



**UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA  
MOLECULAR**

**GABRIELA PATRÍCIA MARTINS DE ALMEIDA BERNARDES**

**IDENTIFICAÇÃO E ANÁLISE *IN SILICO* DE LECTINAS DO PEIXE  
PULMONADO SUL-AMERICANO (*Lepidosirenparadoxa*)  
BASEADO EM TRANSCRIPTOMAS**

**BELÉM/PA**

**2023**

**GABRIELA PATRÍCIA MARTINS DE ALMEIDA BERNARDES**

**IDENTIFICAÇÃO E ANÁLISE *IN SILICO* DE LECTINAS DO PEIXE  
PULMONADO SUL-AMERICANO (*Lepidosirenparadoxa*)  
BASEADO EM TRANSCRIPTOMAS**

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da Universidade Federal do Pará, como requisito parcial para obtenção do grau de Doutor em Genética e Biologia Molecular.

Orientador: Prof.<sup>a</sup> Dra. Maria Paula Cruz Schneider

Co-orientador: Prof. Dr. Agenor Valadares Santos

**BELÉM/PA**

**2023**



## **INSTITUIÇÕES E FONTES FINANCIADORAS**

- Universidade Federal do Pará
- Instituto de Ciências Biológicas
- Centro de Genômica e Biologia do Sistema
- Laboratório de Genômica e Biotecnologia
- Laboratório de Biotecnologia de Enzimas e Biotransformações
- Conselho Nacional de Pesquisa, CNPq
- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, CAPES
- Financiadora de Estudos e Projetos, Finep

Dedico a Deus e a minha família, razão  
de minha existência.

## **AGRADECIMENTOS**

# SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>13</b>
1.1 AS LECTINAS .....	14
1.2 LECTINAS EM PEIXES.....	16
1.3 CLASSIFICAÇÃO DAS LECTINAS EM PEIXES.....	18
<b>1.3.1 Galectinas</b> .....	<b>19</b>
<b>1.3.2 Lectinas Tipo-C</b> .....	<b>21</b>
<b>1.3.3 Lectina ligadora de L-ramnose</b> .....	<b>24</b>
1.4 ATIVIDADE BIOLÓGICA DAS LECTINAS DE PEIXES.....	25
<b>2 OBJETIVOS</b> .....	<b>35</b>
2.1 OBJETIVO GERAL .....	35
2.2 OBJETIVOS ESPECÍFICOS .....	35
<b>3 METODOLOGIA</b> .....	<b>36</b>
3.1 ANIMAIS E AMOSTRAS.....	36
3.2 DADOS TRANSCRIPTÔMICOS .....	36
3.3 ANÁLISES COMPUTACIONAIS.....	36
<b>3.3.1 Análises de sequências</b> .....	<b>36</b>
<b>3.3.2 Caracterização das lectinas</b> .....	<b>37</b>
3.4 MODELAGEM MOLECULAR POR HOMOLOGIA.....	37
3.5 DOCKING MOLECULAR .....	38
3.6 ANOTAÇÃO FUNCIONAL .....	38
<b>4 RESULTADOS</b> .....	<b>39</b>
4.1 IDENTIFICAÇÃO DAS LECTINAS DE PIRAMBOIA .....	39
4.2 CARACTERIZAÇÃO DAS LECTINAS DE PIRAMBOIA.....	40
<b>4.2.1 Lectinas do tipo C</b> .....	<b>40</b>
<b>4.2.4 Intelectinas</b> .....	<b>47</b>
<b>4.2.3 Galectinas</b> .....	<b>50</b>

4.3 ESTUDO DE DOCKING MOLECULAR.....	57
<b>5 DISCUSSÃO .....</b>	<b>62</b>
<b>6 CONCLUSÃO .....</b>	<b>71</b>
<b>REFERÊNCIAS.....</b>	<b>72</b>

## LISTA DE SIGLAS

AH: Atividade hemaglutinante;

AHE: Atividade hemaglutinante específica;

AJL-1: Lectina da *Anguilla japonica* 1;

AJL-2: Lectina da *A. japonica* 2;

BLAST: Basic Local Aligment Search Tools;

Ca<sup>2+</sup>: Cálcio;

CaCl<sub>2</sub>: Cloreto de cálcio;

CLTD: Domínio semelhante a lectina tipo C;

ConCL-s: Lectina da pele de *Congermyrister*;

CRD: Domínio de Reconhecimento de Carboidratos;

CRD RBL: Domínio tipo lectina ligadora de L-ramnose;

CSL: Lectina de Salmão Cão;

CTL: Lectina tipo-C;

DDD: Dispositivos de entrega de drogas;

DNA: Ácido Desoxirribonucleico;

EDTA: Ácido Etilenodiamino Tetra-acético;

EPN: Resíduos de ácido glutâmico, prolina e asparagina;

IAH: Inibição da Atividade Hemaglutinante;

LaBEB: Laboratório de Biotecnologia de Enzimas e Biotransformações;

MEGA: Análise Genética da Evolução Molecular (do inglês, Molecular Evolution GeneticAnalysis);

MnCl<sub>2</sub>: Cloreto de manganês;

NaCl: Cloreto de sódio;

NaOH: Hidróxido de sódio;

NCBI: Centro Nacional de Informações sobre Biotecnologia (do inglês, National Center for Biotechnology Information);

NCCLS: Comitê Nacional para Padrões de Laboratório Clínico (do Inglês, National Committee for Clinical Laboratory Standards);

PCR: Reação em Cadeia de Polimerase;

pH: Potencial hidrogeniônico;

POLR1C: RNA Polimerase I e III, subunidade C

PFL-1: Lectina de peixe pônei 1;

PFL-2: Lectina de peixe pônei 2;

PRR: Receptores de Reconhecimento de Patógenos;

qPCR: PRC em Tempo Real;

QPD: Resíduos de glutamina, prolina e ácido aspártico;

RBL: Lectina ligadora de L-ramnose;

RNA: Ácido Ribonucleico;

Sisbio: Sistema de Autorização e Informação em Biodiversidade

SMART: Ferramenta de Pesquisa de Arquitetura Modular Simples (do inglês, Simple Modular Architecture Research Tool)

STL: Lectina de *Salvelinus leucomaenis*;

Tris/HCL: Tris(hidroximetil)aminometano

WND: Resíduos de triptofano, asparagina e ácido aspártico;

## LISTA DE FIGURAS

Figura 1 - Representação esquemática da interação da lectina com um carboidrato. .....	15
Figura 2 - Foto da Piramboia ( <i>Lepidosiren paradoxa</i> ) .....	32
Figura 3 – Genes de lectinas expressos nos transcriptomas de piramboia durante a regeneração do tecido.....	39
Figura 4 - Esquema da estrutura de <i>LpCTL</i> .....	41
Figura 5 - Figura 5 - Estruturas tridimensionais de <i>LpCTL-1</i> , suas variantes e comparações.....	42
Figura 6 – Alinhamento da sequência de piramboia, bovina, humana, suína e rato de CD302. ....	43
Figura 7 - Esquema da estrutura de <i>LpCD302</i> .....	44
Figura 8 – Estrutura Tridimensional de <i>LpCD302</i> .....	45
Figura 9 - Alinhamento da sequência de p-selectina de piramboia, rato, primatas e humana. Esquema da estrutura da <i>LpSELP</i> . ....	46
Figura 10 - Estrutura tridimensionais de <i>LpSELP</i> .....	47
Figura 11 - Alinhamento das variantes de <i>LpITLN2</i> com a <i>ITLN2</i> da truta arco-íris ( <i>Oncorhynchus mykiss</i> ). ....	48
Figura 12 - Estruturas tridimensionais de <i>LpITLN2</i> , suas variantes e comparações. ....	49
Figura 13 – Alinhamento da sequência de piramboia, búfalo, ovelha, hamster e rato de <i>LpGal-1</i> .....	51
Figura 14 – Esquema da estrutura da <i>LpGal-1</i> . ....	51
Figura 15 – Estrutura tridimensional de <i>LpGal-1</i> .....	52
Figura 16 – Alinhamento de sequências de <i>LpGal-9</i> de piramboia com camundongos, bovina e .....	53
Figura 17 – Esquema da estrutura de <i>LpGal-9</i> .....	54

Figura 18 – Estrutura tridimensional de <i>LpGal-9</i> .....	54
Figura 19 – Sequência de aminoácido de <i>LpGal</i> .....	55
Figura 20 – Estrutura esquemática de <i>LpGal</i> .....	56
Figura 21– Estrutura tridimensional de <i>LpGal</i> . ....	57
Figura 22 – Docking molecular entre as lectinas e carboidratos. ....	58
Figura 23 – Docking molecular entre as lectinas e LPS/Poli (I:C). ....	58
Figura 24 – Detalhe da ligação entre a lectinas <i>LpITLN2-F</i> e o ligante Poli (I:C). ....	58
Figura 25 – Análise dos processos biológicos das lectinas da piramboia .....	60
Figura 26 – Análise funcional das lectinas da piramboia.....	60
Figura 27 – Análise dos compartimentos celulares das lectinas da piramboia .....	61

## LISTA DE TABELAS

Tabela 1 - Atividade biológica das lectinas de peixes. ....	26
Tabela 2 – Genes de Galectinas no baiacu ( <i>Takifugu rubripes</i> ).....	29
Tabela 3 - Genes de lectinas identificados no transcriptoma da pirambóia ( <i>Lepidosiren paradoxa</i> ) .....	33

## RESUMO

As lectinas são proteínas que possuem a capacidade de se ligarem reversivelmente a carboidratos, apresentando uma ampla variedade de atividades biológicas. Essas proteínas estão presentes em quase todos os organismos vivos. Dentre as atividades biológicas que as lectinas podem apresentar destaca-se a atividade antimicrobiana, que consiste na capacidade de inibir o crescimento de microrganismos. Diversas lectinas foram isoladas e posteriormente caracterizadas de vários peixes. A piramboia (*Lepidosiren paradoxa*) é uma espécie de peixe pulmonado que habita as águas do continente Sul-Americano. Podendo ser encontrado nos pântanos e lagoas da região Amazônica no Brasil. Estes peixes possuem o corpo alongado, dois pares de nadadeira lobadas e o corpo coberto de muco que confere proteção ao animal. A presença de lectinas no *L. paradoxa* foi identificada a partir de dados do transcriptoma deste peixe. O presente estudo teve como objetivo identificar os genes de lectina do peixe *L. paradoxa* e analisar *in silico* a capacidade de suas proteínas atuarem como receptores de reconhecimento padrão (PRRs). Dentre os genes diferencialmente expresso nos transcriptomas da piramboia, doze proteínas pertencentes as famílias de lectina do tipo C, intelectinas e galectinas tiveram suas estruturas tridimensionais modeladas. Todas as lectinas apresentaram capacidade de se ligarem a carboidratos. Ademais, as lectinas demonstraram interagirem com ambos os polissacarídeos de superfície, LPS e Poly I:C, com exceção de *LpITLN2-C* que não demonstrou capacidade de se ligar ao Poly I:C. De uma maneira geral as lectinas apresentaram maior afinidade com Poly I:C do que com LPS. Esses achados sugerem que as lectinas de piramboia podem desempenhar um papel importante na defesa imune a patógenos.

**Palavras-Chaves:** Peixe pulmonado; Imunidade; Lectina; Estrutura.

## ABSTRACT

Lectins are proteins that have the ability to reversibly bind to carbohydrates, presenting a wide variety of biological activities. These proteins are present in almost all living organisms. Among the biological activities that lectins may present, antimicrobial activity stands out, which consists of the ability to inhibit the growth of microorganisms. Several lectins were isolated and further characterized from various fish. The piramboia (*Lepidosiren paradoxa*) is a species of lungfish that inhabits the waters of the South American continent. It can be found in the swamps and lakes of the Amazon region in Brazil. These fish have an elongated body, two pairs of lobed fins and a body covered in mucus that protects the animal. The presence of lectins in *L. paradoxa* was identified from transcriptome data from this fish. The present study aimed to identify the lectin genes of the fish *L. paradoxa* and to analyze in silico the ability of its proteins to act as pattern recognition receptors (PRRs). Among the genes differentially expressed in the piramboia transcriptomes, twelve proteins belonging to the C-type lectin, intelectin and galectin families had their three-dimensional structures modeled. All lectins showed the ability to bind to carbohydrates. Furthermore, lectins have been shown to interact with both surface polysaccharides, LPS and Poly I:C, except for LpITLN2-C which has not shown the ability to bind to Poly I:C. In general, lectins showed greater affinity with Poly I:C than with LPS. These findings suggest that piramboia lectins may play an important role in immune defense against pathogens.

**Keywords:** Lungfish; Immunity; Lectin; Structural.

## 1 INTRODUÇÃO

O sistema imunitário é o complexo responsável pela manutenção do equilíbrio do organismo vivo em face a uma ameaça externa ou uma perturbação metabólica. De acordo com Abbas (ABBAS; LICHTMAN; PILLAI, 2019), o sistema imune é constituído por células e moléculas que tem por função a defesa do organismo contra microrganismos infecciosos. Como se pode verificar, o sistema imune é aplicado em vários processos biológicos, através de um grupo de moléculas que trabalham coordenadas para manutenção da vida no organismo.

O sistema imunológico compreende estruturas e processos biológicos cuja função é proteger o organismo de doenças e agentes patogênicos. Todos os organismos vivos, mesmo os organismos unicelulares, possuem um sistema de defesa ou imunitário (SMITH; RISE; CHRISTIAN, 2019). Desde sob a forma de enzimas até os mecanismos mais complexos e evoluídos de defesa. Os vertebrados mandibulados apresentam duas estratégias de defesa: a resposta imune inata e a resposta imune adaptativa (SMITH; RISE; CHRISTIAN, 2019).

Na primeira, temos uma resposta imediata, porém não específica. Enquanto na última, temos uma resposta mais específica, como o próprio nome diz, adaptada ao tipo de patógeno identificado. Contudo a resposta imune adaptativa é ativada através da resposta imune inata (SOMPAYRAC, 2016). É na resposta imune inata onde podemos encontrar uma diversidade de mecanismos e estratégias para identificar o patógeno e deter a infecção.

Dentre esses mecanismos de defesa as lectinas são elementos chaves que desempenham papel importante na detecção e no reconhecimento dos patógenos e na resposta imune inata para deter a invasão de microrganismos patogênicos (RIZZATO, 2009). O entendimento da interação patógeno-hospedeiro é de extrema relevância, pois dessa interação reconhecemos os fatores que desempenharão um papel crucial no desenvolvimento das manifestações clínicas, virulência do patógeno, estado imune e variâncias gênicas do hospedeiro.

As lectinas funcionam como receptores de reconhecimento padrão (PRR), induzindo a proliferação de linfócitos, opsonização, fagocitose ou ativação do

complemento, através do reconhecimento padrões moleculares associados a patógenos (PAMP), os quais são expressos por um amplo espectro de agentes infecciosos (LINO et al., 2015); (MACIEL CARVALHO et al., 2012). A ativação da imunidade inata através da associação destas proteínas com moléculas expressas na superfície celular dos patógenos é o primeiro passo no desenvolvimento da imunidade adquirida contra antígenos específicos (RIZZATO, 2009).

## 1.1 As lectinas

Segundo Rodwell, Bender *et al.* (2021) as lectinas são proteínas capazes de aglutinar células através de sua interação específica e transitória com os carboidratos presentes na superfície celular. Como bem nos asseguram Nelson e Cox (2018), as lectinas são proteínas capazes de reconhecer os arranjos de carboidratos e participam de muitos processos biológicos, como no combate às infecções em gerais, reconhecimento, sinalização e adesão celular.

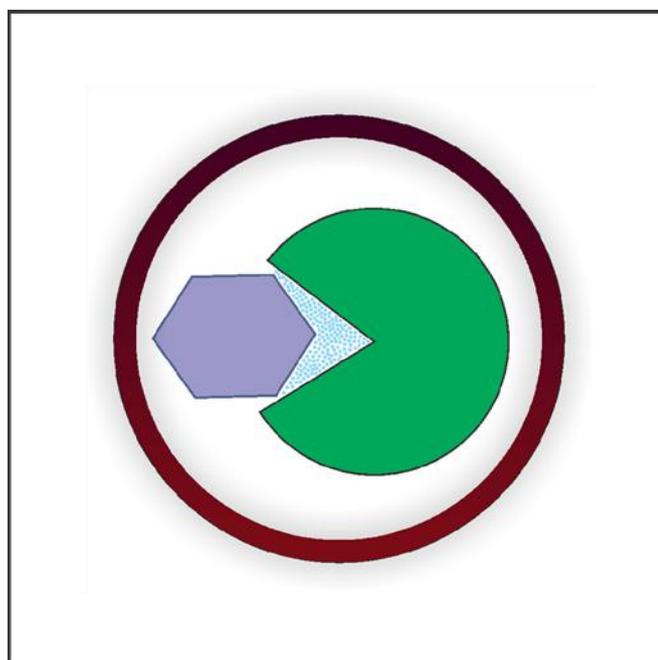
Para Zhang, Xiao-Wen *et al.* (2020) as lectinas desempenham papéis importantes, tanto na imunidade inata quanto na adaptativa. Na imunidade inata, por exemplo, essa molécula funciona como receptores de reconhecimento de padrões (PRRs, do inglês Pattern Recognition Receptors), induzindo a proliferação de linfócitos, opsonização, fagocitose ou ativação do complemento. Essas proteínas identificam células estranhas como "não-próprias" através do reconhecimento de carboidratos expressos na superfície celular dos patógenos e subsequentemente induzem sua destruição por ativação do complemento e/ou células fagocitárias (BRINCHMANN et al., 2018).

