (THU46912.1), sem as regiões de baixa complexidade e repetições internas, e *M. acuminata* (PDB ID 3MIT) (D).

A lectina de *M. balbisiana* (THU46912.1) apresenta uma sequência maior (Figura 4) quando comparada com a *M. acuminata* (3MIT). A porção inicial de *M. balbisiana* (THU46912.1) se refere a regiões de repetições internas e baixa complexidade obtidas através do webserver SMART, que vão do aminoácido 1 até o 56. Essas regiões foram descartadas nas análises de bioinformática do presente trabalho, porém são necessários mais estudos para elucidar a sua importância na estrutura da lectina. Ao investigar por meio de alinhamento local, foi visto que essa região inicial de *M. balbisiana* (THU46912.1) não se repete nas outras lectinas selecionadas nem em outras espécies da família Musaceae.

THU46912.1	1	MVDRTNVTFTYYGKMKTRHDGGTSHETVLEEDEDVVGMEGESLVTSVRCV	50
3MITA	1	-----	0
THU46912.1	51	LTSMVMQNGAIKVGAWGGNGGSAFDMGPAYRIISVKIFSGDVVDAVDVTF	100
		.	
3MITA	1	-----MNGAIKVGAWGGNGGSAFDMGPAYRIISVKIFSGDVVDAVDVTF	44
THU46912.1	101	TYYGKTETRHFSGSGGTPHEIVLQEGEYLVGMAGEVANYHGLVWVGKLG	150
		:	
3MITA	45	TYYGKTETRHFSGSGGTPHEIVLQEGEYLVGMKGEFGNYHGWWWVGKLG	94
THU46912.1	151	STNKKAYGPFNGAGGTPFSLPISAGKISGFFGRDQQLDAIGVYLAP	197
		:	
3MITA	95	STNKKSYGPFNGTGGTPFSLPIAAGKISGFFGRGGDFIDAIGVYLEP	141

Figura 5 Alinhamento feito no EMBOSS Needle entre as lectinas *M. balbisiana* (THU46912.1) e *M. acuminata*. Matriz EBLOSUM62; Identidade 65%; Similaridade 67%; Gaps 28.4%; Score 696 (PDB ID 3MIT).

Uma nova sobreposição foi realizada entre a lectina *M. balbisiana* (THU46912.1) e *M. acuminata* (PDB ID 3MIT) para avaliar o valor e RMSD sem a porção inicial que a difere das outras lectinas do estudo. Esse resultado diminuiu para 0,914 se aproximando da média apresentada pelos outros modelos.

5.3 Docking Molecular

Após a inspeção visual dos modelos, os que melhor se encaixaram nos requisitos escolhidos apresentavam maior valor de RMSD, mais alterações nos sítios de ligação ao carboidrato e menos alterações nos sítios de ligação ao carboidrato

foram submetidas as posteriores análises de refinamento e docking molecular. Para o processo de refinamento apenas três lectinas foram escolhidas, uma do gênero *Ensete*, pois não há informações que elucidem a o perfil de interação a carboidrato das lectinas desse gênero e há poucas informações sobre as atividades biológicas exercidas por elas, e duas do gênero *Musa*, uma representante da espécie *M. balbisiana* devido a carência de informações e estudos e outra da espécie *M. acuminata subsp. malaccensis* em razão da grande identidade com a sequência de BanLec (*M. acuminata*) o que pode acarretar em atividades biológicas semelhantes. As três lectinas foram submetidas ao refinamento pelo software GalaxyRefine (*M. balbisiana* THU46912.1; *E. ventricosum* RWV92385.1; *M. acuminata subsp. malaccensis* XP_018673706.1) e obtiveram um melhor posicionamento de seus aminoácidos quando submetidas ao gráfico de Ramachandran, partindo de uma pontuação de cerca de 67% de aminoácidos em posições favoráveis para 82% em posições favoráveis. A checagem da qualidade dos modelos é feita com base nas classificações de combinações dos ângulos dividida nas quatro regiões do diagrama: as regiões favoráveis, as permitidas, as generosamente permitidas e as proibidas. A região permitida indica conformações permitidas nos limites extremos para os contatos atômicos desfavoráveis. A região generosamente permitida reflete conformações que são permitidas somente se houver certa flexibilidade nos ângulos de ligação. Já a região proibida, as conformações são quase totalmente impedidas por fatores estéricos (RAMACHANDRAN; RAMAKRISHNAN; SASISEKHARAN, 1963).

Tabela 6 Valores obtidos a partir do gráfico de Ramachandran antes e depois do refinamento das lectinas selecionadas (THU46912.1; RWV92385.1; XP_018673706.1).

ID	Antes do refinamento			Após o refinamento		
	Regiões favoráveis	Regiões permitidas	Regiões não permitidas	Regiões favoráveis	Regiões permitidas	Regiões não permitidas
THU46912.1	67,90%	26,70%	5,50%	82,60%	11,90%	5,50%
RWV92385.1	67,30%	28,20%	4,50%	82,70%	13,60%	3,60%
XP_018673706.1	67,30%	28,20%	4,50%	83,60%	12,70%	3,60%

As energias de ligação das três lectinas selecionadas, foram muito próximas da lectina de referência de *M. acuminata* (PDB ID 3MIT). Todos os valores de energia de

ligação variaram de - 4,4 a -3,3. A lectina de *M. acuminata* subsp. *malaccensis* (XP_018673706.1) (-4.4), apresentou maior energia de ligação, porém em uma análise geral, todas as lectinas apresentaram valores muito semelhantes.

Tabela 7 Energias de ligação obtidas através de *docking* molecular para as três lectinas escolhidas para análise (THU46912.1; RWV92385.1; XP_018673706.1)

Pose	Energia de ligação a manose (kcal/mol)			
	<i>M. acuminata</i> (3MIT)	<i>M. balbisiana</i> (THU46912.1)	<i>E. ventricosum</i> (RWV92385.1)	<i>M. acuminata</i> subsp. <i>malaccensis</i> (XP_018673706.1)
1	-5,1	-4,1	-4,1	-4,4
2	-4,9	-3,8	-4,1	-4,1
3	-4,8	-3,7	-4,0	-4,1
4	-4,7	-3,5	-4,0	-4,0
5	-4,7	-3,5	-4,0	-4,0
6	-4,5	-3,4	-3,9	-3,9
7	-4,4	-3,3	-3,9	-3,9
8	-4,3	-3,3	-3,8	-3,9
9	-4,2	-3,3	-3,8	-3,9

Com a análise fina da interação com a manose fornecida pelo *docking molecular*, é possível observar que as interações das três lectinas selecionadas (*M. balbisiana* THU46912.1; *E. ventricosum* RWV92385.1; *M. acuminata* subsp. *malaccensis* XP_018673706.1) as 9 poses apresentaram valores que variaram de - 5,1 a -3,3 sendo que a *M. acuminata* (3MIT) apresenta todos os 9 poses diferentes das demais lectinas em estudo, os menores poses foram encontrados para a lectina de *M. balbisiana* (THU46912.1) as lectinas de *E. ventricosum* RWV92385.1; *M. acuminata* subsp. *malaccensis* XP_018673706.1) possui o maior número de poses iguais, aos sítios que obtivemos como referência no Uniprot. Porém, quando realizamos o *redocking* com a lectina referência *M. acuminata* (PDB ID 3MIT) e analisamos o *output* no *webserver* PLIP, os sítios envolvidos na interação com a

manose são muito divergentes daqueles obtidos no Uniprot, com exceção da pose 6, no qual todos os sítios de ligação estão condizentes com os de referência.