Por esses motivos que as lectinas têm despertado o interesse de muitos pesquisadores, que veem nessas moléculas um grande potencial biotecnológico. Uma vez que elas podem ser utilizadas como receptores de reconhecimento moleculares, ativadoras do sistema complemento, inibidores celulares, adjuvantes no processo de detecção e separação de glicoproteínas entre outras aplicações.

As lectinas reconhecem e interagem de maneira seletiva e bastante específica com os carboidratos presentes na superfície celular através da mediação

dos domínios de reconhecimento de carboidratos (CDR) por intermédio de ligações não covalentes, como ligações iônicas, de hidrogênio, hidrofóbicas e de Van de Walls(DA SILVA, 2018) (Figura 1). Cita-se, como exemplo, as lectinas presente em vários peixes, que tem a capacidade de reconhecer os carboidratos presentes na superfície celular das bactérias patogênicas de peixes de água doce e marinha, inibindo o crescimento desses microrganismos (HUANG et al., 2020; ZHANG et al., 2020).

**Figura 1- Representação esquemática da interação da lectina com um carboidrato.**



Além do importante papel no reconhecimento celular, a interação das lectinas com os carboidratos tem sido explorada em vários campos de pesquisa onde a especificidade ao carboidrato é essencial, sendo consideradas como uma valiosa ferramenta molecular. Esta característica tem sido utilizada em várias aplicações biotecnológicas como em estudos de proliferação de linfócitos (LUCIANO PIRES et al., 2019), atividade antimicrobiana (PIETRZYK-BRZEZINSKA; BUJACZ, 2020), regeneração celular (SUAREZ CARNEIRO et al., 2021) e cicatrização de feridas (GANESHALINGAM et al., 2022).

Nesse sentido, as lectinas permitem identificar a presença de patógenos e/ou moléculas-alvos em estudos variados. A exemplo disso tem-se o uso das lectinas para testes diagnósticos, onde estas são capazes tanto de identificar diretamente os patógenos (PIETRZYK-BRZEZINSKA; BUJACZ, 2020), quanto são eficientes para diferenciar entre suas formas de vidas (ARANDA-SOUZA et al., 2018; MULE et al., 2020), bem como reconhecer padrões de glicosilação entre células normais e tumorais (SHEWELL et al., 2018).

Logo, é importante compreender que essas moléculas estão disponíveis na natureza em todos os organismos vivos e desempenham atividade biológica bem diversificada. Muito já se conhece sobre as lectinas de origem vegetal, que são exaustivamente estudadas desde a sua primeira observação em 1888, por Stillmark. Diferentemente das lectinas animais, que mais recentemente tem despertado o interesse dos pesquisadores, em especial as lectinas encontradas no soro, plasma, muco e ovos de peixes. As lectinas de organismos aquáticos vêm acrescentando a essa área de estudo grandes contribuições.

## **1.2 Lectinas em peixes**

Em meados do século XX o reconhecimento da presença de lectina em tecidos animais ganhou evidência (BEZERRA, 2009; LINO et al., 2013; SHARON, 2008). A história mostra que as lectinas em animais (STOCKERT; MORELL; HERBERT SCHEINBERG, 1974) teriam sido identificadas antes mesmo que as lectinas de plantas (STILLMARK, 1888), mas em um contexto diferente (KILPATRICK, 2002). Apenas quando essas proteínas foram identificadas em mamíferos o cenário mudaria. Desde então, o número de pesquisa de lectinas aumentou exponencialmente (VASTA; AHMED, 2008).

A primeira lectina encontrada em um teleósteo, foi uma lectina carboidrato específica extraída de uma enguia por Watkins & Morgan em (1952), que posteriormente foi identificada como ligadora de L-fucose (RUDIGER; GABIUS, 2013). Desde então, o número de lectinas encontradas nesse grupo de organismo tem aumentado significativamente.

Com o advento da tecnologia do DNA recombinante, a partir dos anos 80, as pesquisas com lectinas puderam ser mais rapidamente desenvolvidas e isso trouxe grandes contribuições nessa área de estudo. Já era conhecido que essas proteínas ocorriam em órgãos internos em animais, entretanto descobriu-se que elas também poderiam ocorrer em secreções da superfície corporal de alguns animais, principalmente em peixes. As primeiras observações da presença dessas proteínas na superfície de peixe datam dos anos 80, quando Kamiya & Shimizu (1980) demonstraram a ocorrência de lectina no muco do corpo do linguado, *Lophopsettamaculata*, e Ingran(1980) propôs que a superfície do corpo de peixes continha lectinas e as descreveu como proteínas ligadoras de carboidratos.

Posteriormente, diversos trabalhos descrevendo a purificação de lectina em peixes foram conduzidos, incluindo bagre do Golfo Árabe, *Ariusthalassinus*(AL-HASSAN et al., 1986), enguia *Congermyriaster*(KAMIYA; MURAMOTO; GOTO, 1988; SHIOMI et al., 1989), dragonete *Repomucenusrichardsonui*(SHIOMI et al., 1990), dojô *Misgurnusanguillicaudatus*(GOTO-NANCE et al., 1995), abadejo *Genypterus capensis*(TODA et al., 1996). Estes trabalhos demonstraram a presença da lectina na superfície desses peixes e que a mesma está associada com a imunidade humoral, apresentando competência equivalente a do complemento e imunoglobulinas.

Da mesma forma, a estrutura molecular e função biológicas destas moléculas que raramente eram relatadas anteriormente, agora se tornou alvo das pesquisas nessa área. Os primeiros estudos relatando simultaneamente a localização e função biológica foram o da lectina presente no tecido epitelial da enguia *Anguilla japonica*, realizado por Suzuki (SUZUKI, 1985) e Suzuki & Kaneko(SUZUKI; KANEKO, 1986). Subsequentemente, a lectina do muco da enguia *Congermyriaster* foi sequenciada por Muramoto e Kamiya (MURAMOTO; KAMIYA, 1992). Em 1999, esses autores apresentaram a classificação estrutural da lectina do muco da superfície corporal da enguia, denominadas de congerinas I e II (MURAMOTO et al., 1999).

Outros estudos estruturais baseados no CRD e na dependência de cálcio ( $Ca^{2+}$ ) da lectina para aglutinar, foram conduzidos, como o da enguia *Anguilla japonica*(TASUMI et al., 2004), *Leiognathusnuchalis*(OKAMOTO et al., 2005) e fugu *Takifugurubripes*(TSUTSUI et al., 2006a). Pesquisas recentes envolvendo

lectinas de peixes se concentram em genes e nas funções biológicas que esses genes podem desempenhar na defesa desses organismos; a exemplo, dos trabalhos com o baiacu *Takifugu obscurus* (HUANG et al., 2022), tilápia do Nilo *Oreochromis niloticus* (YIN et al., 2022) e das enguias do gênero *Anguilla* (TSUTSUI et al., 2019). Em suma, pode se afirmar que as lectinas de peixes têm ampla diversidade molecular e que os estudos dos genes dessas proteínas têm se mostrado promissores.

### 1.3 Classificação das lectinas em peixes

As lectinas animais, de modo geral, são classificadas com base, principalmente, na estrutura dos CRDs. Atualmente essa classificação também leva em consideração a origem evolutiva compartilhada e a similaridade de dobras estruturais (PIETRZYK-BRZEZINSKA; BUJACZ, 2020). Além do requerimento ou não de íons de cálcio para a ligação com o carboidrato específico (ZHANG et al., 2020). Por se tratar de uma classe de proteínas com estruturas bem diversificadas, existem também um número grande de famílias de lectinas. Segundo Elumalai, Preetham *et al.* (ELUMALAI et al., 2019) as lectinas animais estão classificadas atualmente como: lectinas do tipo C, galectinas, tipo P, tipo I, pentraxinas, ligadora de heparina, tipo F, calnexina, tipo M, tipo L, tipo R, F-box, ficolina, tipo quitinase (quilectinas) e intelectinas (tipo X). E ainda, o grupo das lectinas ligadoras de ramnose (RBL) (MU et al., 2022).

Em peixes, as lectinas mais frequentemente caracterizadas são galectinas, lectina do tipo C, lectina ligadora de L-ramnose e lectina do tipo F.

Existem muitas famílias de lectinas animais. Porém esse número ainda pode aumentar, uma vez que os estudos dessas proteínas têm atraído a atenção de vários pesquisadores. A exemplo das lectinas do muco de peixes, que com suas propriedades únicas tem contribuído para a expansão dos números de famílias desse grupo de proteínas (SUN et al., 2019). Estudos sobre o muco da pele de peixes frequentemente visam investigar seu efeito antimicrobiano (CARNEIRO et al., 2022; HUANG et al., 2019; LIRIO; DE LEON; VILLAFUERTE, 2019).

### 1.3.1 Galectinas

As galectinas são uma família de lectinas animais que possuem um CRD conservado, com aproximadamente 130 resíduos de aminoácidos e especificidade para  $\beta$ -galactosídeos. Presente na maioria dos mucos de peixes, a galectina apresenta atividade independente de  $\text{Ca}^{2+}$ , ausência de peptídeo sinal (sequência de aminoácidos responsável pelo endereçamento da proteína) (WU et al., 2021), N-terminal bloqueado (extremidade amina modificada), ausência de glicosilação e a maioria está localizada no citoplasma das células (TIAN et al., 2021a).

Cerca de 15 tipos de galectinas foram identificados até o momento. Eles podem ser subdivididos em três grupos de acordo com a quantidade e disposição dos CRDs: os prototípicos, são aquelas que possuem um CRD (galectina-1, 2, 5, 7, 10, 11, 13, 14 e 15); o tipo quimera, que possui um CRD ligado a um domínio N-terminal rico em prolina, glicina e tirosina (galectina-3); e o tipo repetido em tandem, que possuem dois CRDs distintos conectados por uma ponte de mais de 70 aminoácidos (galectina-4, 6, 8, 9 e 12) (WU et al., 2021). Acredita-se que essas proteínas estejam envolvidas em diversos processos biológicos, como: imunomodulação, adesão celular, inflamação (WU et al., 2021), regulação de crescimento (GANESHALINGAM et al., 2022), apoptose (TIAN et al., 2021a), desenvolvimento e progressão de neoplasias.

Alguns peixes teleósteos apresentam galectinas, como as lectinas encontradas nas enguias *C. myriaster* e *A. japonica* (TSUTSUI et al., 2019). Outras galectinas, como a em tandem presente na garoupa *Epinephelus akaara* (GANESHALINGAM et al., 2022), a quimera do tambor amarelo *Nibeaalbiflora* (WU et al., 2021), e do rodvalho *Scophtalmus maximus* L. (TIAN et al., 2021b). Estas proteínas são consideradas galectinas devido a reação de afinidade de carboidratos apontarem a especificidade com  $\beta$ -galactosídeos. Dessa forma, é possível que outros peixes também possuam galectinas.

A respeito das lectinas isoladas em *C. myriaster*, Muramoto e Kamiya (MURAMOTO; KAMIYA, 1992) observaram que as congerinas I e II apresentavam

atividade hemaglutinante (AH) tanto para eritrócitos de coelho, quanto para os de cavalo e ovelha, além de aglutinar uma bactéria marinha, a *Vibrioanguillarum*. Ambas as lectinas apresentam massa molecular de aproximadamente de 30 kDa, porém cargas elétricas distintas. Ainda segundo estes autores, essas galectinas são exclusivas do muco e apresentam AH independente de cátions divalentes e possuem especificidade a derivados de galactose, particularmente lactose.

A sequência de aminoácido destas lectinas foi determinada, e a estrutura das congerinas foi caracterizada como sendo homodimérica, na qual cada subunidade consiste em 135 resíduos de aminoácidos, incluindo um N-terminal acetilado e sem resíduos de cistina. Congerina II apresenta apenas 47% de homologia com a sequência de congerina I, apesar dessas proteínas apresentarem muitas estruturas secundárias semelhantes. Ambas exibem sequências homologas de aproximadamente 35% com as galectinas de mamíferos e pertencem a família de galectinas prototípicas (MURAMOTO et al., 1999).

*Anguilla japonica* também apresenta duas lectinas, denominadas AJL-1 e AJL-2. Entretanto, apenas AJL-1 exibe especificidade para  $\beta$ -galactosídeos, independente de cátions divalentes. Tasumi, Satoshi et al. (TASUMI et al., 2004) a descreveu como homodimérica sem ligações covalentes, com massa molecular de 30 kDa, composta por 142 resíduos de aminoácidos e apresentando homologia com os membros da família de galectinas prototípicas. Sua expressão foi observada apenas na pele e é capaz de aglutinar a bactéria patogênica *Streptococcus difficile*.

Uma análise comparativa de AJL-1 com as congerinas revelou que esta possui 45% de homologia com a congerina II, lectina mais abundante no muco da *C. myriaster*, e exibe três diferentes resíduos de aminoácidos em seu CRD, similar a estrutura da congerina I (SHIRAI et al., 1999). A similaridade destas proteínas reforça a evidência do papel da lectina do muco no reconhecimento e defesa contra invasões de microrganismos patogênicos.

Recentemente Tsutsui, Shigeyuki et al. (TSUTSUI et al., 2019) determinaram as sequências dos genes que codificam para a galectina na mucosa em todas as espécies do gênero *Anguilla* e verificou que estes genes possuíam cerca de 2,3 a 2,5 kb de comprimento. E que eles apresentavam uma organização de quatro éxons e três íntrons, com exceção da espécie *Anguilla reinhardtii*, cuja porção C-terminal

apresenta uma modificação única. A ligação de  $\beta$ -galactosídeos ocorre com a participação de seis resíduos de aminoácidos, que se apresentam conservados (R73) em três espécies de *Anguilla* e são substituídos nas demais dezesseis espécies por K73. Ainda segundo estes autores, essa substituição não interfere na capacidade da molécula se ligar ao carboidrato. Por fim, a análise filogenética destes genes apontou que a molécula mais ancestral entre as espécies e subespécies de *Anguilla* é a galectina da espécie *Anguilla mossambica*.

Outra galectina presente no peixe asiático tambor amarelo (*Nibeaalbiflora*), denominada YdGal-3, foi caracterizada por Wu, Baolan *et al.* (WU *et al.*, 2021). Essa lectina foi detectada nos doze tecidos estudados e seu gene apresentou um quadro de leitura aberta (ORF) de 1152 pb, que codifica uma proteína de 383 aminoácidos e tem peso molecular de 38,2 kDa. Após um ensaio de desafio imunológico com *Vibrioharveyi*, YdGal-3 teve sua expressão aumentada nos tecidos relacionados a imunidade inata clássicos, sendo que sua expressão na porção cefálica do rim teve um destaque significativo. Indicando que a proteína pode estar envolvida na imunidade inata, o que foi ratificado ao observar que ela exibe capacidade de aglutinar bactérias Gram-negativas e apresentou atividade antibacteriana.

Pela primeira vez relatado em teleosteos, uma galectina do tipo repetido em tandem, obtida da garoupa manchada vermelha (*Epinephelus akaara*) apresentou a capacidade de cicatrização de feridas (GANESHALINGAM *et al.*, 2022). Esta galectina foi denominada EaGal8, seu gene exibiu uma ORF de 960pb, que codifica para uma proteína de cerca de 319 resíduos de aminoácidos, de aproximadamente 35 kDa e ponto isoelétrico de 7,69. A análise filogenética revelou que ela pertence à família de galectina-8. A expressão tecidual desta lectina foi ubíqua, com maior expressão observada no sangue. Após desafio imune sua transcrição foi regulada positivamente de uma maneira dependente do tempo. Ademais, apresentou afinidade de ligação a sete carboidratos diferentes, aglutinando tanto eritrócitos do próprio peixe quanto várias bactérias Gram-positivas e Gram-negativas.

### 1.3.2 Lectinas Tipo-C

As lectinas do tipo-C (CTL) foram inicialmente descritas como proteínas animais que exibem propriedades ligadoras de carboidratos dependente de  $\text{Ca}^{2+}$  e uma variedade estrutural significativa. No entanto, estudos mais recentes demonstraram que a dependência de  $\text{Ca}^{2+}$  pode ou não está presente (LOPES-FERREIRA et al., 2022). Membros desta família apresentam um ou mais domínio de reconhecimento de carboidratos, o C-typelectin-like domain (CLTD), e o domínio funcional pode apresentar 2 ou 3 pares de ligações sulfúricas, formando uma estrutura distinta e estável (HUANG et al., 2020; HUANG; MA; LEI, 2013). Esses domínios são bastante conservados apresentando motivos de ligação de carboidratos.

As CTLs são classificadas em duas principais categorias de proteínas: as manose-específico com motivo EPN (resíduos de ácido glutâmico, prolina e asparagina) e a galactose-específica com motivo QPD (resíduos de glutamina, prolina e ácido aspártico) (VASTA; AHMED, 2008). Funcionalmente as CTLs provaram participar de vários processos imunes, tais como, reconhecimento de patógenos, adesão celular, sinalização intercelular, entre outros. Em peixes já foram descritas várias CTLs, como em linguado *Cynoglossus semilaevis* (HUANG et al., 2019), baiacu *Takifugu obscurus* (HUANG et al., 2020), salmão do Atlântico *Salmo salar* (OJEDA et al., 2020; RICHARDS et al., 2003), falso alabote japonês *Paralichthys olivaceus* (KONDO et al., 2007), grande rodovalho *Scophthalmus maximus* L. (LIU et al., 2022; ZHANG; HU; SUN, 2010), peixe sapo Cano *Thalassophryne maculosa* (LOPES-FERREIRA et al., 2022), peixe escorpião *Scorpaenaplumieri* (ANDRICH et al., 2015) e corvina amarela gigante *Larimichthys crocea* (LV; ZHANG; WANG, 2016).

Embora seja comumente encontrada no interior dos organismos, como todas as CTLs citadas acima, existem relatos da presença dessas proteínas também na superfície corporal de alguns peixes. A AJL-2 é uma lectina que foi extraída do muco que recobre o corpo da *A. japonica* e apresenta estrutura molecular similar a lectina tipo C, com atividade específica de lactose e tanto aglutina *Escherichia coli* quanto inibe o seu crescimento (TASUMI et al., 2002). Adicionalmente, conCL-s é outra CTL, encontrada no muco da *C. myriaster*, com atividade manose-específica e que aglutina a levedura *Saccharomyces cerevisiae*. Porém estas lectinas não apresentam atividade dependente de cátions bivalentes (TSUTSUI et al., 2007).

Estruturalmente AJL-2 é um homodímero ligado covalentemente por ligações dissulfeto, resultante de 4 resíduos de cisteínas. Apresenta aproximadamente 30% de homologia com várias CTLs, com sequência de aminoácidos altamente conservados do CRD, apesar de mostrar atividade independente de  $\text{Ca}^{2+}$ . Esta lectina é produzida nas células claras do epitélio e secretada para a superfície corporal, onde é exclusivamente encontrada como um componente do muco. Além disso, AJL-2 possui a capacidade de aglutinar *Escherichia coli* e adicionalmente inibir o seu crescimento numa concentração de 1  $\mu\text{g/ml}$ , sugerindo sua participação na defesa do hospedeiro (TASUMI et al., 2002).

De acordo com Tsutsui et al. (2007), conCL-s possui estrutura tetramérica, consistindo em dímeros associados não covalentemente que são ligados para formar a estrutura tetramérica por pontes dissulfetos. Esta lectina é produzida nas células claras do epitélio, mas pode ser encontrada tanto nos tecidos internos quanto na superfície corporal do peixe. Em comparação com outras CTLs, conCL-s exibiu alta homologia com várias lectinas do tipo-C e apresentou motivos EPN no seu CRD e WND (resíduos de triptofano, asparagina e ácido aspártico) no sítio de ligação do íon divalente. Possui atividade aglutinante a levedura *S. cerevisiae* e a bactéria *E. coli*.

Não há dúvidas que ambas as lectinas, AJL-2 e conCL-s, fazem parte da família de lectinas tipo C, mesmo apresentando atividade independente de cátion divalente. Em comparação com as congerinas, lectinas encontradas no muco da *C. myriaster* e que pertence a família das galectinas, AJL-2 não apresenta homologia com essas proteínas. Com relação a conCL-s, apesar de serem produzidas pelo mesmo tipo de células, e serem encontradas no muco do mesmo peixe, as congerinas diferem estruturalmente, pertencendo a famílias distintas de lectinas. Mesmo AJL-2 e conCL-s que pertencem a mesma superfamília de lectinas, estão agrupadas em famílias distintas devido a suas diferenças estruturais e especificidade a carboidratos diferentes. Contudo, ambas lectinas são capazes de aglutinar *E. coli*, sendo que AJL-2 também é capaz de inibir o crescimento destas bactérias, enquanto conCL-s apenas aglutina a bactéria citada e adicionalmente levedura (SHIOMI et al., 1989; TASUMI et al., 2002; TSUTSUI et al., 2007).

### 1.3.3 Lectina ligadora de L-ramnose

Anteriormente descrita em ouriço do mar e em ovos de peixes, a lectina ligadora de L-ramnose (RBL) foi extraída da secreção corporal de peixe pônei, *Leiognathus nuchalis*, que é um peixe sem escamas que pertence à família de Leiognathidae e da ordem Perciformes. As RBLs apresentam de dois a três CRDs homólogos em tandem (repetições organizadas em sequências) de cerca de 95 resíduos de aminoácidos (TATENO et al., 2002).

Duas RBLs foram isoladas do peixe pônei, PFL-1 e PFL-2. Conforme descrito por Okamoto, Masaki et al. (OKAMOTO et al., 2005), PFL-1 exibe uma sequência de aminoácidos com alta homologia com lectinas ligadoras de L-ramnose descritas anteriormente e possui ainda um domínio de reconhecimento de carboidratos conservado típico de lectinas ligadoras de L-ramnose (CRD RBL), com repetições em tandem e está presente apenas na superfície corporal do peixe pônei. PFL-2 apresenta uma sequência N-terminal altamente idêntica à PFL-1, sugerindo que PFL-2 seja uma isoforma de PFL-1.

Em contraste com as RBLs anteriormente descritas, que estão frequentemente associadas a gônadas e ovos de peixes, a exemplo da RBL dos ovos da truta arco-íris *Oncorhynchus mykiss* (TATENO et al., 2001) e da truta branca *Salvelinus leucomaenis* (TATENO et al., 2002), que ocorrem em diversos tecidos e especificamente nos ovários, as lectinas isoladas no peixe pônei são apenas expressas na superfície corporal deste peixe, não possuindo nenhuma ligação com a lectina expressa nos ovários desse peixe.

Algumas RBLs, em ovos de peixes, apresentam habilidade de aglutinar patógenos, como as lectinas do salmão cão (CSLs) e as lectinas de truta branca (STLs), que são capazes de aglutinar *E. coli* e *Bacillus subtilis* (SHIINA et al., 2002; TATENO et al., 2002). No entanto, as lectinas encontradas no muco do peixe pônei não exibiu atividade aglutinante para as seis bactérias nas quais foram testadas, sugerindo que no futuro essas lectinas sejam testadas em um número maior de bactérias e outros patógenos, uma vez que as puffedlectinas apresentaram atividade aglutinante bastante específica para um número reduzido de bactérias e foi a

primeira lectina do muco de peixes descrita que possui a capacidade de se ligar a parasitas (OKAMOTO et al., 2005).

Vários tipos de lectinas estão presentes no muco de peixes. Essa diversidade molecular é muito importante para a elucidação da função destas moléculas nesses organismos. RBLs são conhecidas por estarem envolvidas em mecanismos de resposta inflamatória, atuando como moléculas de reconhecimento e transporte celular, assim como ativadores da cascata inflamatória através de citocinas (VASTA et al., 2011).

#### **1.4 Atividade biológica das lectinas de peixes**

Segundo Ángeles Esteban (ÁNGELES ESTEBAN, 2012), as lectinas dos peixes apresentam propriedades antimicrobianas contra infecções de patógenos e tem sido demonstrado em diferentes peixes. Watanabe, Yasuharu *et al.* (WATANABE et al., 2013), admite que o muco dos peixes funciona como um mecanismo de barreira bioquímica.

Entre as moléculas que compõem o muco do peixe, as lectinas possuem a capacidade de aglutinar células, através do reconhecimento de carboidratos específicos. Esta capacidade de aglutinar das lectinas de peixe previne a entrada de patógenos no organismo desses animais. O muco de peixes permite o controle de muitos eventos no organismo vivo, tais como fertilização, desenvolvimento e reconhecimento do próprio/não-próprio, pela variação estrutural do glicocálice das células (SUZUKI et al., 2003).

Como se pode verificar, lectinas em geral são proteínas que apresentam diversas atividades biológicas. Existem relatos de atividade antibacteriana (CARVALHO et al., 2015; MOURA et al., 2015; PROCÓPIO et al., 2017; SILVA et al., 2016; TASUMI et al., 2002, 2004), atividade antiviral (BARTON et al., 2014; MOLCHANOVA et al., 2007; SATO et al., 2015), atividade inseticida (COELHO; MARANGONI; MACEDO, 2007; SÁ et al., 2009), atividade antifúngica (ALBUQUERQUE et al., 2014; DA SILVA, 2018; REGENTE et al., 2014; SITOHY; DOHEIM; BADR, 2007; TSUTSUI et al., 2007; WEIDENBACH et al., 2016), atividade

antitumoral (KAWANO et al., 2009; PETROSSIAN; BANNER; OPPENHEIMER, 2007) entre outras, indicando que essas proteínas funcionam no mecanismo de defesa do organismo vivo. Evidentemente estas proteínas podem ser utilizadas para aplicações clínicas e biotecnológicas, como marcadores celulares, transporte de moléculas, a exemplo de fármacos que agem em células-alvos pelo sistema de transporte de drogas (DDD, do inglês Drugs Delivery Development) (BIES; LEHR; WOODLEY, 2004; PIETRZYK-BRZEZINSKA; BUJACZ, 2020), testes diagnósticos e outras inúmeras possibilidades. Uma pesquisa inédita no Brasil utilizou a lectina obtida de um peixe para produção de implantes dentários, aumentando a eficiência desses implantes (ANJOS, 2015).

Alguns estudos têm mostrado que as lectinas de peixes podem se ligar a bactérias (HUANG et al., 2019, 2022; LV; ZHANG; WANG, 2016), fungos (TSUTSUI et al., 2007), parasitas (BRINCHMANN et al., 2018; NAKAMURA et al., 2012; TSUTSUI et al., 2006a) e podem desempenhar outras atividades biológicas (Tabela 1). Essas macromoléculas são altamente específicas para o seu ligante e podem impedir a aderência de potenciais patógenos ao corpo dos peixes, inibindo sua invasão. Devido a essas propriedades, lectinas são consideradas potenciais agentes antimicrobianos presentes no muco.

Pesquisas com lectinas de plantas são mais abundantes e mais profundamente exploradas. Mas a descoberta de novas lectinas animais, especificamente as lectinas de peixes, tem trazido uma nova dimensão a biologia e imunologia de peixes.

**Tabela 1 - Atividade biológica das lectinas de peixes.**

Atividade	Lectina	Peixe	Especificidade de ligação	Família de lectina	Referências
Antiparasitária	Congerinas	Enguia congro ( <i>C. myriaster</i> )	$\beta$ -galactosídeos	Galectina	(NAKAMURA et al., 2012)
Antifúngica	ConCL-s	Enguia congro ( <i>C. myriaster</i> )	Manose	Tipo C	(TSUTSUI et al., 2007)
Antibacteriano	MaCTL	Botia ( <i>Misgurnus anguillicaudatus</i> )	Galactose	Tipo C	(ZHANG et al., 2020)

Encapsulação celular	LcGal9	Corvina amarela ( <i>L. crocea</i> )	$\beta$ -galactosídeos	Galectina	(ZHANG et al., 2016)
		Corvina amarela ( <i>L. crocea</i> )	Antibacteriano	Lyc CTLR	(AO et al., 2015)
Antibacteriana	LcNNTC	Corvina amarela ( <i>L. crocea</i> )	Manose	Tipo C	(LV; ZHANG; WANG, 2016)
Efeito citotóxico	SauFBP32	Salmão do Atlântico ( <i>Salmo salar</i> )	Fucose	Tipo C	(EWART; JOHNSON; ROSS, 1999)
Antibacteriana	CSL (1-3)	Salmão cão ( <i>O. keta</i> )	Ramnose	RBL	(SHIINA et al., 2002)
Antiparasitário	Pufflectinas	Baiacu ( <i>T. rubripes</i> )	Manose	Tipo C	(TSUTSUI et al., 2006b)
Antibacteriano	Drgal-1 (L1, L2, L3)	Peixe zebra ( <i>D. rerio</i> )	LaNAc	Galectina	(CHEN et al., 2018)
Antibacteriano	AJL-1	Enguia japonesa ( <i>A. japonica</i> )	$\beta$ -galactosídeos	Galectina	(TASUMI et al., 2004)

**Tabela 1 – Atividade biológica das lectinas de peixes (continuação)**

Atividade	Lectina	Peixe	Especificidade de ligação	Família de lectina	Referências
Antibacteriano	AJL-2	Enguia japonesa ( <i>A. japonica</i> )	Lactose	Tipo C	(MU et al., 2022; TASUMI et al., 2002)
Aglutina patógenos bacteriano	OnMBL	Tilápia do Nilo ( <i>O. niloticus</i> )	Manose	Tipo C	(MU et al., 2017)
Regula a fagocitose de monócitos	OnMR	Tilápia do Nilo ( <i>O. niloticus</i> )	Manose	Tipo C	(YIN et al., 2022)
Induz produção IFN- $\gamma$	OniL	Tilápia do Nilo ( <i>O. niloticus</i> )	Manose	Tipo C	(DA SILVA et al., 2012)
Antibacteriana	EaGal8	Garoupa	$\beta$ -galactosídeos	Galectina	(GANESHALINGAM et al., 2022)

<b>Antiviral</b>		<i>(E. akaara)</i>			
<b>Cicatrização de feridas</b>					
<b>Aglutina patógenos bacterianos</b>	YdGal3	Tambor amarelo <i>(N. albiflora)</i>	$\beta$ -galactosídeos	Galectina	(WU et al., 2021)
<b>Antibacteriana</b>	ComaSeL	Tambaqui <i>(C. macropomum)</i>	Galactose	Tipo C	(MACIEL CARVALHO et al., 2012)
<b>Inibição do crescimento bacteriano</b>	SalntL	Bagre do Atlântico <i>(S. asotus)</i>	Manose	Intelectina	(TSUTSUI et al., 2011)

Fonte: Adaptado de (ELUMALAI et al., 2019).

### 1.5 Aspectos genéticos das lectinas em peixes

A respeito dos genes de lectinas em peixes as informações disponíveis ainda são limitadas. Isto pode ser justificado, em parte, pela quantidade de peixes que tiveram seus genomas montados e pela qualidade que esses genomas foram anotados. De acordo com Lu e Luo (LU; LUO, 2020) até 2019 cerca de 270 genomas de peixes, das mais de 34.000 espécies de peixes conhecidas, estavam disponíveis no Centro Nacional de Informações sobre Biotecnologia do Genoma dos EUA (NCBI) e que tais sequências variavam em muito na qualidade, devido tanto as tecnologias de sequenciamento quanto as ferramentas de software que haviam sido utilizadas. Os autores ainda ressaltam que a maioria desses genomas havia sido sequenciada utilizando tecnologia de sequenciamento de nova geração, amplamente utilizado na atualidade, que geram leituras curtas, o que torna o trabalho de montagem do genoma bastante desafiador.

Os genes de lectinas já foram identificados em várias espécies de peixes. Estudos atuais em lectinas de peixe têm evidenciado esses genes, o que tem contribuído para o conhecimento acerca do perfil de expressão dessas proteínas e do papel delas na defesa imune. A exemplo das naterinas, que são lectinas que apresentam um domínio de ligação a carboidrato, semelhante a jacalina, e um outro domínio de ligação a toxina, semelhante a aerolisina. Esta estrutura semelhante a aerolisina sugere que a proteína é capaz de lisar patógenos (BRINCHMANN et al.,

2018). Essa teoria foi demonstrada em peixe zebra (*Danio rerio*), quando a expressão do gene de naterina aumentou após desafio com *Staphylococcus aureus* e a eliminação das bactérias aumentou quando antes do desafio era administrado uma injeção da proteína (CHEN et al., 2018). Além disso, o gene da naterina (Aep1) foi expresso no estágio inicial do desenvolvimento, o que sugere que esse gene está relacionado com o sistema imune inato, uma vez que nesse estágio do desenvolvimento o sistema imune adaptativo ainda não é expresso (CHEN et al., 2018).

O baiacu (*Takifugurubripes*) é um dos peixes que tiveram seu genoma sequenciado, o que permitiu que os genes de lectinas pudessem ser mais bem identificados e caracterizados. Atualmente há 3 versões de montagens do genoma do baiacu disponíveis publicamente no NCBI. Vários genes de lectinas foram identificados no baiacu, a exemplo da família de galectinas demonstrada na tabela 2.

**Tabela 2–Genes de Galectinas no baiacu (*Takifugurubripes*)**

	Comprimento		Cromossomo	Éxons	ID do gene
	Transcritos (pb)	Proteínas (aa)			
Gal1 (proto)					
<b>LOC115248679</b>	956	135	-	11	115248679
<b>LOC115249819</b>	1095	135	5	12	115249819
<b>LOC115248474</b>	744	135	-	6	115248474
<b>LOC105418406</b>	1341	135	5	6	105418406
<b>LOC115248680</b>	1433	135	-	4	115248680
<b>LOC115248678</b>	689	135	-	5	115248678
<b>LOC115248677</b>	1076	135	-	6	115248677
<b>LOC115248467</b>	913	135	-	5	115248467
Gal2 (proto)					
<b>Igals2b</b>	674	131	17	4	101071380
<b>LOC115248468</b>	1578	132	-	4	115248468
Gal3 (quimera)					
<b>LOC101079160</b>	1571	356	16	6	101079160

<b>Igals3bp</b>	1665	479	5	4	101064756
<b>LOC101076673</b>	2102	550	20	6	101076673
<b>Igals3b</b>	1003	240	2	7	105418870
<b>Igals3a</b>	1876	373	2	6	101061771
Gal4 (tandem)					
<b>LOC101073226</b>	1562	452	11	14	101073226
Gal8 (tandem)					
Igals8	1061	336	16	9	101076544
<b>LOC101074025</b>	1373	296	2	8	101074025

**Tabela 2 – Genes de Galectinas no baiacu (*Takifugurubripes*). (Continuação)**

	Comprimento		Cromossomo	Éxons	ID do gene
	Transcritos (pb)	Proteínas (aa)			
Gal9 (tandem)					
<b>LOC101079175</b>	1341	346	11	8	101079175
Outras galectinas					
<b>LOC101073084</b>	909	175	17	5	101073084
<b>Igalsl</b>	4270	178	5	4	101072765
<b>Igals2a</b>	908	136	5	4	101063946

Fonte: banco de dados do NCBI.