Ao examinar os resultados do *webserver* PLIP, as todas as lectinas selecionadas apresentaram proximidade com os sítios de ligação da referência. Quando observamos os aminoácidos envolvidos na interação da lectina *M. acuminata subsp. malaccensis* podemos concluir que seus resíduos são os que mais de diferem da lectina de *M. acuminata* (PDB ID 3MIT (30 -38; 60-63; 105-106; 130-133)).

Tabela 8 Resíduos de interação com a manose obtidos através do ??? molecular para as lectinas selecionadas para a análise (THU46912.1; RWV92385.1; XP_018673706.1).

<i>Resíduos de interação com a manose</i>				
Pose	<i>M. acuminata</i> (3MIT)	<i>M. balbisia</i> (THU46912.1)	<i>E. ventricosum</i> RWV92385.1	<i>M. acuminata subsp.</i> <i>malaccensis</i> (XP_018673706.1)
1	34,35,36,60	36,37,38,39,41,63,86	36,37,64,108,109	79,96,116,117
2	33,63,105,106	37,38,39,40,41,63,86	36,37,108,109	79,97,98,104,116,117
3	33,63,105,106	37,38,39,41,63,86	36,64,109	79,116,117
4	34,35,36,38	38,39,41	36,64,66,108,109	36,64,108,109
5	33,63,105,106	57,59,131	36,64,108,109	36,37,64,66,108
6	34,35,36,60	37,38,39,63,86	36,37,64,109	64,66,108,109
7	34,35,36,38,60	38,40,41,62,63	21,29,129,130	37,38,39,63,86
8	34,35,36,38,60	57,59,131	9,21,129,130	64,66,108,109
9	35,36,60	59,132,133	36,37,64,108,109	64,66,108,109

A lectina *M. acuminata* (PDB ID 3MIT) utilizada para o *redocking* teve os seus ligantes retirados para a análise de interação com a molécula de manose através da mesma técnica de *docking* aplicada para as outras lectinas. Ao sobrepormos a estrutura da *M. acuminata* (PDB ID 3MIT) com o *redocking* analisando a ligação com a manose, é possível notar uma grande semelhança na distribuição espacial das moléculas de manose, contendo apenas uma das interações da cristalografia sem sobreposição de alguma das poses do *redocking*.

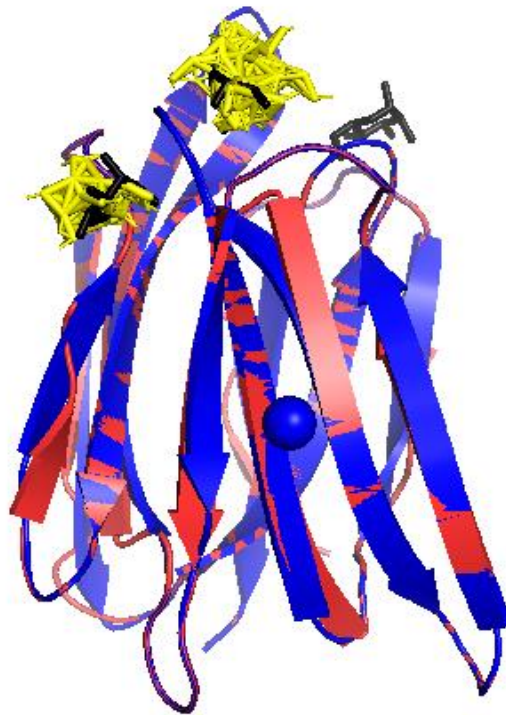


Figura 6 Sobreposição da *M. acuminata* referência (PDB ID 3MIT) em azul com as manoses em preto com o seu *redocking* em vermelho com as manoses ligantes em amarelo.

Ao analisar, no Pymol, as poses geradas pelo *docking* utilizando a manose como ligante, as três lectinas selecionadas (*M. balbiana* THU46912.1; *E. ventricosum* RWV92385.1; *M. acuminata subsp. malaccensis* XP_018673706.1) apresentaram proximidade a disposição espacial das moléculas do ligante com a referência. Todas as lectinas apresentaram a disposição espacial dos carboidratos próximas aos de referência (PDB ID 3MIT) (Figura 7).