O número de cópias dos genes de lectinas varia significadamente entre as espécies de peixes. Estudos de análise de expressão em diferentes tecidos revelaram diferentes números de cópia de lectinas ligadora de ramnose (RBL) em diferentes espécies de peixes. Existem três RBLs que são expressas no salmão cão (*Oncorhynchusketa*) e na truta arco-íris (*Oncorhynchusmykiss*), quatro RBLs na tilápia do Nilo (*Oreochromisniloticus*), seis cópias foram identificadas no peixe-gato do canal (*Ictaluruspunctatus*) e oito no bagre do sul (*Silurusmeridionales*) (HUANG et al., 2019). O maior número de RBLs foram encontradas no zebrafish, que foi previsto possuir 22 RBLs potenciais (OGAWA et al., 2011). O mesmo é verdade para outros tipos de lectinas, como observamos na enguia europeia (*Anguilla anguilla*) que

apresenta sete fuclectinas, e são diferencialmente expressas (HONDA et al., 2000; VASTA et al., 2017). Enquanto um gene é expresso no fígado e secretada no soro, associado a uma defesa sistêmica, o outro é expresso nas brânquias e está associada a proteção da mucosa (BRINCHMANN et al., 2018).

A localização da expressão desses genes é também bastante variável. É o que acontece com os genes homólogos de galectinas encontradas no bacalhau do Atlântico (*Gadus morhua*). Estes genes podem ser encontrados tanto na superfície da pele e muco, quanto no reto. Genes de uma mesma lectina podem apresentar expressão específica em tecido e célula. Assim como podem ser expressos em diferentes estágios do desenvolvimento do peixe. Por fim, genes de diferentes lectinas podem ser identificados em um mesmo peixe.

#### **1.60 peixe pulmonado sul-americano (*Lepidosirenparadoxa*)**

O *Lepidosirenparadoxa* (Fitzinger, 1837) é popularmente conhecido como pirambóia, que significa peixe-cobra na língua tupi ("pirá" - peixe e "mboîa" - cobra). Mas também pode ser conhecido por outros nomes, como pirarucu-boia, traíra-boia ou caramuru. Este animal possui o corpo alongado, semelhante a enguia, uma cabeça achatada e dois pares de nadadeira lobadas (peitoral e pélvica), podendo atingir cerca de 1,5 m de comprimento. No Brasil, a piramboia é encontrada nos pântanos e lagos com baixo teor de oxigênio da região Amazônica. Na época de chuva esses animais mantêm uma vida ativa, enquanto na época de estiagem eles permanecem pouco ativos, logo após cavarem um canal de aproximadamente 1 metro de profundidade com uma câmara no final, onde permanecem durante toda estação seca (ALMEIDA-VAL et al., 2010).

A piramboia é um peixe pulmonado, que pertencem ao grupo dos peixes com nadadeiras lobadas, os Sarcopterygii. Eles recebem esse nome por possuir um pulmão primitivo que permite que esses animais sejam capazes de extrair o oxigênio do ar atmosférico. Os Sarcopterygii compreendem as classes Actinistia (celacantos) e Dipnoi (peixes pulmonados). Os Dipnoi são exclusivamente de água doce, embora no Devoniano existisse espécies também marinhas. Existem 4 gêneros vivos de

peixes pulmonados atualmente, porém apenas 3 pertencem a classe dos Dipnoi que são os gêneros *Neoceratodus*, *Lepidosiren* e *Protopterus*, que são endêmicos da Austrália, América do sul e África respectivamente (MAINA, 2004). Segundo Xiaoni, Gan *et al.*(XIAONI *et al.*, 2011), estes animais ocupam uma posição muito importante na filogenia dos vertebrados. Pois são considerados um grupo irmão dos tetrápodes e, portanto, constituem-se em modelo animal para entender a aquisição de características que permitiram a transição do ambiente aquático para o terrestre. Estes animais acumulam adaptações morfológicas, fisiológicas e moleculares que os permitiram a tolerância a um ambiente pobre em oxigênio (LÜDEMANN *et al.*, 2020). Dentre essas adaptações incluem genes envolvidos na imunidade (AMEMIYA *et al.*, 2013) e que protege os animais de ataques de patógenos favorável ao ambiente aquático.

**Figura 2 - Foto da Piramboia (*Lepidosiren paradoxa*)**



De acordo com Medzhitov e Janeway(MEDZITOV; JANEWAY JR., 1997) o sistema imune dos organismos multicelulares evoluiu mediante uma pressão seletiva imposta pelos microorganismos, e isso culminou no desenvolvimento de diversas estratégias de defesa, os quais protegem o hospedeiro por meio da neutralização e/ou destruição de patógenos invasores. O reconhecimento do patógeno constitui a primeira etapa para que a resposta imune seja estabelecida. E isso acontece através de proteínas capazes de reconhecer, desencadear e amplificar a resposta imune contra uma grande variedade de invasores (RIZZATO, 2009).

A pele, inclusive o muco que recobre a pele, é uma das principais linhas de defesa no sistema imunológico dos peixes (MAGNADÓTTIR, 2006). Porém seus

mecanismos moleculares, que dão início a resposta imune, continuam obscuros em peixes (ÁNGELES ESTEBAN, 2012). Conhecer esses mecanismos imunes é muito importante, tanto para o manejo dos peixes quanto para a saúde do homem, pois devido as propriedades antibactericidas o muco da pele do peixe pode ser considerado uma fonte potencial de componentes antibacterianos, podendo ser usados pela indústria farmacêutica na produção medicamentos alternativos para o tratamento de infecção por microrganismo e resistência a medicamentos.

Durante o período de estivação esses animais secretam um muco que recobre o seu corpo e os mantém na temperatura ideal durante toda a estação. A manipulação desses animais em viveiros e laboratórios também estimula a secreção de mucos nestes peixes. O muco atua como uma barreira física, servindo como uma armadilha para aprisionar patógenos, e uma barreira química, através da sua composição de moléculas como lectinas, lisozimas, peptídeos antimicrobianos (AMPs) entre outros, que agem como neutralizadores de patógenos. O período de estivação permite que o animal sobreviva em ambientes desfavoráveis mais aumenta sua vulnerabilidade a patógenos (DARBY HEIMROTH et al., 2021). Apenas os peixes pulmonados africano e sul-americanos possuem a capacidade de estivação, diferentemente do representante australiano.

O muco dos peixes pulmonados não é tão estudado quanto os dos demais peixes teleósteos, mas análises moleculares e comparativa demonstraram a presença de algumas moléculas da imunidade humoral no muco da pele do peixe pulmonado africano, *Protopterus dolloi* (ÁNGELES ESTEBAN, 2012; SMITH; RISE; CHRISTIAN, 2019). Estudo recente demonstrou que durante o período de estivação o casulo formado a partir do muco secretado pelo peixe pulmonado africano possui propriedade antimicrobiana promovendo um papel protetor a patógenos durante o estágio de dormência desses animais (DARBY HEIMROTH et al., 2021). O que nos levou a perguntar qual é o papel das lectinas na defesa desses animais na composição do muco, uma vez que num estudo preliminar, em bancos de dados genéticos, nosso grupo detectou a presença de vários genes de lectinas no transcriptoma da pirambóia (Tabela 3).

**Tabela 3 - Genes de lectinas identificados no transcriptoma da pirambóia (*Lepidosiren paradoxa*)**

INTELECTINA	GALECTINAS	SUPERFAMÍLIA TIPO C				RE	TIPO I
		Tipo c	Colectinas	Selectinas	Cell Killer		
ITLN 1	LGALSL	CLEC2B	COLEC11	SELP	KLRD1	MLEC	SIGLEC5
ITLN 2	LGALS1	CLEC2D	COLEC12	SELE	KLRF1	ERLEC1	SIGLEC7
	LGALS2	CLEC2L		SELPLG	KLRG1		SIGLEC1 5
	LGALS3	CLEC2B					
	LGALS8 (a)	CLEC3A					
	LGALS8 (b)	CLEC4					
		(E, L, M)					
		CLEC10A					
		CLEC11A					
		CLEC14A					
		CLEC17A					
		CLEC19A					
		CD207					

Neste projeto propomos isolar, caracterizar e avaliar a atividade antimicrobiana da lectina presente no muco do peixe pulmonado sul-americano. Analisar o padrão de expressão dos genes da lectina na mucosa do animal durante o período de estivação. Esperamos desta forma prover informações acerca da resposta imune inata que permite aos peixes pulmonados sobreviverem em ambiente desafiador.

## 2 OBJETIVOS

### 2.1 Objetivo geral

Identificar genes de lectinas em *Lepidosirenparadoxa* com base nos transcriptomas e prever as características e funções das proteínas através de análise *in silico*.

### 2.2 Objetivos específicos

- Identificar genes de *L.paradoxa* nos transcriptomas disponíveis;
- Modelar as estruturas da lectina do peixe *Lepidosirenparadoxa*;
- Caracterizar as lectinas da piramboia;
- Prever a função da lectina na imunidade contra microrganismos.

### **3 METODOLOGIA**

#### **3.1 Animate amostras**

Espécimes de *Lepidosiren paradoxa*, foram obtidos de fontes naturais no Estado do Pará (licença Sisbio nº 79802-2). Os peixes foram aclimatados às condições de laboratório por pelo menos uma semana. Durante o período de adaptação, os peixes foram alimentados com ração comercial para peixe (36% PB) uma vez ao dia. Os cuidados com os animais foram conduzidos de acordo com o Comitê de Ética em Pesquisa Animal da Universidade Federal do Pará, sob o protocolo CEUA Nº 8251300622. Após sedação dos peixes, foi realizada uma lesão napele dos animais com o uso de Punch dermatológico 6mm de diâmetro.

#### **3.2 Dados transcriptômicos**

Os dados gerados dos transcriptomas da piramboia foram utilizados para buscar genes de lectinas, dentre os genes diferencialmente expressos (DEGs). Dois conjuntos de dados de transcriptomas da piramboia foram analisados, referente a diferentes tempos (2, 7 e 21 dias pós-injúria - dpi) e estruturas regeneradas (pele e nadadeira). Após mapeamento dos genes (DEGs), as sequências dos genes de lectinas referentes a 2 e 7 dias foram recuperadas do banco de dados de transcriptoma da piramboia (dados não publicados, SCHEIDER, I., 2023) e os dados referentes a 21 dias foram recuperados dos dados publicados (NOGUEIRA et al., 2016) e depositados no NCBI, sob nº. de acesso: BioProject, PRJNA301439.

#### **3.3 Análises computacionais**

##### **3.3.1 Análises de sequências**

As sequências de nucleotídeos obtidas a partir dos dados publicados do transcriptoma de pirambóia ((NOGUEIRA et al., 2016); BioProject, número de acesso: PRJNA301439) serão analisadas usando o NCBI Basic Local Alignment Search Tool (BLAST). A ferramenta ClustalW (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) para analisar a identidade e a plataforma NCBI ORF finder (<https://www.ncbi.nlm.nih.gov/orffinder/>) para determinar a estrutura de quadro de leitura aberta. As relações evolutivas destas lectinas serão analisadas através do software MEGA, para identificação da família de lectina.

### 3.3.2 Caracterização das lectinas

As propriedades físico-químicas das lectinas serão preditas com a ferramenta ExPASyProtParam (<https://web.expasy.org/protparam/>). Para prever a estrutura tridimensional da proteína a plataforma SWISS-Model (<https://www.expasy.org/resources/swiss-model>) será utilizada. As regiões de domínios conservadas das sequências de aminoácidos serão determinadas usando a ferramenta SMART (<http://smart.embl-heidelberg.de/>).

## 3.4 Modelagem molecular por homologia

A modelagem molecular foi realizada utilizando-se duas estratégias. As sequências que apresentaram valores de identidades razoáveis com os moldes depositados no RCSB Protein Data Bank (BERMAN et al., 2000) foram modeladas através do software Modeller 10.4 (WEBB; SALI, 2016). O alinhamento dessas sequências com os seus respectivos moldes fora realizado pelo Promals3D.

Um total de 100 modelos foram gerados, dessa forma considerando diferentes conformações e elencados de acordo com as pontuações Molpdf e DOPE. Os modelos foram gerados satisfazendo restrições espaciais como comprimento de ligação, diferentes ângulos, diedros e interações não-ligadas, além do emprego do refinamento pelo servidor GalaxyWEB(KO et al., 2012).

Após a construção do modelo, a qualidade estereoquímica será avaliada através do servidor SAVES do UCLA-DOE. O SAVES concentra diversos tipos de programas que foram utilizadas como o PROCHECK (LASKOWSKI; THORNTON; MACARTHUR, 2006), que avalia a qualidade estereoquímica através do gráfico de Ramachandran. O Verify3D (EINSENBURG; LUTHY; BOWEI, 1997) irá avaliar a qualidade do enovelamento do modelo e finalmente o ERRAT (COLOVOS; YEATES, 1993) será computado para mensurar a qualidade local do modelo.

### **3.5 Docking molecular**

A simulação da docagem molecular foi realizada usando o programa GOLD (*Genetic Optimisation for Ligand Docking*) versão 5.5 (JONES et al., 1997), que realiza docagem baseada em algoritmos genéticos e tem se mostrado eficiente no *Screening* virtual, otimização e identificação de modos de ligação de diversas moléculas bioativas. O conjunto de moléculas selecionadas como ligantes foram encaixadas pelo GOLD, aplicando a função de pontuação GoldScore (VERDONK et al., 2003) com 100% de eficiência e interação com resíduos de 10,0 Å no sítio de ligação à carboidratos das moléculas.

### **3.6 Anotação funcional das lectinas**

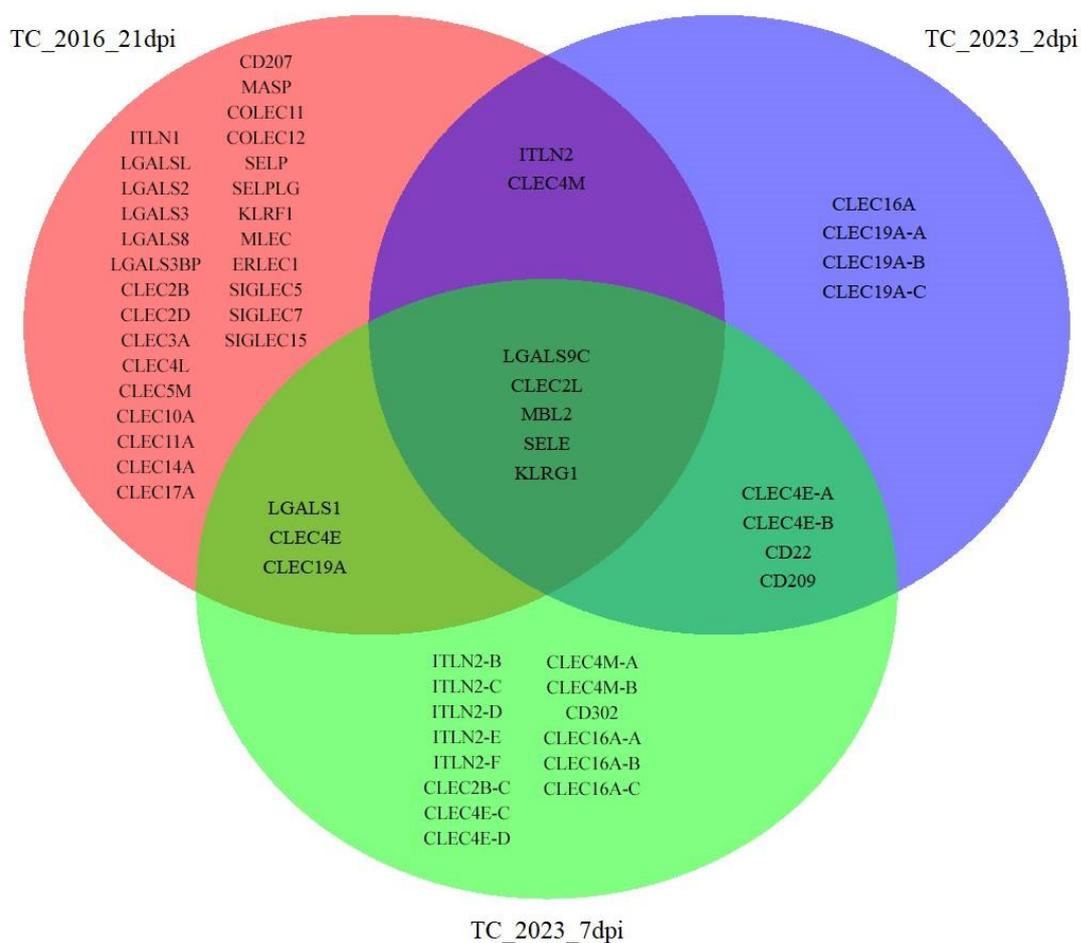
A análise da anotação funcional das lectinas foi verificada utilizando a plataforma DAVID (*Database for Annotation, Visualization and Integrated Discovery*) versão v2023q2, por meio do módulo GoCharts, que atribui genes a categorias funcionais de Gene Ontology com base nas classificações selecionadas pelo usuário e no nível de especificidade do termo. As sequências das proteínas obtidas do UniProt foram adicionadas na plataforma DAVID, que analisou para três tipos de funções: 1) processos biológicos, 2) função molecular e 3) compartilhamento celular. As sequências foram comparadas com as do peixe pulmonado africano (*Protopterus annectens*) e foi utilizado o grau de especificidade 2.

## 4 RESULTADOS

### 4.1 Identificação das lectinas de piramboia

As sequências nucleotídicas das lectinas foram recuperadas do banco de dados dos transcriptomas da piramboia (NOGUEIRA et al., 2016;SCHEIDER, I.,2023 – dados não publicado). Ao todo identificamos 78 genes de lectinas, sendo 15 nas amostras de 2 dpi, 26 nas amostras de 7dpi e 36 nas amostras de 21dpi (Figura 3).

**Figura 3– Genes de lectinas expressos nos transcriptomas de piramboia durante a regeneração do tecido.**



Dentre os genes diferencialmente expressos conseguimos caracterizar *in silico* o produto de sete desses genes, referentes a três famílias de lectinas: três do tipo C (CLEC19A, CD302/CLEC13A e SELP), uma intelectina (ITLN2) e três galectinas (LGALS1, LGALS9 e LGALSL). Os genes CLEC19A apresentou três variantes e ITLN2 exibiu quatro, totalizando doze lectinas identificadas e caracterizadas. A tabela atualizada com o total de lectinas identificadas por famílias no transcriptomas da piramboia encontra-se no apêndice.

## 4.2 Caracterização das lectinas de piramboia

Dentre as lectinas identificadas, doze apresentaram moldes acurados que permitiram modelar essas proteínas. As lectinas foram agrupadas, de acordo com a análise das suas respectivas sequências, em três famílias distintas:

### 4.2.1 Lectinas do tipo C

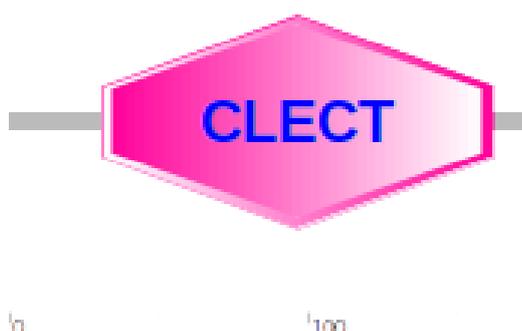
A análise dos transcriptomas da piramboia revelou três variantes do gene CLEC19A: CLEC19A-A, CLEC19A-B e CLEC19A-C. O produto desse gene foi denominado *LpCTL* e suas variações *LpCTL-1*, *LpCTL-2* e *LpCTL-3*, respectivamente.

Os resultados mostraram que o comprimento total da ORF do gene CLEC19A-A compreende 528 pb e codifica uma proteína de 175 aminoácidos (ver sequência no Apêndice 1), contendo um peptídeo sinal, formado por 18 resíduos N-terminais e um domínio de reconhecimento de carboidrato (CRD). A predição do peptídeo sinal sugere que esta proteína seja secretada. O peso molecular previsto de *LpCTL-1* foi de 20,36 kDa e o seu ponto isoelétrico foi de 5,24. O índice de instabilidade para esta proteína foi de 36,34, classificando-a como estável. Sendo uma lectina da superfamília do tipo C, *LpCTL-1* apresentou um CRD conservado (resíduos 30-167) com um motivo de ligação galactose-específico, QPD, e outro específico para a ligação com o  $\text{Ca}^{2+}$ , WND.

Para avaliar a relação de lectinas do tipo C (CTL), a sequência de aminoácido de *LpCTL* foi alinhada com outras CTLs conhecidas por ClustalW.\*O resultado dos alinhamentos mostrou que *LpCTL* apresenta as quatro cisteínas conservadas, que seriam responsáveis pela formação de pontes dissulfeto intra-cadeia que estabiliza a estrutura da proteína.

A modelagem por homologia previu que a estrutura tridimensional de *LpCTL-1* tem 3  $\alpha$ -hélices,  $\alpha$ -1 (WVNAEAYCQRN),  $\alpha$ -2 (LEENNYIYELMG) e  $\alpha$ -3 (NDYI) e 8 folhas- $\beta$  (Figura 5<sup>a</sup>). A modelagem também mostrou que o sítio de ligação de  $\text{Ca}^{2+}$  e de ligação ao açúcar estavam expostos na porção C-terminal.

**Figura 4 - Esquema da estrutura de *LpCTL***

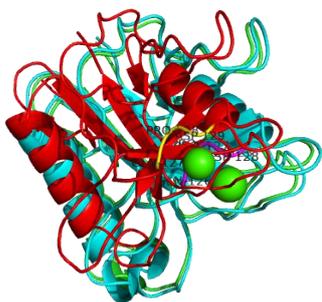


A estrutura tridimensional de *LpCTL-2* (Figura 5B) é muito semelhante a *LpCTL-1* (Figura 5A), apresentando uma leve diferença conformacional na sua estrutura secundária. Porém ambas as proteínas apresentam 175 aminoácidos, pesos moleculares de 20,50 kDa e 20,36 kDa respectivamente e pontos isoelétricos de 5,24. A variante *LpCTL-3* (Figura 5C) é um pouco mais curta, 172 aminoácidos, possui um peso molecular de 19,66 kDa, ponto isoelétrico de 4,78. Todas as proteínas foram consideradas estáveis (índice de instabilidade 28,55 e 27,25) e apresentavam o motivo de ligação galactose-específico, QPD, na mesma posição da sequência. A estrutura tridimensional de *LpCTL-3* mostrou que essa proteína exibia uma  $\alpha$ -hélice a menos ( $\alpha$ -3). O gráfico de Ramachandran mostrou que 94,1%, 94,12% e 92,4%, respectivamente, dos resíduos nas estruturas de *LpCTL-1*, *LpCTL-2* e *LpCTL-3* (Figura 5) estão em posições favoráveis energeticamente. O fator de qualidade geral e compatibilidade para as proteínas modeladas foram observados

como 91,41%, 85,82% e 78,9% no ERRAT e 83,09%, 75,74% e 85,29% no Verify3D, respectivamente. Esses resultados confirmam que o modelo apresenta boa qualidade e confiabilidade.

**Figura 5 - Estruturas tridimensionais de LpCTL-1, suas variantes e comparações.**

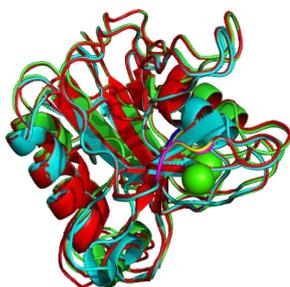
A.



B.

C.

D.



E.

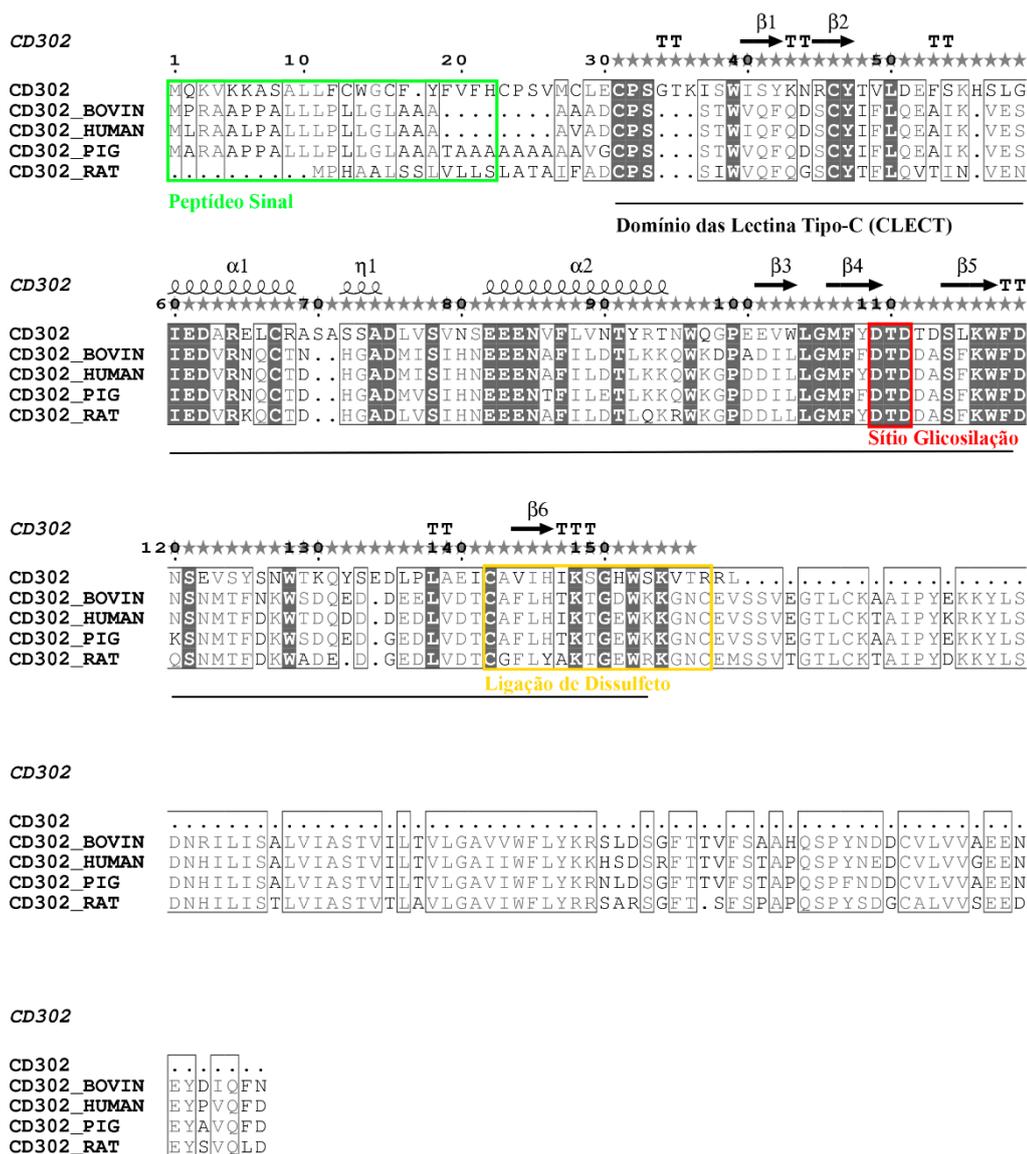
F.

(A) LpCTL-1, (B) LpCTL-2 e (C) LpCTL-3. Comparações entre (D) LpCTL-1 e LpCTL-2, (E) LpCTL-1 e LpCTL-3 e (F) LpCTL-2 e LpCTL-3.

O ORF de CD302 foi de 480 pb, codificando uma proteína de 159 aminoácidos. E *LpCD302* é uma proteína estrutural, com um domínio transmembrana (7-29 aminoácidos) e um domínio de lectina tipo C (31-159 aminoácidos) (Figura 7). O peso molecular predito de *LpCD302* foi de 18,38 kDa com um pI teórico de 5.5.

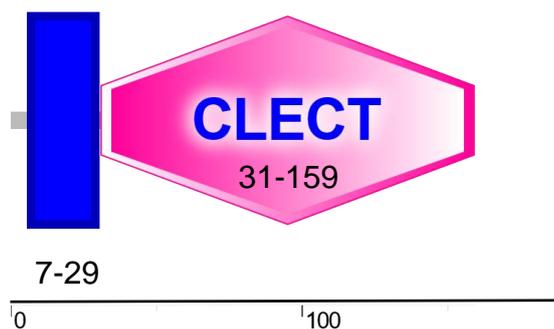
O alinhamento das sequências demonstra que a lectina *LpCD302* compartilha um domínio de lectina tipo C com as lectinas CD302 de outras espécies (Figura 6). Além do domínio transmembrana previsto na sequência de aminoácido, o alinhamento revelou um peptídeo sinal (22 resíduos) na região N-terminal e os resíduos DTD, na posição 109-111, foram preditos como um sítio de glicosilação. A estrutura molecular de *LpCD302* (Figura 8) apresenta 6 cadeias folhas- $\beta$  e 2 cadeias  $\alpha$ -hélices, as posições dos resíduos são energeticamente favoráveis, pois o gráfico de Ramachandran mostrou 87,42% para a estrutura. Quando observado no ERRAT (94,6%) e no Verify3D (71,65%), o fator de qualidade geral e compatibilidade confirmam que o modelo apresenta boa qualidade e confiabilidade.

**Figura 6– Alinhamento da sequência de piramboia, bovina, humana, suína e rato de CD302.**

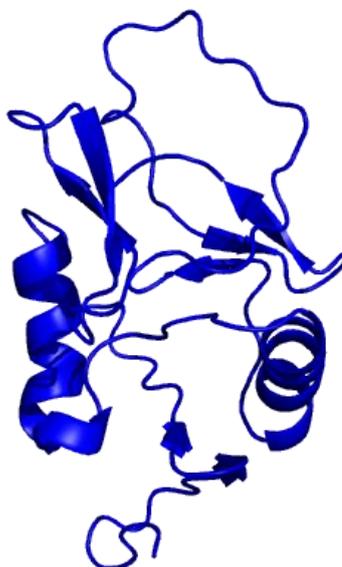


Os resíduos homólogos estão marcados em cinza escuro. O peptídeo sinal está destacado de verde, os resíduos responsáveis pela formação das pontes dissulfeto, em amarelo e os sítios de glicosilação marcados de vermelho. O domínio de reconhecimento de carboidratos está sublinhado de preto e as cadeias  $\alpha$ -hélices e folhas  $\beta$  sinalizadas acima da seqüência. *Lepidosiren paradoxa*, *Bos taurus*, *Homo sapiens*, *Sus scrofa domestica* e *Rattus norvegicus*.

Figura 7 - Esquema da estrutura de *LpCD302*

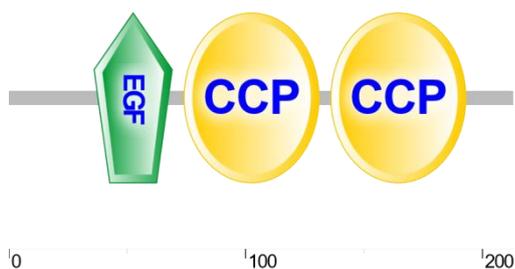


**Figura 8– Estrutura Tridimensional de *LpCD302***



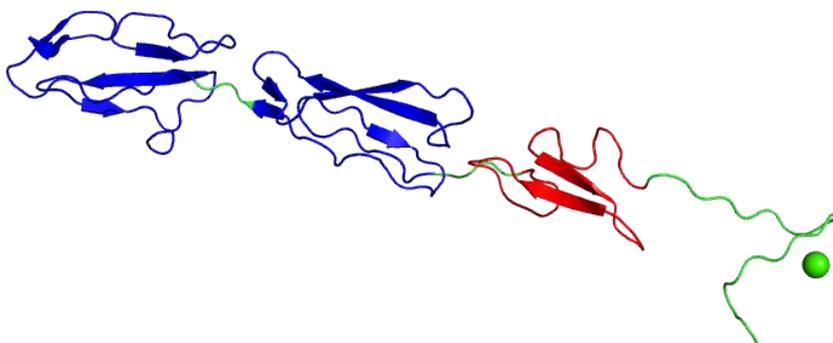
O gene SELP apresenta uma ORF que compreende 642 pb e codifica uma proteína de 213 aminoácidos. O peso molecular previsto e o ponto isoelétrico teórico desta proteína foram 23,34 kDa e 6,24 respectivamente. O índice de instabilidade foi de 39,91, classificando esta proteína como estável. Sendo uma lectina membro da superfamília do tipo C, a P-selectina de piramboia (*LpSELP*) apresentou três diferentes domínios como as outras selectinas, com um domínio de reconhecimento de carboidratos (CRD) na região N-terminal, um domínio semelhante ao fator de crescimento epidermal (EGF, 36aa – 69aa) e dois domínios de ligação ao complemento (CRP, 74aa – 131aa e 136aa – 193aa). O alinhamento da sequência das proteínas revelou a presença de 6 resíduos de cisteínas, que formam três pontes dissulfetos, e um sítio de glicosilação N-terminal.





A estrutura molecular predita revelou que *LpSELP* exibe oito folhas- $\beta$  e o domínio de ligação ao carboidrato se encontra na região C-terminal da proteína (Figura 10). O fator de qualidade geral e confiabilidade foram considerados favoráveis da estrutura tridimensional de *LpSELP* (Figura 14) exibindo gráfico de Ramachandran (94,80%), ERRAT (97,87%) e no Verify3D (67,04%).