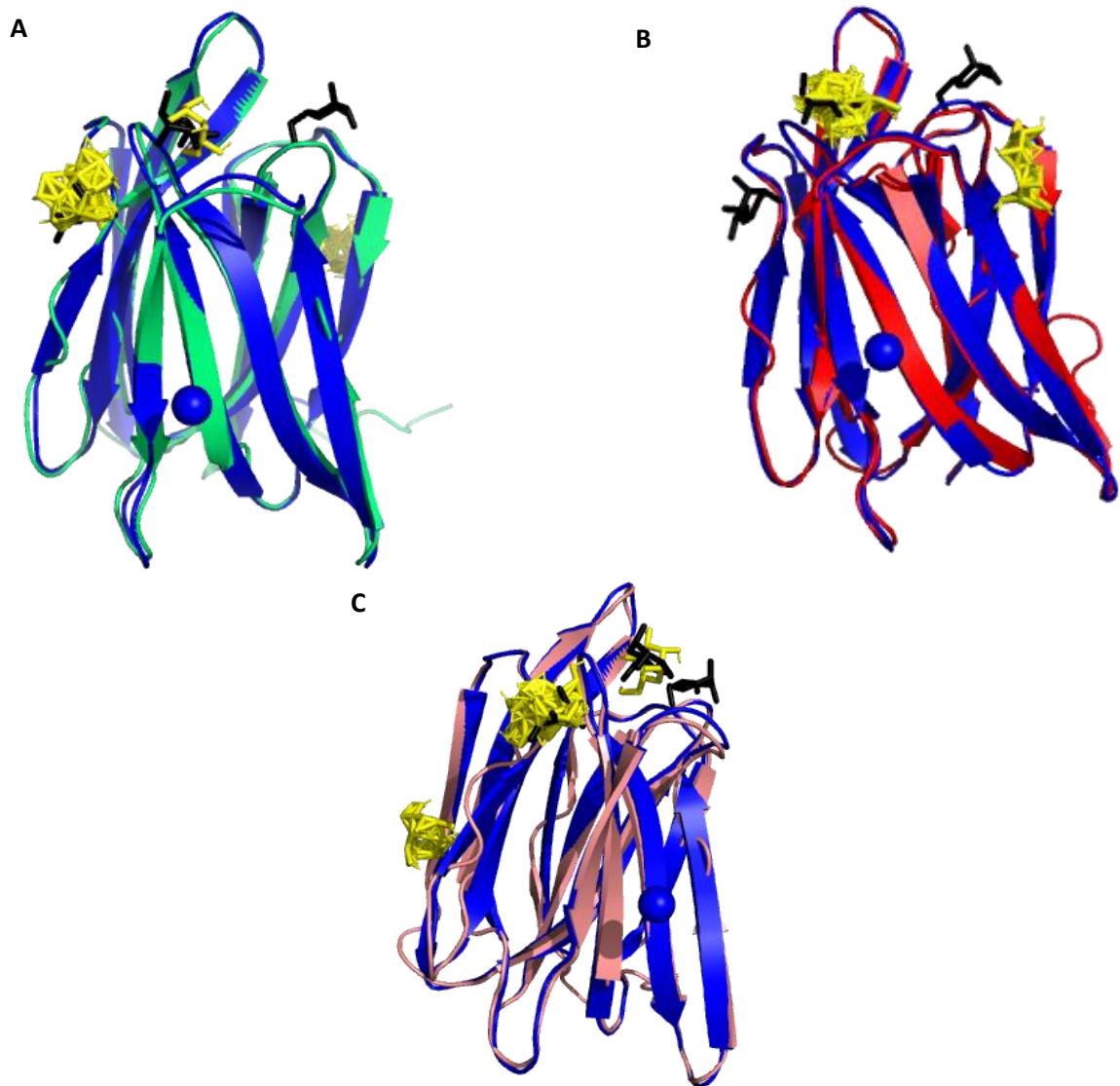


Figura 7 Sobreposição do resultado do docking, utilizando a manose como ligante, para a análise da disposição dos ligantes. *E. ventricosum* RWV92385.1 (verde claro) (A) *M. balbisiana* THU46912.1 (rosa escuro) (B) *M. acuminata* subsp. *malaccensis* XP_018673706.1 (rosa claro) (C) com a *M. acuminata* (PDB ID 3MIT) (azul escuro). Ligantes *M. acuminata* (PDB ID 3MIT) (preto); ligante dos modelos (amarelo).

6. Discussão

Há uma grande necessidade de descoberta de novas biomoléculas que desempenhem papéis estratégicos em diversas áreas da biotecnologia. O conhecimento detalhado da estrutura da proteína e da interação estabelecida com os carboidratos é um pré-requisito para a melhor compreensão dos produtos biológicos para o desenvolvimento de processos biotecnológicos e aplicações terapêuticas (ARNAUD; AUDFRAY; IMBERTY, 2013).

As atividades biológicas das lectinas são elucidadas em diversas espécies e apresentam uma ampla gama de ações, porém, como relatado ao longo deste trabalho, muitas lectinas da família Musaceae não são caracterizadas, o que limita o grande potencial que essa família pode vir a apresentar.

Existem diversos estudos que utilizam extratos totais de espécies diferentes de bananas, comestíveis e não comestíveis, avaliando diversas atividades biológicas podemos citar, antiviral, antibacteriana e cicatrizante (IMAM; AKTER, 2011; JAIN *et al.*, 2011; KAPADIA; PUDAKALKATTI; SHIVANAİKAR, 2015; PANDA *et al.*, 2020b; PEREIRA; MARASCHIN, 2015; SAROWAR HOSSAIN *et al.*, 2011). Os estudos utilizam extratos totais, onde estão presentes diversos compostos que podem apresentar essas atividades, inclusive as lectinas.

Diversas lectinas da família relacionadas a jacalina (JRL's), como a BanLec (Lectina de *M. acuminata*) possuem atividades antivirais, cicatrizantes, antioxidantes, antibacterianas estabelecidas (DE CAMARGO *et al.*, 2020; GOLDSTEIN, 2002b; LIU; LI; ZHANG, 2014; SINGH; DEVI; NG, 2014; STOJANOVIĆ *et al.*, 2010; SWANSON *et al.*, 2010).

Poucos estudos caracterizam as amostras totais da extração, o que contribui para o pouco que se sabe a respeito do potencial biotecnológico das lectinas dessa família. Explorar as lectinas da família Musaceae é de grande importância, devido as diversas propriedades relatadas a partir dos extratos totais de diferentes espécies aliada a grande necessidade de novos compostos para o tratamento de diversas doenças (DENARO *et al.*, 2020; KAISER *et al.*, 2002).

No presente trabalho foram encontradas 22 sequências de lectinas que apresentaram 100% de similaridade com o input *M. acuminata* (PDB ID 1X1V). As lectinas encontradas eram de apenas três espécies diferentes representando dois gêneros *Musa* e *Ensete*, além disso, ainda havia diversas isoformas da mesma lectina,

o que reforça ainda mais a escassez de informações disponíveis sobre as lectinas da família Musaceae.

Os sítios de ligação a carboidratos foram analisados após uma busca na literatura e nos bancos de dados, utilizando os resultados do alinhamento global. A princípio, os sítios de ligação a carboidrato da *M. acuminata* (PDB ID: 1X1V), lectina aplicada como referência, foram retirados do banco de dados Uniprot.

Posteriormente, comparando o artigo utilizado como fonte e os sítios expostos no Uniprot, foram encontradas diferenças entre as informações (SHARMA; VIJAYAN, 2011a). Diante desse impasse, foi decidido manter os sítios retirados do Uniprot, pois eram os mesmos para ambas as lectinas utilizadas como referência neste estudo *M. acuminata* (PDB ID: 1X1V) e *M. acuminata* (PDB ID: 3MIT).

Assim, com um referencial adotado, foi possível realizar a comparação dos sítios de interação ao ligante das lectinas selecionadas. Elas apresentaram alterações de aminoácidos com propriedades químicas diferentes nos sítios de ligação. Essas alterações podem representar grandes mudanças nas propriedades biológicas das lectinas, até mesmo a mudança de um único carboidrato pode alterar o perfil de ligação ao carboidrato alterando drasticamente as atividades antes apresentadas. Um dos maiores exemplos é a versão recombinante da BanLec H84T, a alteração de uma histidina para uma treonina na posição 84 diminuiu a sua capacidade mitogênica aumentando o interesse do uso biotecnológico da atividade antiviral dessa lectina (SWANSON *et al.*, 2015). Em outros casos a mudança de um aminoácido pode não ter alteração alguma nas atividades desempenhadas pela proteína (ROBERTO *et al.*, 2015). Porém, apenas a informação de que há certa alteração nos sítios não é suficiente para inferir qualquer teoria acerca da mudança do perfil de interação a carboidratos e da atividade biológica de uma lectina, são necessárias análises mais detalhadas e a validação dos resultados *in vitro* (NASCIMENTO *et al.*, 2018).

Estudos acerca da predição da interação fina entre ligante e proteína podem ser feitos através do método de *docking* molecular, que se baseia na utilização de softwares que avaliam a energia de ligação das prováveis interações entre a lectina e o carboidrato (NECO *et al.*, 2018; PINTO-JUNIOR *et al.*, 2017a, 2017b). Para essa análise é necessária a obtenção da estrutura tridimensional das proteínas, assim, através do *software* I-TASSER foi obtida o modelo 3D de todas as lectinas que não possuíam a estrutura depositada no PDB. Todos os modelos obtidos foram sobrepostos a lectina *M. acuminata* (PDB ID 1X1V) e apresentaram valores de RMSD

extremamente baixos, demonstrado que as lectinas da família Musaceae são estruturalmente semelhantes.

Antes das análises de *docking* molecular, foram necessárias etapas de refinamento dos modelos estruturais obtidos, com o intuito de melhorá-los a fim de obter uma predição mais próxima possível do modelo *in vitro* (FEIG, 2017; FEIG; MIRJALILI, 2016; HIRANUMA *et al.*, 2021). Para esse processo apenas três lectinas foram escolhidas, uma do gênero *Ensete*, pois não há informações que elucidem a o perfil de interação a carboidrato das lectinas desse gênero e há poucas informações sobre as atividades biológicas exercidas por elas, e duas do gênero *Musa*, uma representante da espécie *M. balbisiana* devido a carência de informações e estudos e outra da espécie *M. acuminata subsp. malaccensis* em razão da grande identidade com a sequência de BanLec (*M. acuminata*) o que pode acarretar em atividades biológicas semelhantes.