**Figura 10 - Estrutura tridimensionais de *LpSELP*.**

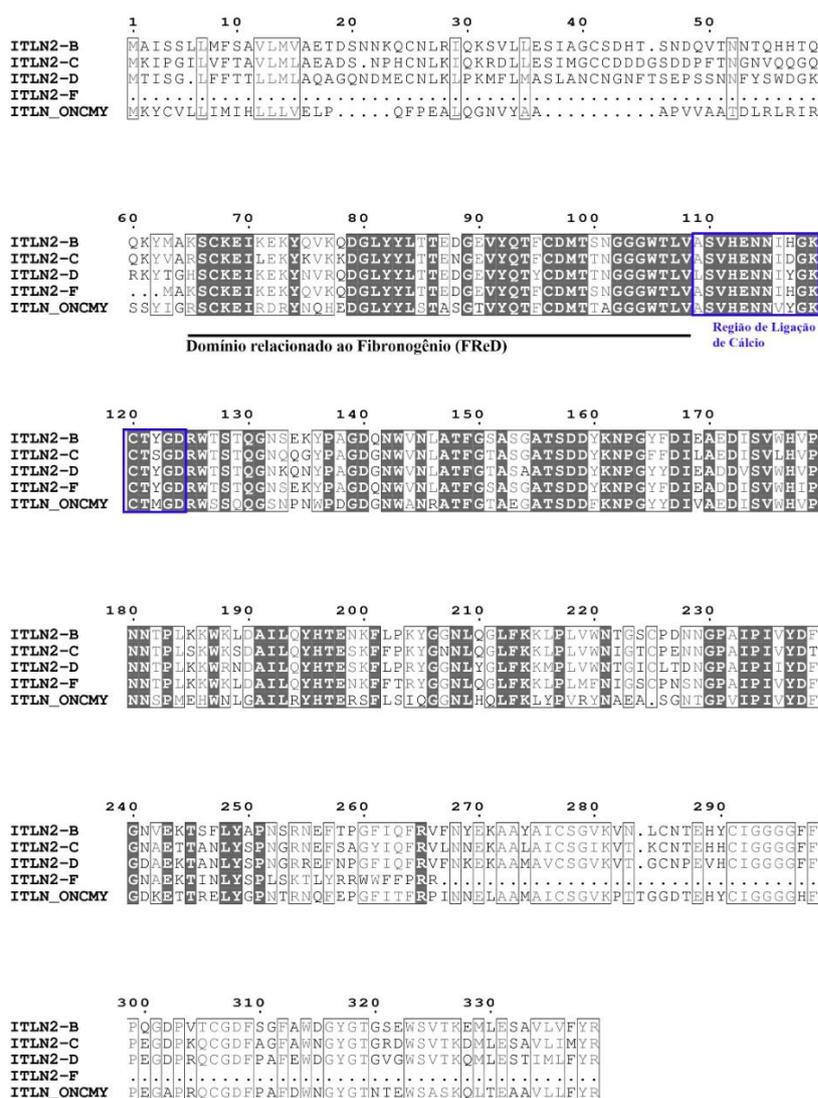


#### 4.2.4 Intelectinas

As sequências de intelectinas são bastante conservadas, apresentando um domínio semelhante ao fibrinogênio (FReD) na região N-terminal (Figura 11) e um domínio específico de lectina na região C-terminal. O domínio FReD contém aproximadamente 142 aminoácidos, enquanto o domínio de intelectina apresenta cerca de 130 aminoácidos. *ITLN2* apresenta uma ORF de 1020 pb, produzindo uma proteína de 339 aminoácidos (Figura 11).

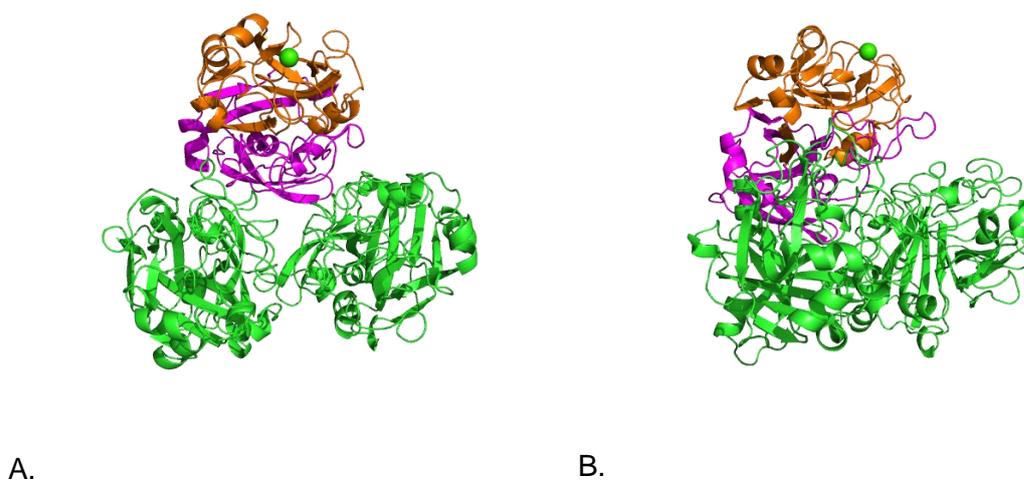
Foram caracterizadas 4 variáveis de *LpITLN2* neste estudo, denominadas: *LpITLN2-B*, *LpITLN2-C*, *LpITLN2-D* e *LpITLN2-F* (Figura 12). As variáveis B, D e F apresentam estruturas semelhantes, apresentando ORFs de 1020 pb e codificando proteínas de 339 aminoácidos, e posição do domínio FReD (62-204 aminoácidos) e domínio intelectina conservados (206-336 aminoácidos). As proteínas apresentam pesos moleculares de 37,77, 37,19 e 37,97 kDa respectivamente, enquanto *LpITLN2-F* possui peso molecular de 23,29 kDa. Os pontos isoelétricos teóricos são respectivamente 5.4, 5.2 e 5.6, enquanto *LpITLN2-F* apresenta ponto isoelétrico de 7,63. As proteínas apresentam índice de instabilidade de 34,81, 27,44, 35,75 e 47,16, sendo as três primeiras sendo consideradas estáveis e a última instável.

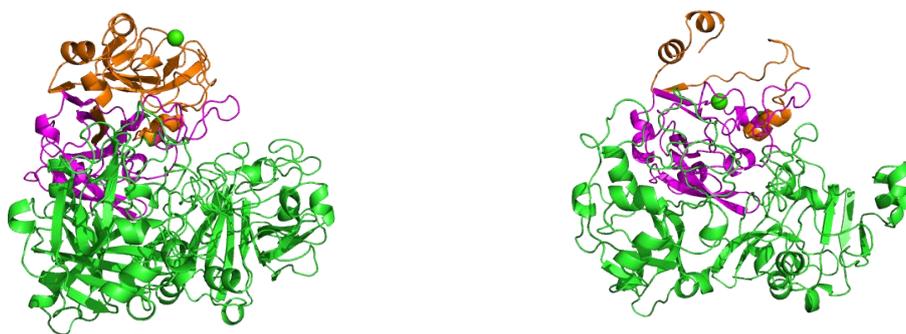
**Figura 11 - Alinhamento das variantes de *LpITLN2* com a ITLN2 da truta arco-íris (*Oncorhynchus mykiss*).**



A estrutura tridimensional de *LpITLN2* é trimérica (Figura 12), contendo 11 folhas- $\beta$ , com exceção da variante *LpITLN2-F* que apresenta apenas 8 folhas- $\beta$ , alternadas por 6 (*LpITLN2-D* e *LpITLN2-F*) ou 7  $\alpha$ -hélices. Todas as variantes apresentam um sítio de ligação ao  $\text{Ca}^{2+}$  e as variantes *LpITLN2-D* e *LpITLN2-F* apresentaram um peptídeo sinal na sua porção N-terminal. O alinhamento das sequências das quatro variantes com a ITLN2 da truta arco-íris (*Oncorhynchus mykiss*), *OnITLN2*, mostrou a presença de 10 resíduos de cisteína conservados em intelectinas. Ademais, o sítio de ligação ao carboidrato dessas proteínas exibe 12 aminoácidos conservados com substituições em dois resíduos de *LpITLN2-C* (H117D e Y122S) e de *LpITLN2-D* (A109L e H117Y). As estruturas tridimensionais exibiram fator de qualidade geral e confiabilidade favoráveis, através do gráfico de Ramachandran (83,3%, 84,4%, 83,8% e 89,9%), ERRAT (92,20%, 94,41%, 91,05% e 91,77%) e no Verify3D (90,29%, 90,73%, 92,42 e 79,53%). Os demais alinhamentos individuais das variantes de *LpITLN2* se encontram no apêndice 2.

**Figura 12 - Estruturas tridimensionais de *LpITLN2*, suas variantes e comparações.**





D.

E.

(A) LpITLN2-B, (B) LpITLN2-C, (C) LpITLN2-D e (D) LpITLN2-F.

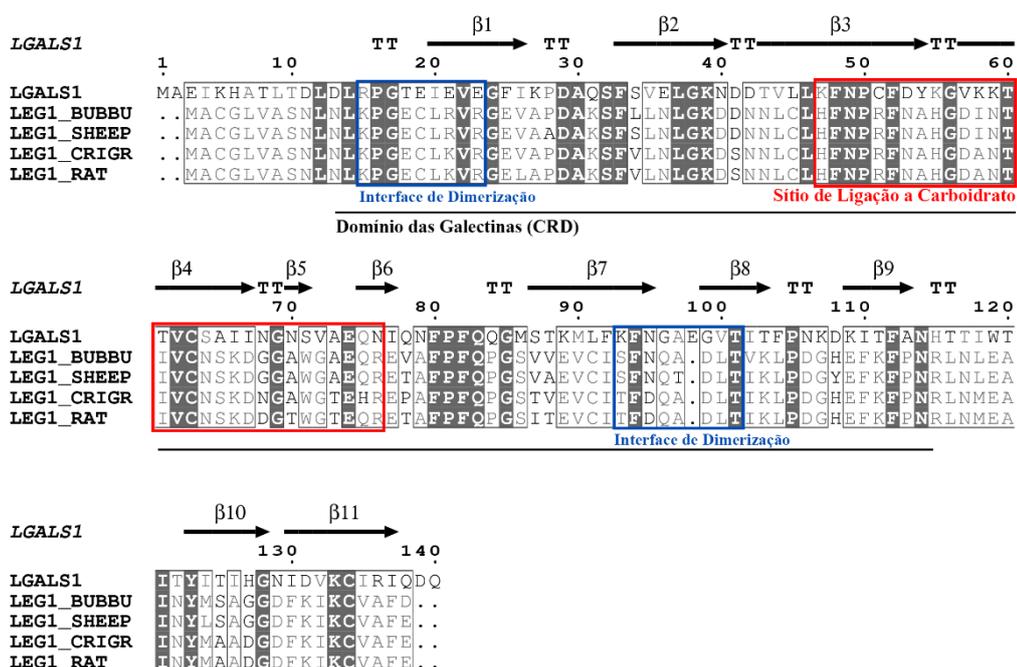
#### 4.2.3 Galectinas

Neste trabalho, três galectinas foram identificadas e caracterizadas através dos transcriptomas da piramboia: LGALS1, LGALS9C e LGALS1. O gene LGALS1 apresenta uma ORF de 423 pb e codifica uma proteína da família de galectina-1 de 140 aminoácidos, a *LpGal-1*. Como um membro protótipo da família das galectinas, a sequência de *LpGal-1* exibe apenas um domínio de reconhecimento de carboidratos (Figura 13). *LpGal-1* apresenta peso molecular de aproximadamente 15,67 kDa e ponto isoelétrico teórico 5,87. O índice de instabilidade desta proteína é de 31,95, portanto é considerada uma proteína estável. O CRD é conservado, ocupando a posição 9-138 aminoácidos, não dependente de cálcio (Figura 14).

A estrutura molecular tridimensional prevista de *LpGal-1* revelou que ela era uma proteína homodimérica não covalente, formada por uma estrutura  $\beta$ -sanduíche com 2 folhas antiparalelas (Figura 15). Uma bolsa de ligação ao açúcar com os resíduos atípicos expostos foi observada na estrutura de *LpGal-1*. Ao invés dos motivos conservados em protótipos de galectinas, HNR, VN e W-ER, *LpGal-1* exibiu os seguintes resíduos: 47K, 49N, 51C, 62V, 64S, 71S, 74E e 76N (KNC, VS e S-EN). A estrutura tridimensional de *LpGal-1* foi validada pelo gráfico de Ramachandran que demonstrou que 86,7% dos resíduos estão em posições

favoráveis energeticamente. O fator de qualidade geral e compatibilidade para as proteínas modeladas foram observados como 89,86% no ERRAT e 75,81% no Verify3D, respectivamente. Esses resultados confirmam que o modelo apresenta boa qualidade e confiabilidade.

Figura 13– Alinhamento da sequência de piramboa, búfalo, ovelha, hamster e rato de LpGal-1



Os resíduos homólogos estão marcados em cinza escuro. O domínio de reconhecimento de carboidratos (CRD) das galectinas está sublinhado de preto e as cadeias  $\alpha$ -hélices e folhas- $\beta$  sinalizadas acima da sequência. Os sítios de ligação a carboidrato e a interface de dimerização estão marcados de vermelho e azul, respectivamente. *Lepidosirenparadoxa*, *Bubalus bubalis*, *Ovis aries*, *Cricetus griseus* e *Rattus norvegicus*

Figura 14– Esquema da estrutura da LpGal-1.

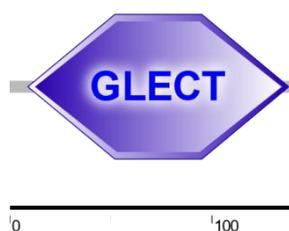
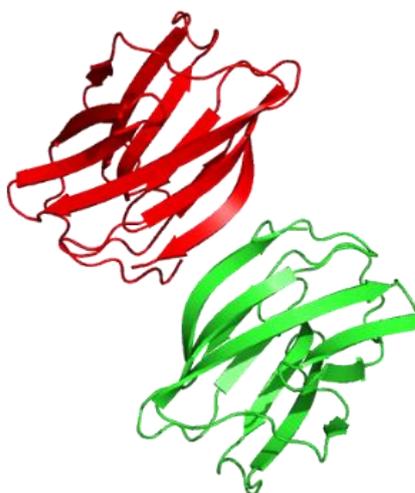


Figura 15– Estrutura tridimensional de *LpGal-1*.

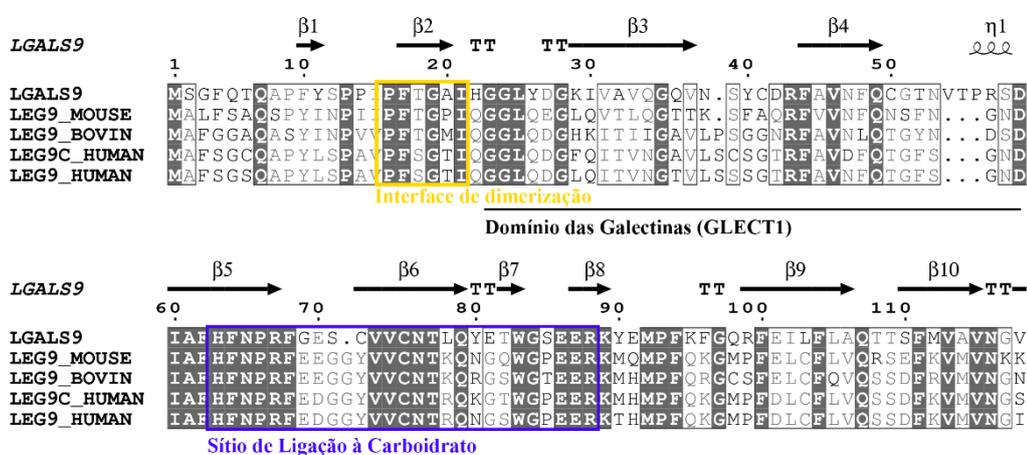


O produto do gene LGALS9 apresenta uma proteína dimérica *em tandem* de 338 aminoácidos (Figura 16), codificada por uma ORF de 1017 pb. A galectina-9 da piramboia, *LpGal-9*, apresenta dois CRDs conectados (15-149 aa e 208-338 aa) (Figura 17). O peso molecular predito de *LpGal-9* é de aproximadamente 37,30 kDa e ponto isoelétrico teórico 6,79. O índice de instabilidade desta proteína é de 36,32, sendo considerada uma proteína estável.

A estrutura secundária de *LpGal-9* revelou que cada CRD exibiu um arranjo  $\beta$ -sanduíche consistindo em folhas  $\beta$ -pregueadas antiparalelas de 10 a 11 fitas, juntamente com  $\alpha$ -hélices curtas e uma região de dobradiça enrolada (Figura 18). Os monômeros N- e C-terminais se encontram em lados opostos do dímero, e os resíduos de aminoácidos responsáveis pela ligação ao açúcar foram localizados em bolsões de ligação de açúcar. *LpGal-9* exibiu motivos de ligação ao glicano conservados em seu CRD N-terminal (H\_NPR e WG-EER) enquanto no CRD C-terminal apresentou uma variação nos resíduos em um dos motivos (H-AVR e WG-EER).

A análise do gráfico de Ramachandran para a estrutura tridimensional de LpGal-9 demonstrou que 85,3% dos resíduos estão posicionados energeticamente favoráveis. Além disso, as análises no ERRAT (89,86%) e no Verify3D (75,81%) conferem a estrutura modelada um fator de qualidade confiável.

Figura 16– Alinhamento de seqüências de LpGal-9 de piramboa com camundongos, bovina e humana.







A proteína do gene LGALS apresenta 178 aminoácidos, codificada por uma ORF de 537 pb (Figura 19). A proteína relacionada com galectina, *LpGal*, é um protótipo de galectina, apresentando um único CRD (43-177 aa) (Figura 20). O peso molecular predito de *LpGal* é de aproximadamente 19,85 kDa e ponto isoelétrico teórico 5,87. O índice de instabilidade desta proteína é de 50,20, sendo considerada uma proteína instável.

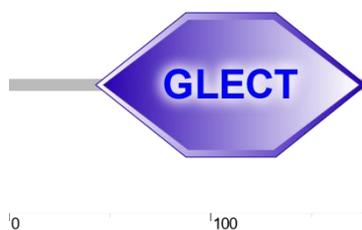
**Figura 19– Sequência de aminoácido de *LpGal*.**

```

1 ATGGCGGGCTCTGTGGCAGCAGAGAAGAACGGAATACTGAGTCCT
  M A G S V A A E K N G I L S P
46 AACAGAGAAGAGGGGCACTTGACCCTACTGAGCAACTCACCGCAG
  N R E E G H L T L L S N S P Q
91 TCTCCAGAGCAAGAATCCCATTACCACCACGTCTGACTGTTCCCT
  S P E Q E F P L P P R L T V P
136 TTTTGTGGTCATATCAAAGGAGGAATGAAACCTGGCAAGAAGATT
  F C G H I K G G M K P G K K I
181 AAAGTGATGGGGATAGTAGATGTCTGTCCAGAAAGTTTGACATC
  K V M G I V D V C P E S F D I
226 AGTTTAACTTGTGGAACATCAGAAGAGCCACCTGCAGATGTGGCT
  S L T C G T S E E P P A D V A
271 TTTGAACTAAAAGTGAGATTTCTCAAAGGCAAGTTGTTCCGAAT
  F E L K V R F P Q R Q V V R N
316 GCTCAAATTGACAGTGAATGGGAAAAGAAGAGCTGTCAATTTCC
  A Q I D S E W G K E E L S I S
361 TACTTCCCTTTCATACCAGATCAGCCATTTAGAGTTGAGATCCAC
  Y F P F I P D Q P F R V E I H
406 TGTGAACACCAGCGCTTCAGAGTATTTGTCGATGGCCACCAGCTG
  C E H Q R F R V F V D G H Q L
451 TTTGATTTCTACCACCGTGTCTAGCACTGCAAGCCATTGATACA
  F D F Y H R V L A L Q A I D T
496 ATAAAGATTAATGGAGACCTTCAGCTGACAAAGCTTGGCTAA
  I K I N G D L Q L T K L G *

```

**Figura 20– Estrutura esquemática de LpGal.**



A modelagem por homologia previu que a estrutura de *LpGal* é monomérica, formada por  $\beta$ -sanduíche com duas folhas antiparalelas (Figura 21). Uma bolsa de ligação e resíduos conservados foi observado, mas sem previsão para peptídeo sinal ou domínio transmembrana. Para a estrutura modelada de *LpGal*, o gráfico de Ramachandran demonstrou que 91,30% dos resíduos estão posicionados energeticamente favoráveis. Além disso, as análises no ERRAT (87,70%) e no Verify3D (80,60%) conferem a estrutura modelada um fator de qualidade geral e confiabilidade.

Figura 21– Estrutura tridimensional de LpGal.



### 4.3 Estudo de docking molecular

As estruturas modeladas foram submetidas ao estudo de docking molecular para demonstrar sua afinidade de ligação aos carboidratos específicos, na presença de Cálcio. Além de analisadas com os carboidratos simples (Figura 22), as proteínas foram testadas também com polissacarídeos de superfície microbiana, LPS e Poli (I:C), para avaliar suas capacidades de se ligarem a bactérias.

As análises confirmam a afinidade da ligação das lectinas do tipo C que apresentam o motivo de ligação QPD, que é galactose específica demonstrando que sua maior afinidade entre os carboidratos foi com a lactose. Já a lectina *LpCD302*, que não apresenta o referido motivo, apresenta maior afinidade com sacarose. As intelectinas apresentam maior afinidade com maltose, enquanto as galectinas apresentam tanto afinidade com lactose, como esperado, quanto com outros açúcares, apresentando com estes até uma ligação mais forte, como sacarose (*LpGal-1*) e maltose (*LpGal-9*).

**Figura 22– Docking molecular entre as lectinas e carboidratos.**

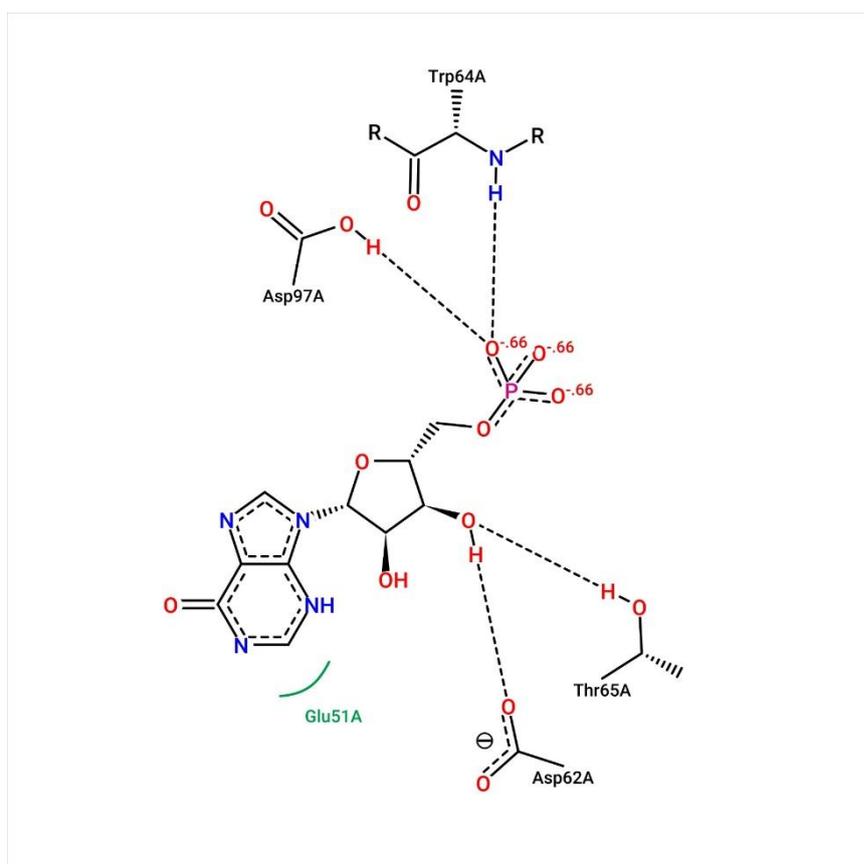
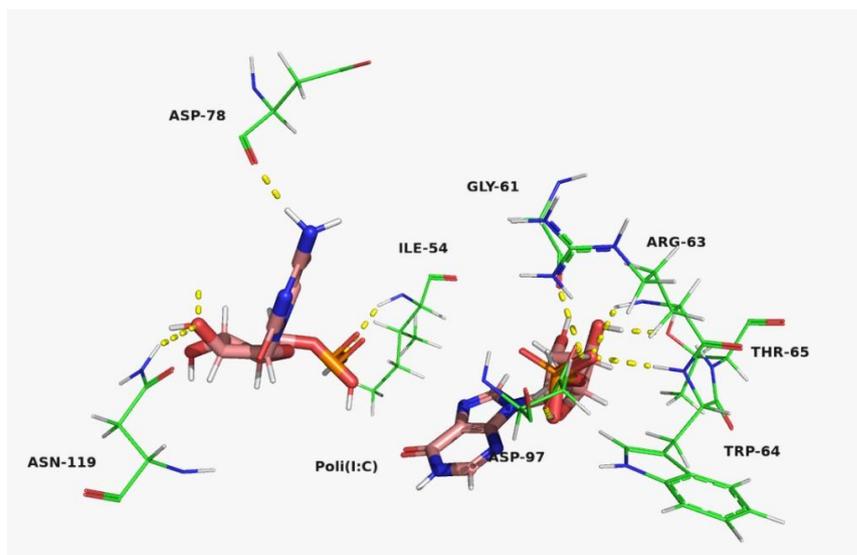
	CD302	CLEC19A-A	CLEC19A-B	CLEC19A-C	ITLN2-B	ITLN2-C	ITLN2-D	ITLN2-F	LGALS1	LGALS9	LGALS	SELP
Arabnose	27,00	48,47	51,46	43,98	55,91	49,46	61,79	50,07	30,59	30,88	39,95	26,94
Frutose	32,70	55,59	57,28	50,06	64,02	53,84	56,83	53,34	36,75	44,21	48,79	31,90
Fucose	26,29	50,56	48,38	44,46	57,06	49,64	61,79	50,79	32,54	32,28	41,36	28,57
Galactose	27,83	52,43	50,16	47,93	60,72	51,38	63,99	51,66	34,81	40,57	45,25	30,43
Glicose	27,86	52,54	50,47	49,13	60,71	52,26	63,99	51,66	35,04	40,55	44,26	30,47
Lactose	35,68	<b>72,42</b>	62,75	<b>64,97</b>	70,71	70,50	70,19	68,46	51,03	59,73	63,44	36,90
Maltose	36,78	68,43	60,69	56,30	<b>72,46</b>	<b>70,66</b>	<b>70,26</b>	69,73	51,30	60,23	59,94	36,24
Manose	27,89	52,30	43,00	47,93	60,53	52,26	64,09	51,66	36,61	41,15	43,83	30,62
N-acetilglicosamina	31,62	54,08	45,99	51,40	62,13	49,54	57,45	53,13	42,04	45,53	52,18	33,53
Ramnose	25,75	50,70	48,38	44,46	57,06	49,64	61,79	50,08	32,14	31,52	41,60	28,04
Ribose	27,75	47,60	54,42	49,30	62,60	51,10	59,12	47,08	35,70	36,47	43,61	29,44
Sacarose	39,87	69,22	61,84	63,60	69,68	62,15	69,73	68,38	52,27	56,65	<b>71,26</b>	38,54

Quanto a afinidade com os ligantes LPS e Poli (I:C) para avaliar a capacidade das proteínas a se ligarem a microrganismos (bactérias e vírus), de maneira geral todas as proteínas apresentaram uma boa afinidade a uma das duas moléculas. Com destaque para as proteínas *LpITLN2-F*, que apresentou maior afinidade com a molécula de Poli (I:C) (Figura 24), sugerindo seu reconhecimento em moléculas de RNA viral, seguida das proteínas *LpCTL-2*, *LpGal-1*, *LpGal-9* e *LpCD302*, que apresentaram afinidade menores. A *LpITLN2-C* não apresentou interação com a Poli (I:C), sugerindo não reconhecer a molécula viral. Porém a *LpITLN2-C* apresenta uma boa afinidade com LPS, sugerindo o reconhecimento e ligação com a parede bacteriana.

**Figura 23– Docking molecular entre as lectinas e LPS/Poli (I:C).**

	CD302	CLEC19A-A	CLEC19A-B	CLEC19A-C	ITLN2-B	ITLN2-C	ITLN2-D	ITLN2-F	LGALS1	LGALS9	LGALS	SELP
LPS	51,20	57,85	39,16	44,85	64,03	62,58	57,23	76,84	52,13	64,37	58,20	51,37
Poli(I:C)	<b>58,93</b>	62,27	<b>64,03</b>	54,90	64,91	-6,39	64,54	<b>106,35</b>	<b>61,08</b>	<b>61,55</b>	59,24	<b>60,94</b>

**Figura 24– Detalhe da ligação entre a lectinas *LpITLN2-F* e o ligante Poli (I:C).**

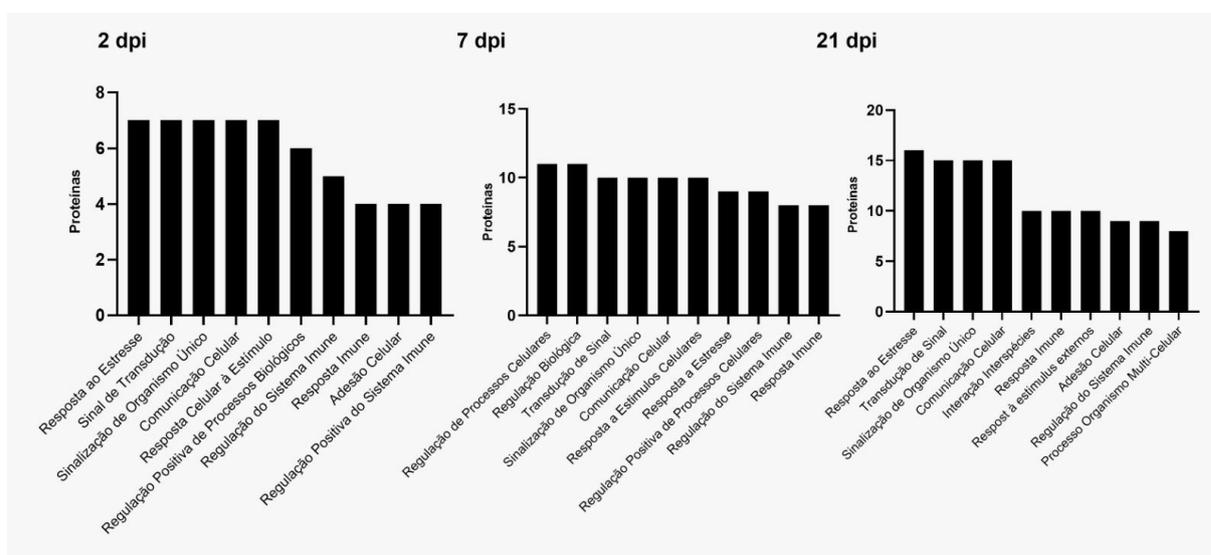


#### 4.4 Anotação funcional das lectinas da piramboia

Para prever em quais funções as lectinas caracterizadas estão envolvidas, foi realizada uma análise de anotação funcional na plataforma DAVID, para três categorias de funções: a) processos biológicos; b) função molecular e c) compartimento celular.

A análise dos processos revelou que as lectinas de piramboia estão envolvidas na resposta ao estresse, na transdução de sinal, sinalização de organismo único, comunicação celular e resposta celular à estímulos. Além disso, também estão envolvidas na regulação do sistema imune e adesão celular

**Figura 25– Análise dos processos biológicos das lectinas da piramboia**



A análise da função molecular mostrou que a maioria das lectinas estão associadas a funções de reconhecimento aos carboidratos, mas também podem desempenhar atividade de receptor e reconhecimento de partículas virais (figura 26).

**Figura 26– Análise funcional das lectinas da piramboia**



## 5 DISCUSSÃO

Ostranscriptomas da piramboia revelaram 78 genes de lectinas em três períodos de regeneração do tecido. De acordo Bonnardelet *al.*(2021) , os organismos aquáticos apresentam um maior repertório de lectinas, uma vez que esses animais utilizam as lectinas para aglutinar bactérias presentes no ambiente. Essa diversidade no repertório de lectinas do tipo C também foi observada no estudo conduzido por Sacoet *al.*(2023). Neste estudo, genes de lectinas do tipo C (CTLs) da piramboia foram identificados e o produto desses genes foram caracterizados. As sequências de aminoácido destas proteínas apresentam o domínio de ligação ao carboidrato bastante conservado, característico desta família de proteínas.

As lectinas do tipo C da piramboia, *Lepidosirenparadoxa*, *LpCTLs*, são compostas por quatro regiões: peptídeo sinal, ponte dissulfeto intra-cadeia, domínio das lectinas tipo C (CTLD) e sítio de ligação ao Cálcio. Outras CTLs de peixes também compartilharam dessa estrutura, como a *PfCTL* do peixe gato amarelo (*Pelteobagrusfulvidraco*) (TANG et al., 2021) e *MaCTL* do dojô (*Misgurnusanguillicaudatus*)(ZHANG et al., 2020). Além disso, a modelagem tridimensional da estrutura dessas proteínas revelou a presença de até três alfa-hélices e 8 folhas-beta.

De uma maneira geral, as CTLs são caracterizadas por apresentarem seus CTLDs conservados, apresentando tanto o motivo de ligação ao carboidrato quanto as quatro cisteínas que formam as duas ligações dissulfetos que estabilizam o domínio de lectina. As CTLs da piramboia apresentam o motivo de ligação galactose-específico, formado pelos aminoácidos glutamina (Gln ou Q), prolina (Pro ou P) e aspartato ou ácido aspártico (Asp ou D), portanto denominado motivo QPD, em oposição ao motivo manose-específico, formado pelos aminoácidos glutamato (Glu ou E), prolina (Pro ou P) e asparagina (Asn ou N), portanto denominado motivo EPN, sugerindo que estas proteínas tenham afinidades a galactose (RAPOSO; CANELAS; BARROS, 2021). Esses motivos estão envolvidos na mediação do

reconhecimento aos glicanos da superfície microbiana. Apesar de serem bem conservados, esses motivos podem apresentar variações, como QPT(substituição do aspartato pela treonina – Thr ou T) ou QPG (substituição do aspartato pela glicina, Gly ou G), EPD(substituição da asparagina peloglutamato – Glu ou D) ou EPQ (substituição da asparagina pela glutamina – Gln ou Q)sem que a sua função seja prejudicada(CHOI et al., 2022).

A presença do motivo WND, importante para a ligação com os carboidratos de uma maneira dependente de  $Ca^{2+}$  também foi identificado. O sítio de ligação de cálcio mostrou ser bem conservado, apresentando os resíduos D<sub>105</sub>, E<sub>109</sub>, D<sub>130</sub> e D<sub>136</sub> comparado com outros organismos. Esses mesmos resíduos também estão presente em *MaCTL*, uma CTL do peixe dojô(ZHANG et al., 2020),demonstrando que essas proteínas são capazes de se ligar ao  $Ca^{2+}$ .