O perfil de interação e a disposição espacial do ligante na proteína são sempre comparados a um modelo, nesse estudo a lectina utilizada foi *M. acuminata* (PDB ID: 3MIT), por já possui em seu depósito a interação com a manose realizada por cristalografia de Raios-X.

Além da realização do *docking* molecular com as lectinas selecionadas, foi realizado um *redocking* com a lectina de *M. acuminata* (PDB ID: 3MIT), para permitir a comparação da predição realizada *in silico* e *in vitro*. O *redocking* apresentou resultado muito similares aos gerados anteriormente *in vitro*, o que revalida os resultados obtidos por esse estudo (SHARMA; VIJAYAN, 2011a). Todas as energias de ligação obtidas na análise foram próximas aos valores do *redocking* da lectina de *M. acuminata* (PDB ID: 3MIT), demonstrado que as lectinas selecionadas podem apresentar afinidade pela manose.

Na análise dos resíduos envolvidos na interação no webserver PLIP, os três modelos apresentaram diversos aminoácidos que coincidiram, ou eram muito próximos, com os dos sítios apresentados como referência extraído do Uniprot. Na análise da disposição espacial dos ligantes, as três lectinas analisadas apresentaram a maioria das poses semelhantes às da *M. acuminata* (PDB ID 3MIT) que teve a sua estrutura resolvida *in vitro*.

Essas análises corroboram com a teoria que lectinas de outros gêneros da família Musaceae podem apresentar um perfil interação ao carboidrato semelhante ao da Banlec, podendo apresentar atividades biológicas de grande interesse para a

biomedicina em tratamentos antivirais contra doenças como Ebola, H1N1, SARS-COV 2 e HIV.

A capacidade da lectina de ligação a manose pode fazer com que desenvolva uma atividade antiviral através da ligação ao capsídeo viral, impedindo que se ligue na célula e se multiplique, impedir o contato vírus-célula ao se ligar nos epítomos utilizados para a ancoragem viral na célula e ainda, se ligar a células do sistema imune estimulando a resposta. Um dos exemplos, é a atividade de lectinas JRL's frente ao HIV, que interagem com proteínas glicosiladas com alto teor de manose (gp120 e gp41) do complexo Env que é responsável pela fusão do envelope viral e a membrana do hospedeiro.

Conjuntamente, as lectinas dessa família apresentam a capacidade de se ancorar em glicanos com alto teor de manose na superfície das células TCD4+ impedindo a infecção das mesmas (COVÉS-DATSON *et al.*, 2019; KEYAERTS *et al.*, 2007; MITCHELL; RAMESSAR; O'KEEFE, 2017).

No campo da oncologia, as JRL's apresentam diversas atividades de interesse, constituindo um potencial terapêutico antitumoral. A lectina de *Artocarpus heterophyllus* possui afinidade a manose e exibe potente atividade anticâncer em células mielóides leucêmicas humanas NB4, através da interrupção de MMP em uma via independente de caspase. A lectina de banana Del Monte (*Musa acuminata*), que apresenta o potencial de se ligar a frutose ou manose, interage com os oligossacarídeos contendo manose na superfície celular, e retarda a proliferação de células cancerígenas das linhagens L1210 e HepG2, implicando em lectinas potencialmente exploráveis para utilização terapêutica anticâncer (BHUTIA *et al.*, 2019; COVÉS-DATSON *et al.*, [s. d.], 2020; DE MEJÍA; PRISECARU, 2005b; HAMID *et al.*, 2013; KEYAERTS *et al.*, 2007; SOHRAB *et al.*, 2020).

7. Conclusão

Através das técnicas de bioinformática foi possível elucidar parcialmente a capacidade biotecnológica das lectinas da família Musaceae, demonstrando semelhanças de interação a carboidrato entre lectinas com a estrutura não resolvida e um modelo obtido por cristalografia. Os resultados obtidos no presente trabalho geraram dados que podem ser utilizadas para impulsionar diversas pesquisas relacionadas as lectinas da família Musaceae, onde apresentou dados acerca da interação a carboidrato e das possíveis atividades biológicas que essas lectinas podem desempenhar.

8. Referências

- ADIYAMAN, Recep; MCGUFFIN, Liam James. Methods for the refinement of protein structure 3D models. **International Journal of Molecular Sciences**, [S. l.], v. 20, n. 9, 2019. Disponível em: <https://doi.org/10.3390/ijms20092301>
- AJAY; MURCKO, Mark A. Computational Methods to Predict Binding Free Energy in Ligand-Receptor Complexes. **Journal of Medicinal Chemistry**, [S. l.], v. 38, n. 26, p. 4953–4967, 1995. Disponível em: <https://doi.org/10.1021/jm00026a001>. Acesso em: 1 jun. 2021.
- ALTSCHUL, Stephen F. *et al.* Basic local alignment search tool. **Journal of Molecular Biology**, [S. l.], v. 215, n. 3, p. 403–410, 1990. Disponível em: [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2). Acesso em: 15 ago. 2020.
- BAKER, D.; SALI, A. **Protein structure prediction and structural genomics**. [S. l.]: Science, 2001. Disponível em: <https://doi.org/10.1126/science.1065659>. Acesso em: 15 mar. 2021.
- BENKERT, Pascal; KÜNZLI, Michael; SCHWEDE, Torsten. QMEAN server for protein model quality estimation. **Nucleic Acids Research**, [S. l.], v. 37, n. SUPPL. 2, p. W510, 2009. Disponível em: <https://doi.org/10.1093/nar/gkp322>. Acesso em: 25 mar. 2021.
- BHATTACHARYA, Debswapna. RefinedD: Improved protein structure refinement using machine learning based restrained relaxation. **Bioinformatics**, [S. l.], v. 35, n. 18, p. 3320–3328, 2019. Disponível em: <https://doi.org/10.1093/bioinformatics/btz101>. Acesso em: 21 mar. 2021.
- BHUTIA, Sujit K. *et al.* **Plant lectins in cancer therapeutics: Targeting apoptosis and autophagy-dependent cell death**. [S. l.]: Academic Press, 2019. Disponível em: <https://doi.org/10.1016/j.phrs.2019.04.001>
- BHUTIA, Sujit K.; MALLICK, Sanjaya K.; MAITI, Tapas K. In vitro immunostimulatory properties of Abrus lectins derived peptides in tumor bearing mice. **Phytomedicine**, [S. l.], v. 16, n. 8, p. 776–782, 2009. Disponível em: <https://doi.org/10.1016/j.phymed.2009.01.006>
- BIENERT, Stefan *et al.* The SWISS-MODEL Repository-new features and functionality. **Nucleic Acids Research**, [S. l.], v. 45, n. D1, p. D313–D319, 2017. Disponível em: <https://doi.org/10.1093/nar/gkw1132>. Acesso em: 26 mar. 2021.
- BONNEAU, R.; BAKER, D. **Ab initio protein structure prediction: Progress and prospects**. [S. l.]: Annu Rev Biophys Biomol Struct, 2001. Disponível em: <https://doi.org/10.1146/annurev.biophys.30.1.173>. Acesso em: 15 mar. 2021.
- BROOIJMANS, Natasja; KUNTZ, Irwin D. **Molecular recognition and docking algorithms**. [S. l.]: Annu Rev Biophys Biomol Struct, 2003. Disponível em: <https://doi.org/10.1146/annurev.biophys.32.110601.142532>. Acesso em: 2 jun. 2021.
- CHEUNG, Allen H. K.; WONG, Jack H.; NG, T. B. Musa acuminata (Del Monte banana) lectin is a fructose-binding lectin with cytokine-inducing activity. **Phytomedicine**, [S. l.], v. 16, n. 6–7, p. 594–600, 2009. Disponível em: <https://doi.org/10.1016/j.phymed.2008.12.016>. Acesso em: 8 jul. 2020.
- CORBEIL, Christopher R.; WILLIAMS, Christopher I.; LABUTE, Paul. **Variability in docking success rates due to dataset preparation**. [S. l.]: Springer, 2012. Disponível em: <https://doi.org/10.1007/s10822-012-9570-1>. Acesso em: 3 abr. 2021.

COVÉS-DATSON, Evelyn M. *et al.* Targeted disruption of pi-pi stacking in Malaysian banana lectin reduces mitogenicity while preserving antiviral activity. **Cancer Biology Program**, [S. l.], v. 48109, n. 11, p. 12, [s. d.]. Disponível em: <https://doi.org/10.1038/s41598-020-80577-7>. Acesso em: 14 mar. 2021.

COVÉS-DATSON, Evelyn M. *et al.* Inhibition of Ebola Virus by a Molecularly Engineered Banana Lectin. **PLoS Neglected Tropical Diseases**, [S. l.], v. 13, n. 7, p. 1–20, 2019. Disponível em: <https://doi.org/10.1371/journal.pntd.0007595>

COVÉS-DATSON, Evelyn M. *et al.* A molecularly engineered antiviral banana lectin inhibits fusion and is efficacious against influenza virus infection in vivo. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 117, n. 4, p. 2122–2132, 2020. Disponível em: <https://doi.org/10.1073/pnas.1915152117>. Acesso em: 20 jul. 2020.

D'HONT, Angélique *et al.* The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. **Nature**, [S. l.], v. 488, n. 7410, p. 213–217, 2012. Disponível em: <https://doi.org/10.1038/nature11241>

DAMME, Els J. M. Van *et al.* Plant Lectins: A Composite of Several Distinct Families of Structurally and Evolutionary Related Proteins with Diverse Biological Roles. **Critical Reviews in Plant Sciences**, [S. l.], v. 17, n. 6, p. 575–692, 1998. Disponível em: <https://doi.org/10.1080/07352689891304276>. Acesso em: 30 abr. 2020.

DE CAMARGO, Laura Junqueira *et al.* Antiviral activity of native banana lectin against bovine viral diarrhea virus and bovine alphaherpesvirus type 1. **International Journal of Biological Macromolecules**, [S. l.], v. 157, p. 569–576, 2020. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2020.04.125>

DE MEJÍA, Elvira González; PRISECARU, Valentin I. Lectins as bioactive plant proteins: A potential in cancer treatment. **Critical Reviews in Food Science and Nutrition**, [S. l.], v. 45, n. 6, p. 425–445, 2005 a. Disponível em: <https://doi.org/10.1080/10408390591034445>

DE MEJÍA, Elvira González; PRISECARU, Valentin I. Lectins as bioactive plant proteins: A potential in cancer treatment. **Critical Reviews in Food Science and Nutrition**, [S. l.], v. 45, n. 6, p. 425–445, 2005 b. Disponível em: <https://doi.org/10.1080/10408390591034445>. Acesso em: 11 maio. 2021.

DENARO, Marcella *et al.* **Antiviral activity of plants and their isolated bioactive compounds: An update**. [S. l.]: John Wiley and Sons Ltd, 2020. Disponível em: <https://doi.org/10.1002/ptr.6575>. Acesso em: 14 maio. 2021.

DUTTA, N. K. *et al.* Anti-variola and antivaccinia principles from seeds of Banakadali (*Ensete superbum*, Cheesm, Musaceae). **Indian Journal of Medical Research**, [S. l.], v. 56, n. 5, p. 735–741, 1968.

FEIG, Michael. Computational protein structure refinement: almost there, yet still so far to go. **Wiley Interdisciplinary Reviews: Computational Molecular Science**, [S. l.], v. 7, n. 3, p. e1307, 2017. Disponível em: <https://doi.org/10.1002/wcms.1307>. Acesso em: 21 mar. 2021.

FEIG, Michael; MIRJALILI, Vahid. Protein structure refinement via molecular-dynamics simulations: What works and what does not? **Proteins: Structure, Function and Bioinformatics**, [S. l.], v. 84, n. S1, p. 282–292, 2016. Disponível em: <https://doi.org/10.1002/prot.24871>. Acesso em: 6 maio. 2021.

- FISER, András; DO, Richard Kinh Gian; ŠALI, Andrej. Modeling of loops in protein structures. **Protein Science**, [S. l.], v. 9, n. 9, p. 1753–1773, 2000. Disponível em: <https://doi.org/10.1110/ps.9.9.1753>. Acesso em: 25 mar. 2021.
- GAUTIER, Romain *et al.* HELIQUEST: A web server to screen sequences with specific α -helical properties. **Bioinformatics**, [S. l.], v. 24, n. 18, p. 2101–2102, 2008. Disponível em: <https://doi.org/10.1093/bioinformatics/btn392>. Acesso em: 25 mar. 2021.
- GOLDSTEIN, Irwin J. *et al.* Carbohydrate binding properties of banana (*Musa acuminata*) lectin: II. Binding of laminaribiose oligosaccharides and β -glucans containing β 1,6-glucosyl end groups. **European Journal of Biochemistry**, [S. l.], v. 268, n. 9, p. 2616–2619, 2001. Disponível em: <https://doi.org/10.1046/j.1432-1327.2001.02149.x>. Acesso em: 20 jul. 2020.
- GOLDSTEIN, Irwin J. Lectin structure-activity: The story is never over. *In*: 2002a, **Journal of Agricultural and Food Chemistry**. [S. l.: s. n.] p. 6583–6585. Disponível em: <https://doi.org/10.1021/jf0201879>
- GOLDSTEIN, Irwin J. Lectin structure-activity: The story is never over. *In*: 2002b, **Journal of Agricultural and Food Chemistry**. : J Agric Food Chem, 2002. p. 6583–6585. Disponível em: <https://doi.org/10.1021/jf0201879>. Acesso em: 29 mar. 2021.
- GUEX, Nicolas; PEITSCH, Manuel C.; SCHWEDE, Torsten. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. **ELECTROPHORESIS**, [S. l.], v. 30, n. S1, p. S162–S173, 2009. Disponível em: <https://doi.org/10.1002/elps.200900140>. Acesso em: 26 mar. 2021.
- HAMID, Rabia *et al.* Lectins: Proteins with Diverse Applications. **Journal of Applied Pharmaceutical Science**, [S. l.], v. 3, n. 4, p. 93–103, 2013. Disponível em: <https://doi.org/10.7324/JAPS.2013.34.S18>. Acesso em: 11 maio. 2021.
- HEGYI, Hedi; GERSTEIN, Mark. The relationship between protein structure and function: A comprehensive survey with application to the yeast genome. **Journal of Molecular Biology**, [S. l.], v. 288, n. 1, p. 147–164, 1999. Disponível em: <https://doi.org/10.1006/jmbi.1999.2661>
- HEO, Lim; PARK, Hahnbeom; SEOK, Chaok. GalaxyRefine: Protein structure refinement driven by side-chain repacking. **Nucleic acids research**, [S. l.], v. 41, n. Web Server issue, p. W384–W388, 2013. Disponível em: <https://doi.org/10.1093/nar/gkt458>. Acesso em: 5 abr. 2021.
- HILLISCH, Alexander; PINEDA, Luis Felipe; HILGENFELD, Rolf. **Utility of homology models in the drug discovery process**. [S. l.]: Elsevier Current Trends, 2004. Disponível em: [https://doi.org/10.1016/S1359-6446\(04\)03196-4](https://doi.org/10.1016/S1359-6446(04)03196-4)
- HIRANUMA, Naozumi *et al.* Improved protein structure refinement guided by deep learning based accuracy estimation. **Nature Communications**, [S. l.], v. 12, n. 1, p. 1–11, 2021. Disponível em: <https://doi.org/10.1038/s41467-021-21511-x>. Acesso em: 6 maio. 2021.
- IMAM, Mohammad Zafar; AKTER, Saleha. *Musa paradisiaca* l. and *musa sapientum* l. : A phytochemical and pharmacological review. **Journal of Applied Pharmaceutical Science**, [S. l.], v. 1, n. 5, p. 14–20, 2011.
- JAIN, Preeti *et al.* Antibacterial and antioxidant activities of local seeded banana fruits. **African Journal of Pharmacy and Pharmacology**, [S. l.], v. 5, n. 11, p. 1398–1403, 2011. Disponível em: <https://doi.org/10.5897/AJPP11.294>
- JONES, Gareth *et al.* Development and validation of a genetic algorithm for flexible docking.

Journal of Molecular Biology, [S. l.], v. 267, n. 3, p. 727–748, 1997 a. Disponível em: <https://doi.org/10.1006/jmbi.1996.0897>. Acesso em: 2 jun. 2021.

JONES, Gareth *et al.* Development and validation of a genetic algorithm for flexible docking. **Journal of Molecular Biology**, [S. l.], v. 267, n. 3, p. 727–748, 1997 b. Disponível em: <https://doi.org/10.1006/jmbi.1996.0897>. Acesso em: 3 abr. 2021.

KAISER, Annette *et al.* **The necessity to develop drugs against parasitic diseases**. [S. l.: s. n.] Disponível em: <https://europepmc.org/article/med/12611273>. Acesso em: 14 maio. 2021.

KAPADIA, Suraj Premal; PUDAKALKATTI, Pushpa S.; SHIVANAİKAR, Sachin. Detection of antimicrobial activity of banana peel (*Musa paradisiaca* L.) on *Porphyromonas gingivalis* and *Aggregatibacter actinomycetemcomitans*: An in vitro study. **Contemporary Clinical Dentistry**, [S. l.], v. 6, n. 4, p. 496–499, 2015. Disponível em: <https://doi.org/10.4103/0976-237X.169864>. Acesso em: 3 maio. 2021.

KEYAERTS, Els *et al.* Plant lectins are potent inhibitors of coronaviruses by interfering with two targets in the viral replication cycle. **Antiviral Research**, [S. l.], v. 75, n. 3, p. 179–187, 2007. Disponível em: <https://doi.org/10.1016/j.antiviral.2007.03.003>

KOSHTE, V. L. *et al.* Isolation and characterization of BanLec-I, a mannoside-binding lectin from *Musa paradisiaca* (banana). **Biochemical Journal**, [S. l.], v. 272, n. 3, p. 721–726, 1990. Disponível em: <https://doi.org/10.1042/bj2720721>

KOSHTE, V. L. *et al.* The potent IgG4-Inducing antigen in banana is a mannose-binding lectin, BanLec-I. **International Archives of Allergy and Immunology**, [S. l.], v. 97, n. 1, p. 17–24, 1992. Disponível em: <https://doi.org/10.1159/000236090>

LANNOO, Nausicaä; VAN DAMME, Els J. M. **Lectin domains at the frontiers of plant defense**. [S. l.]: Frontiers Research Foundation, 2014. Disponível em: <https://doi.org/10.3389/fpls.2014.00397>

LASKOWSKI, R. A. *et al.* PROCHECK: a program to check the stereochemical quality of protein structures. **Journal of Applied Crystallography**, [S. l.], v. 26, n. 2, p. 283–291, 1993. Disponível em: <https://doi.org/10.1107/s0021889892009944>

LASKOWSKI, Roman A. *et al.* **PDBsum: A Web-based database of summaries and analyses of all PDB structures**. [S. l.]: Elsevier Ltd, 1997. Disponível em: [https://doi.org/10.1016/S0968-0004\(97\)01140-7](https://doi.org/10.1016/S0968-0004(97)01140-7). Acesso em: 25 mar. 2021.

LESK, Arthur M. **Introduction to bioinformatics**. 4^o ed. [S. l.: s. n.]. *E-book*. Disponível em: <https://doi.org/10.1201/b13091-1>

LI, Wenlin *et al.* Estimation of Uncertainties in the Global Distance Test (GDT_TS) for CASP Models. **PLOS ONE**, [S. l.], v. 11, n. 5, p. e0154786, 2016. Disponível em: <https://doi.org/10.1371/journal.pone.0154786>. Acesso em: 26 mar. 2021.

LIU, Xiao Yu; LI, Huan; ZHANG, Wei. The lectin from *Musa paradisiaca* binds with the capsid protein of tobacco mosaic virus and prevents viral infection. **Biotechnology and Biotechnological Equipment**, [S. l.], v. 28, n. 3, p. 408–416, 2014. Disponível em: <https://doi.org/10.1080/13102818.2014.925317>

LOVELL, Simon C. *et al.* Structure validation by C α geometry: ϕ, ψ and C β deviation. **Proteins: Structure, Function and Genetics**, [S. l.], v. 50, n. 3, p. 437–450, 2003. Disponível em: <https://doi.org/10.1002/prot.10286>

- MAGHRABI, Ali H. A.; MCGUFFIN, Liam J. ModFOLD6: An accurate web server for the global and local quality estimation of 3D protein models. **Nucleic Acids Research**, [S. l.], v. 45, n. W1, p. W416–W421, 2017. Disponível em: <https://doi.org/10.1093/nar/gkx332>. Acesso em: 25 mar. 2021.
- MEAGHER, Jennifer L. *et al.* Crystal structure of banana lectin reveals a novel second sugar binding site. **Glycobiology**, [S. l.], v. 15, n. 10, p. 1033–1042, 2005. Disponível em: <https://doi.org/10.1093/glycob/cwi088>. Acesso em: 3 abr. 2021.
- MENG, Xuan-Yu *et al.* Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. **Current Computer Aided-Drug Design**, [S. l.], v. 7, n. 2, p. 146–157, 2012. Disponível em: <https://doi.org/10.2174/157340911795677602>
- MITCHELL, Carter A.; RAMESSAR, Koreen; O'KEEFE, Barry R. Antiviral lectins: Selective inhibitors of viral entry. **Antiviral Research**, [S. l.], v. 142, p. 37–54, 2017. Disponível em: <https://doi.org/10.1016/j.antiviral.2017.03.007>
- MO, Hanqing *et al.* Carbohydrate binding properties of banana (*Musa acuminata*) lectin. **European Journal of Biochemistry**, [S. l.], v. 268, n. 9, p. 2609–2615, 2001. Disponível em: <https://doi.org/10.1046/j.1432-1327.2001.02148.x>. Acesso em: 17 jul. 2020.
- MORRIS, Garrett M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, [S. l.], v. 30, n. 16, p. 2785–2791, 2009. Disponível em: <https://doi.org/10.1002/jcc.21256>. Acesso em: 26 mar. 2021.
- MORRIS, Garrett M.; LIM-WILBY, Marguerita. Molecular docking. **Methods in Molecular Biology**, [S. l.], v. 443, p. 365–382, 2008. Disponível em: https://doi.org/10.1007/978-1-59745-177-2_19. Acesso em: 22 mar. 2021.
- NASCIMENTO, Kyria Santiago *et al.* Homology modeling, molecular docking, and dynamics of two α -methyl-d-mannoside-specific lectins from *Arachis* genus. **Journal of Molecular Modeling**, [S. l.], v. 24, n. 9, 2018. Disponível em: <https://doi.org/10.1007/s00894-018-3800-y>
- NECO, Antonio Hadson Bastos *et al.* Structural analysis, molecular docking and molecular dynamics of an edematogenic lectin from *Centrolobium microchaete* seeds. **International Journal of Biological Macromolecules**, [S. l.], v. 117, p. 124–133, 2018. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2018.05.166>
- ORENGO, Christine A.; TODD, Annabel E.; THORNTON, Janet M. From protein structure to function. **Current Opinion in Structural Biology**, [S. l.], v. 9, n. 3, p. 374–382, 1999. Disponível em: [https://doi.org/10.1016/S0959-440X\(99\)80051-7](https://doi.org/10.1016/S0959-440X(99)80051-7). Acesso em: 8 fev. 2021.
- PAGADALA, Nataraj S.; SYED, Khajamohiddin; TUSZYNSKI, Jack. Software for molecular docking: a review. **Biophysical Reviews**, [S. l.], v. 9, n. 2, p. 91–102, 2017. Disponível em: <https://doi.org/10.1007/s12551-016-0247-1>
- PANDA, Sujogya Kumar *et al.* Antiviral and Cytotoxic Activity of Different Plant Parts of Banana (*Musa spp.*). **Viruses**, [S. l.], v. 12, n. 5, p. 549, 2020 a. Disponível em: <https://doi.org/10.3390/v12050549>. Acesso em: 4 mar. 2021.
- PANDA, Sujogya Kumar *et al.* Antiviral and Cytotoxic Activity of Different Plant Parts of Banana (*Musa spp.*). **Viruses**, [S. l.], v. 12, n. 5, p. 549, 2020 b. Disponível em: <https://doi.org/10.3390/v12050549>. Acesso em: 3 maio. 2021.

- PEITSCH, M. C. Large scale protein modelling and model repository. **Proceedings. International Conference on Intelligent Systems for Molecular Biology**, [S. l.], v. 5, p. 234–236, 1997. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/9322042>. Acesso em: 15 ago. 2020.
- PEREIRA, Aline; MARASCHIN, Marcelo. **Banana (Musa spp) from peel to pulp: Ethnopharmacology, source of bioactive compounds and its relevance for human health**. [S. l.]: Elsevier Ireland Ltd, 2015. Disponível em: <https://doi.org/10.1016/j.jep.2014.11.008>
- PEUMANS, Willy J. *et al.* Fruit-specific lectins from banana and plantain. **Planta**, [S. l.], v. 211, n. 4, p. 546–554, 2000 a. Disponível em: <https://doi.org/10.1007/s004250000307>
- PEUMANS, Willy J. *et al.* Fruit-specific lectins from banana and plantain. **Planta**, [S. l.], v. 211, n. 4, p. 546–554, 2000 b. Disponível em: <https://doi.org/10.1007/s004250000307>. Acesso em: 3 abr. 2021.
- PEUMANS, Willy J.; VAN DAMME, Els J. M. **Lectins as Plant Defense Proteins** 'Archives of Microbiology'. [S. l.: s. n.].
- PINTO-JUNIOR, Vanir Reis *et al.* Molecular modeling, docking and dynamics simulations of the Dioclea lasiophylla Mart. Ex Benth seed lectin: An edematogenic and hypernociceptive protein. **Biochimie**, [S. l.], v. 135, p. 126–136, 2017 a. Disponível em: <https://doi.org/10.1016/j.biochi.2017.02.002>
- PINTO-JUNIOR, Vanir Reis *et al.* Structural studies of a vasorelaxant lectin from Dioclea reflexa Hook seeds: Crystal structure, molecular docking and dynamics. **International Journal of Biological Macromolecules**, [S. l.], v. 98, p. 12–23, 2017 b. Disponível em: <https://doi.org/10.1016/j.ijbiomac.2017.01.092>
- PLOETZ RANDY *et al.* Banana and plantain – an overview with emphasis on the Pacific Island cultivars. **Permanent Agriculture Resources (PAR)**, [S. l.], 2007. Disponível em: https://www.researchgate.net/publication/228645617_Banana_and_plantain_-_an_overview_with_emphasis_on_the_Pacific_Island_cultivars. Acesso em: 12 mar. 2021.
- R. RICHARD BUEGESS; MURRAY P. DEUTSCHER. **Guide to Protein Purification** . [S. l.: s. n.]. *E-book*. Disponível em: <https://books.google.com.br/books?id=f6Lp4yna4hoC&printsec=frontcover&hl=pt-BR#v=onepage&q&f=false>. Acesso em: 2 mar. 2021.
- RAMACHANDRAN, G. N.; RAMAKRISHNAN, C.; SASISEKHARAN, V. **Stereochemistry of polypeptide chain configurations**. [S. l.]: J Mol Biol, 1963. Disponível em: [https://doi.org/10.1016/S0022-2836\(63\)80023-6](https://doi.org/10.1016/S0022-2836(63)80023-6). Acesso em: 4 jun. 2021.
- ROBERTO, Paulo *et al.* **Genética Geral Para Universitários**. [S. l.: s. n.]. *E-book*. Disponível em: https://www.repository.ufrpe.br/bitstream/123456789/2355/1/livro_geneticageralweb.pdf. Acesso em: 10 maio. 2021.
- SANTOS FILHO, Osvaldo Andrade; BICCA DE ALENCASTRO, Ricardo. Protein homology modeling. **Química Nova**, [S. l.], v. 26, n. 2, p. 253–259, 2003. Disponível em: <https://doi.org/10.1590/s0100-40422003000200019>. Acesso em: 13 mar. 2021.
- SAROWAR HOSSAIN, M. *et al.* Antidiarrheal, antioxidant and antimicrobial activities of the musa sapientum seed. **Avicenna Journal of Medical Biotechnology**, [S. l.], v. 3, n. 2, p. 95–105, 2011.

SHARMA, Alok; VIJAYAN, Mamannamana. Influence of glycosidic linkage on the nature of carbohydrate binding in β -prism I fold lectins: An X-ray and molecular dynamics investigation on banana lectin-carbohydrate complexes. **Glycobiology**, [S. l.], v. 21, n. 1, p. 23–33, 2011 a. Disponível em: <https://doi.org/10.1093/glycob/cwq128>. Acesso em: 3 maio. 2021.

SHARMA, Alok; VIJAYAN, Mamannamana. Influence of glycosidic linkage on the nature of carbohydrate binding in β -prism I fold lectins: An X-ray and molecular dynamics investigation on banana lectin-carbohydrate complexes. **Glycobiology**, [S. l.], v. 21, n. 1, p. 23–33, 2011 b. Disponível em: <https://doi.org/10.1093/glycob/cwq128>. Acesso em: 8 mar. 2021.

SHARMA ALOK; VIJAYAN MAMANNAMANA. Influence of glycosidic linkage on the nature of carbohydrate binding in -prism I fold lectins: An X-ray and molecular dynamics investigation on banana lectin-carbohydrate complexes. **Glycobiology**, [S. l.], 2011. Disponível em: https://www.researchgate.net/publication/45799778_Influence_of_glycosidic_linkage_on_the_nature_of_carbohydrate_binding_in_-prism_I_fold_lectins_An_X-ray_and_molecular_dynamics_investigation_on_banana_lectin-carbohydrate_complexes. Acesso em: 17 jul. 2020.

SINGH, D. D. *et al.* Unusual sugar specificity of banana lectin from *Musa paradisiaca* and its probable evolutionary origin. Crystallographic and modelling studies. **Glycobiology**, [S. l.], v. 15, n. 10, p. 1025–1032, 2005. Disponível em: <https://doi.org/10.1093/glycob/cwi087>. Acesso em: 12 nov. 2020.

SINGH, Senjam Sunil; DEVI, Sanjenbam Kunjeshwori; NG, Tzi Bun. Banana lectin: A brief review. **Molecules**, [S. l.], v. 19, n. 11, p. 18817–18827, 2014. Disponível em: <https://doi.org/10.3390/molecules191118817>

SOHRAB, Sayed S. *et al.* The Emergence of Human Pathogenic Coronaviruses: Lectins as Antivirals for SARS-CoV-2. **Current Pharmaceutical Design**, [S. l.], v. 26, n. 41, p. 5286–5292, 2020. Disponível em: <https://doi.org/10.2174/1381612826666200821120409>

SOUZA, Maria A. *et al.* **The immunomodulatory effect of plant lectins: A review with emphasis on ArtinM properties**. [S. l.]: Glycoconj J, 2013. Disponível em: <https://doi.org/10.1007/s10719-012-9464-4>. Acesso em: 13 maio. 2021.

SRINIVAS, Balaji Kyathgowdanadoddi; SHIVAMADHU, Madhu Chakkere; JAYARAMA, Shankar. Angio-Suppressive Effect of Partially Purified Lectin-like Protein from *Musa acuminata* pseudostem by Inhibition of VEGF-Mediated Neovascularization and Induces Apoptosis Both In Vitro and In Vivo. **Nutrition and Cancer**, [S. l.], v. 71, n. 2, p. 285–300, 2019. Disponível em: <https://doi.org/10.1080/01635581.2018.1540714>

STOJANOVIĆ, Marijana M. *et al.* In vitro stimulation of Balb/c and C57 BL/6 splenocytes by a recombinantly produced banana lectin isoform results in both a proliferation of T cells and an increased secretion of interferon-gamma. **International Immunopharmacology**, [S. l.], v. 10, n. 1, p. 120–129, 2010. Disponível em: <https://doi.org/10.1016/j.intimp.2009.10.007>

SWANSON, Michael D. *et al.* A lectin isolated from bananas is a potent inhibitor of HIV replication. **Journal of Biological Chemistry**, [S. l.], v. 285, n. 12, p. 8646–8655, 2010. Disponível em: <https://doi.org/10.1074/jbc.M109.034926>

SWANSON, Michael D. *et al.* Engineering a Therapeutic Lectin by Uncoupling Mitogenicity

from Antiviral Activity. **Cell**, [S. l.], v. 163, n. 3, p. 746–758, 2015. Disponível em: <https://doi.org/10.1016/j.cell.2015.09.056>

TROTT, Oleg; OLSON, Arthur J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. **Journal of Computational Chemistry**, [S. l.], v. 31, n. 2, p. NA-NA, 2009. Disponível em: <https://doi.org/10.1002/jcc.21334>. Acesso em: 15 ago. 2020.

VAN DAMME, Els J. M. Lectins as tools to select for glycosylated proteins. **Methods in Molecular Biology**, [S. l.], v. 753, p. 289–297, 2011. Disponível em: https://doi.org/10.1007/978-1-61779-148-2_19

VAN DAMME, Els J. M. **History of plant lectin research**. [S. l.]: Humana Press Inc., 2014. Disponível em: https://doi.org/10.1007/978-1-4939-1292-6_1

VASTA, Gerardo R. *et al.* **Galectins as self/non-self recognition receptors in innate and adaptive immunity: An unresolved paradox**. [S. l.]: Frontiers Media SA, 2012. Disponível em: <https://doi.org/10.3389/fimmu.2012.00199>. Acesso em: 13 mar. 2021.

WANG, Zhe *et al.* Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: The prediction accuracy of sampling power and scoring power. **Physical Chemistry Chemical Physics**, [S. l.], v. 18, n. 18, p. 12964–12975, 2016. Disponível em: <https://doi.org/10.1039/c6cp01555g>. Acesso em: 2 abr. 2021.

WATERHOUSE, Andrew *et al.* SWISS-MODEL: Homology modelling of protein structures and complexes. **Nucleic Acids Research**, [S. l.], v. 46, n. W1, p. W296–W303, 2018. Disponível em: <https://doi.org/10.1093/nar/gky427>. Acesso em: 26 mar. 2021.

WHITTINGTON, Ruth; FAULDS, Diana. Interleukin-2: A Review of its Pharmacological Properties and Therapeutic Use in Patients with Cancer. **Drugs**, [S. l.], v. 46, n. 3, p. 446–514, 1993. Disponível em: <https://doi.org/10.2165/00003495-199346030-00009>. Acesso em: 15 mar. 2021.

WONG, Jack Ho; NG, T. B. Isolation and characterization of a glucose/mannose-specific lectin with stimulatory effect on nitric oxide production by macrophages from the emperor banana. **International Journal of Biochemistry and Cell Biology**, [S. l.], v. 38, n. 2, p. 234–243, 2006 a. Disponível em: <https://doi.org/10.1016/j.biocel.2005.09.004>

WONG, Jack Ho; NG, T. B. Isolation and characterization of a glucose/mannose-specific lectin with stimulatory effect on nitric oxide production by macrophages from the emperor banana. **International Journal of Biochemistry and Cell Biology**, [S. l.], v. 38, n. 2, p. 234–243, 2006 b. Disponível em: <https://doi.org/10.1016/j.biocel.2005.09.004>. Acesso em: 8 jul. 2020.

WRIGHT, Peter E.; DYSON, H. Jane. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. **Journal of Molecular Biology**, [S. l.], v. 293, n. 2, p. 321–331, 1999. Disponível em: <https://doi.org/10.1006/jmbi.1999.3110>

XU, Jinrui; ZHANG, Yang. How significant is a protein structure similarity with TM-score = 0.5? **Bioinformatics**, [S. l.], v. 26, n. 7, p. 889–895, 2010. Disponível em: <https://doi.org/10.1093/bioinformatics/btq066>. Acesso em: 25 mar. 2021.

ZEMLA, Adam. LGA: A method for finding 3D similarities in protein structures. **Nucleic Acids Research**, [S. l.], v. 31, n. 13, p. 3370–3374, 2003. Disponível em: <https://doi.org/10.1093/nar/gkg571>. Acesso em: 26 mar. 2021.

ZHANG, Yang. I-TASSER server for protein 3D structure prediction. **BMC Bioinformatics**, [S. l.], v. 9, n. 1, p. 1–8, 2008. Disponível em: <https://doi.org/10.1186/1471-2105-9-40>. Acesso em: 15 mar. 2021.