O alinhamento das sequencias das CTLs da piramboia com outros organismos confirmou a presença do peptídeo sinal nas *LpCTLs*, sugerindo que estas proteínas podem ser produzidas e secretadas para outros tecidos com a finalidade de exercer uma função imunológica (ZHANG et al., 2020). Essa suposta função imunológica é corroborada pela capacidade dessas proteínas se ligarem a polissacarídeos de superfícies, presente em bactérias e vírus. As variantes de *LpCTLs* demonstraram afinidades de ligação como lipopolissacarídeo bacteriano, LPS, e com o ácido policitídílico, Poly I:C, sendo essa afinidade, maior para Poly I:C e menor para LPS. Esses resultados sugerem que essas lectinas podem atuar como receptores de reconhecimento padrão que ligam partículas virais e bacterianas.

Outras CTLs de teleósteos demonstraram afinidades a estes polissacarídeos e apresentaram expressão induzida ao organismo ser desafiado com LPS e Poly I:C, como a lectina do tipo C do peixe gato amarelo, *Pelteobagrus fulvidraco*, *PfCTL*, que apresentou sua maior expressão com LPS(TANG et al., 2021). O oposto foi observado nas CTLs de piramboia que apresentaram maior afinidade para Poly I:C, sugerindo que estas proteínas teriam um papel na imunidade contra bactérias e vírus. A CTL do dojô, *Misgurnus anguillicaudatus*, *MaCTL*, também exibiu ligação ao LPS e atividade antibacteriana, sendo capaz de inibir o crescimento tanto de bactérias Gram +/-, demonstrando que essas proteínas desempenham um papel na imunidade bacteriana(ZHANG et al., 2020).

Vários estudos evidenciam o papel na imunidade antimicrobiana das CTLs (BROWN; WILLMENT; WHITEHEAD, 2018; CHOI et al., 2022; HUANG et al., 2020; ZHANG et al., 2022). No entanto para que haja a resposta imune é imprescindível que haja primeiro a ligação ao carboidrato. Em estudos que buscavam entender a modulação de CTLs após estimulação de um ligante de lectina,  $\beta$ -glucano, foi observado diferentes tipos de resposta a este mesmo ligante em organismos diferentes. Enquanto foi sugerido que lectinas podem representar um mecanismo de feedback negativo para impedir a hiperestimulação de células imunes na carpa, durante a infecções fúngicas e bacterianas (PETIT et al., 2019), no salmão pode-se observar a atuação destas proteínas no reconhecimento e na transdução de sinal durante tais infecções (KIRON et al., 2016).

De acordo com Brown *et al.* (2018), CTLs possuem a capacidade de reconhecer ligantes próprios e não próprios. *LpCTLs* demonstraram possuir capacidade de reconhecer carboidratos de maneira dependente de  $Ca^{2+}$  como outras CTLs, que exibiram além da capacidade de se ligarem e aglutinarem microrganismos patogênicos, demonstraram atividade antimicrobiana, inibindo o crescimento desses patógenos. Sugerindo que estas proteínas podem atuar na resposta imune desses animais como receptores de reconhecimento padrão.

Além das *LpCTLs*, este trabalho também caracterizou outras CTLs de piramboia, como o receptor de membrana *LpCD302* e a p-selectina *LpSELP*. *LpCD302* trata-se de um receptor de membrana do tipo C (CRL) que exibiu um domínio de lectina tipo C e demonstrou capacidade de se ligar a maltose. Ademais, a ligação com LPS e Poly I:C foi observada, sendo que a proteína apresentou maior afinidade com Poly I:C. Na tilápia do Nilo, foi identificado um receptor de membrana, semelhante a *LpCD302*, denominado *OnCD302* (ZHANG et al., 2022). Esse receptor demonstrou capacidade de se ligar e aglutinar duas bactérias, *Streptococcus agalactiae* e *Aeromonas hydrophila*. Ambos patógenos foram capazes de induzir a expressão de *OnCD302* na tilápia, porém apenas *S. agalactiae* teve seu crescimento inibido pelo CLR (ZHANG et al., 2022).

A P-selectina é um receptor de membrana que media a adesão molecular em mamíferos e é expressa na membrana de plaquetas e células endoteliais (LUDWIG; SCHÖN; BOEHNCKE, 2007). Em peixes, P-selectina não é frequentemente relatada,

mas sabe-seestar presente no genoma do peixe-zebra (*Danio rerio*) e do fugu (*Takifugurubripes*).Entretanto, não se sabe ao certo em qual tipo de célula essa proteína é expressa(SUN et al., 2010).Na piramboia, *LpSELP* foi expressa durante a regeneração de tecidos e apresentou um domínio de reconhecimento de carboidrato (CRD) na porção N-terminal, um domínio semelhante ao fator de crescimento epidérmico (EGF)e um domínio repetido de ligação ao complemento (CRP).

Estudos comparativos da sequência de aminoácido de P-selectina entre peixes e mamíferos demonstraram que a estrutura primária dessa proteína é relativamente conservada(BAÏSSE et al., 2019; SUN et al., 2010). A P-selectina dopeixe-zebra apresentou um peptídeo sinal de 25 aminoácido e os cinco domínios funcionais conservados de P-selectina: um domínio de ligação a carboidrato na porção N-terminal, um domínio semelhante a EGF, um domínio de dez repetições de ligação ao complemento, um domínio transmembranar e um domínio citoplasmático(SUN et al., 2010).O domínio CRP é variável entre espécies, *LpSELP* por exemplo só apresentou dois domínios CRPs, em comparação com os dez presente no peixe-zebra.Além disso, não foi identificado na P-selectina da piramboia o domínio transmembranar e o domínio citoplasmático relatado em peixe-zebra e em mamíferos.Nota-se que o domínio de lectina e o de fator de crescimento epidérmico parecem determinar a função da P-selectina, que está envolvida com a promoção da inflamação, enquanto a função do CRP permanece desconhecida.

Outra família de lectina identificada nos transcriptomas da piramboia foi as intelectinas (ITLN). A intelectina é um grupo recente de lectinas e recebe esse nome por ter sido identificadas nas células de Paneth do intestino delgado do camundongo (NONNECKE et al., 2022). Porém foi identificada pela primeira vez no ovócito do *Xenopuslaevis* e era denominada de X-lectina (LI et al., 2021). Recebeu essa denominação por não apresentar um domínio de reconhecimento de carboidrato padrão das lectinas. Em seu lugar, havia um domínio semelhante ao fibrinogênio (FReD) e um domínio específico de intelectina na porção C-terminal (CHEN; LI; YANG, 2020).

As intelectinas são amplamente distribuídas, desde placozoários a mamíferos. E já foram relatadas em várias espécies de peixes, como o peixe-zebra, truta arco-íris, carpa, bagre do canal entre outros (CHEN; LI; YANG, 2020). Em mamíferos

foram relatadas apenas duas cópias de genes de intelectina, enquanto peixes apresentam mais cópias. Demonstrando que essas proteínas têm um papel efetivo na resposta imune, devido à pressão ambiental nesses organismos.

A sequência de aminoácido e a estrutura tridimensional das intelectinas são bastante conservadas. O domínio FReD, encontra-se na região N-terminal e apresenta cerca de 200 aminoácidos, sendo a região que apresenta maior variação (CHEN; LI; YANG, 2020). O domínio específico de intelectina, está posicionado na região C-terminal e é mais conservado que a porção N-terminal. A piramboia apresentou quatro variações de INTL-2 no seus transcriptomas.

As variantes de *LpITLN2* apresentam uma ORF de 1020 pb, codificando proteínas de 339 aminoácidos, com exceção da *LpITLN2-F* que apresentou uma ORF de 615 pb e codificou uma proteína de 204 aminoácidos. O alinhamento entre as sequências das *LpITLN2* revelou que as sequências são bem conservadas entre si, apresentando 10 resíduos de cisteínas conservados como as intelectinas da maioria dos mamíferos (CHEN; LI; YANG, 2020). Além disso, sítio de ligação de carboidrato apresentam 12 aminoácidos conservados com variação em dois aminoácidos de duas variantes. A *LpITLN2-C* apresentou mutações nos aminoácidos H117D (substituição de histidina para glutamato) e Y122S (substituição de tirosina para serina). E na *LpITLN2-D* as mutações apresentadas foram nos aminoácidos A109L (substituição de alanina para leucina) e H117Y (substituição de histidina para tirosina, sendo esse último semelhante ao da truta arco-íris (*Oncorhynchus mykiss*). Variações como essa foram observadas entre as intelectinas humana no sítio de ligação aos carboidratos (NONNECKE et al., 2022). No entanto as mutações não interferiram na ligação com cálcio nas variantes das intelectinas de piramboia.

As estruturas tridimensionais das intelectinas de piramboia são triméricas, contendo 11 folhas- $\beta$ , com exceção da variante *LpITLN2-F* que apresenta 8 folhas- $\beta$ , alternadas por 7  $\alpha$ -hélices, com exceção das variantes *LpITLN2-D* e *LpITLN2-F*, que apresentam 6  $\alpha$ -hélices. Semelhantes a estrutura de ITLN1 humana (*hITLN1*), que é uma estrutura trímica, contendo duas folhas- $\beta$  altamente distorcida, cercadas por 7  $\alpha$ -hélices curtas e regiões de bobinas aleatórias (CHEN; LI; YANG, 2020). Ademais, as estruturas das intelectinas da lamprea, *Lethenteron camtschaticum*, (PDB:7E1E) e

de *Xenopus laevis* epidérmica, XEEL, (PDB:5ZC0) (WANGKANONT et al., 2016) também apresentam estruturas similares.

As intelectinas de piramboia demonstraram afinidade de ligação para o dissacarídeo maltose, mas também para outros glicanos, como lactose e N-acetilglicosamina de maneira dependente de Cálcio. Wangkanont(2016) demonstrou que as intelectinas humana e de *Xenopus* apresentam estrutura semelhantes, inclusive no sítio de ligação a carboidrato. E que ambas preferem se ligar a glicanos com grupos 1,2-hidroxila, o que também ocorre com a ligação a N-acetilglicosamina. Corroborando com o que sugere Chen *et al.*(2020), que essas proteínas teriam evoluído em uma única direção para reconhecer microrganismos com epítomos contendo grupos terminais 1,2-hidroxila.

Estudos revelaram que os genes de intelectinas participam das defesas imunológicas contra infecções bacterianas (LI et al., 2021; SINGRANG et al., 2022; WANGKANONT et al., 2016). As lectinas de piramboias, *LpITLN2*, demonstraram afinidade de ligação ao LPS e Poly I:C, sendo que maior afinidade para Poly I:C, com exceção da variante *LpITLN2-C*, que não demonstrou afinidade a Poly I:C. Opondo-se a variante *LpITLN2-F*, que apresentou a maior afinidade a Poly I:C das *LpITLN2*. Sugerindo que as intelectinas de piramboia reconheçam tanto microrganismos bacterianos quanto virais.

Além das infecções bacterianas, as intelectinas estão associadas com condições fisiológicas, como polispermia, desenvolvimento embrionário e metabolismo do ferro (CHEN; LI; YANG, 2020). Mas também participam de processos de dano ao tecido, doenças inflamatória e pulmonar, obesidade e câncer (CHEN; LI; YANG, 2020; CUI et al., 2023; LI et al., 2021; NONNECKE et al., 2022; SINGRANG et al., 2022). Recentemente, Cui et al. (2023), que a intelectina de *Megalobrama ambycephala*(*MaITLN*), um tipo de sargo, aumentava a eficiência fagocitária dos macrófagos. E que LPS havia induzido a expressão da proteína *MaITLN*, sugerindo que estava envolvida na resistência de macrófagos a patógenos (CUI et al., 2023).

Por fim, os transcriptomas da piramboia revelaram outra família de lectinas. Foram identificadas e caracterizadas três galectinas da piramboia, *LpGal-1*,

*LpGal-9* e *LpGal*. As galectinas são membros da família de lectinas que possuem a capacidade de ligação a  $\beta$ -galactosídeos, independente de  $\text{Ca}^{2+}$ , durante a resposta imune contra microrganismos patogênicos (TIAN et al., 2021; WU et al., 2021). Pode-se classificar as galectinas com base em sua estrutura molecular. Desta forma, existem três tipos de galectinas: os protótipos com um único CRD, as quimeras, que apresentam um domínio rico em prolina e glicina na região N-terminal e o tipo repetida em tandem, contendo dois CRDs distintos ligados por um peptídeo (WU et al., 2023). Por mais que estas proteínas não exibam peptídeo sinal, elas podem ser secretadas por uma via não clássica (GANESHALINGAM et al., 2022).

A análise da sequência da galectina de piramboia, membro da família da galectina-1, *LpGal-1*, revelou que a proteína possuía um CRD com importantes sítios de ligação ao carboidrato e que estes diferiam em mamíferos. O que pode influenciar na sua capacidade de ligação e especificidade ao glicano. Isso é corroborado com os resultados de docking molecular que demonstrou a capacidade de *LpGal-1* se ligar a lactose, mas mostrou maior afinidade para maltose e sacarose que não são  $\beta$ -galactosídeos. De acordo com Bonnardelet *al.* (2021), as lectinas podem apresentar multivalência, podendo reconhecer diferentes carboidratos. Isso porque esses carboidratos presentes nos bolsões de ligação são muito importantes para prever a função destas proteínas. A tilápia do Nilo (*Oreochromis niloticus*) também exibiu uma galectina protótipo que não apresentavam os resíduos H e R no sítio de ligação (NIU et al., 2021).

A estrutura tridimensional de *LpGal-1* revelou a dobra de  $\beta$ -sanduíche comum com duas folhas  $\beta$  antiparalelas, exibindo os resíduos no sítio de ligação. *LpGal-1* forma ainda um homodímero por intermédio de interações não covalente entre suas subunidades. Esse tipo de estrutura é comumente encontrado em outras galectinas protótipos (NIU et al., 2021; ZHU et al., 2019). Segundo Liu e Stowell (2023), as galectinas prototípicas existem em equilíbrio monômero-dímero com o estado estacionário diferindo para cada galectina.

Diversas galectinas foram identificadas e caracterizadas, em várias espécies de peixes, como receptores de reconhecimento padrão por serem capazes de reconhecer a porção do carboidrato associada a membrana do patógeno (GANESHALINGAM et al., 2022; LIU et al., 2013; WU et al., 2023; ZHU et

al., 2019). A galectina-1 da piramboia exibiu capacidade de reconhecer os polissacarídeos de membrana LPS e Poly I:C. LPS é um componente representativo da membrana de bactérias, enquanto Poly I:C assemelha-se a estrutura do RNA de fita dupla, sendo, portanto, associado à infecção viral. A galectina-1 da carpa capim (*Ctenopharyngodonidella*), um peixe de importância econômica de água doce, originário da China, demonstrou capacidade de ligação ao LPS e Poly I:C, sendo significativamente modulada pela infecção com LPS e Poly I:C (ZHU et al., 2019). De igual forma, a galectina da tainha vermelha (*Liza hematocheila*), um importante recurso pesqueiro da Coreia do Sul, exibiu a capacidade de se ligar ao LPS e ao Poly I:C, de maneira dose dependente. Esta proteína foi significativamente modulada positivamente pela infecção destas duas moléculas.

A galectina-9 foi outra galectina que foi identificada e caracterizada em nosso trabalho. *LpGal-9* é uma galectina do tipo repetida em tandem, onde os CRDs N-terminal e C-terminal estão conectados por um peptídeo curto. Adicionalmente, *LpGal-9* mostrou possuir, em seus CRDs, os motivos conservados H-NPR e WG-EER que estão associados a interação específica com os  $\beta$ -galactosídeos (LIANG et al., 2020). Embora estes motivos sejam conservados e estão presentes no CRD das maiores das galectinas, existem diferenças de aminoácidos dentro dos motivos entre diferentes espécies (ZHANG; JIANG; SUN, 2021). Estas variações podem afetar na ligação com os carboidratos, seja inibindo a sua capacidade de ligação ao glicano ou apresentando afinidade com os  $\beta$ -galactosídeos. O CRD N-terminal da *LpGal-9* apresentou os motivos de ligação clássicos, mas o CRD C-terminal tem os motivos H-AVR e WG-EER.

A estrutura tridimensional de *LpGal-9* demonstrou o caráter heterodimérico da proteína. O modelo previsto se assemelha ao de outras galectinas do tipo repetida em tandem e a especificidade de ligação a carboidrato de ambos CRDs não parece ter sido afetada pela substituição ocorrida no CRD C-terminal, pois *LpGal-9* mostrou capacidade de ligação a lactose e a maltose. Ao contrário do relatado com a galectina-8 humana, onde a substituição de histidina (H) em H-NPR por asparagina (N) eliminava a capacidade de ligação com o carboidrato (NISHI et al., 2003; STOWELL et al., 2008), enquanto a substituição de arginina (R) em H-NPR por asparagina (N) resultou em perda da ligação de carboidrato pela galectina-9 humana (MATSUSHITA et al., 2000).

Além da capacidade de se ligar aos dissacarídeos, *LpGal-9* demonstrou capacidade de se ligar aos polissacarídeos de membrana LPS e Poly I:C, sugerindo que essa proteína pode atuar na resposta inata a microrganismos patogênicos. Relatórios anteriores demonstraram que galectinas-9 de peixe são capazes de se ligar ao LPS e Poly I:C e podem ligar e aglutinar bactérias, impedindo a infecção em peixes (LIANG et al., 2020; WANG et al., 2021; XU et al., 2022). Como a *LpGal-1*, a galectina-9 da piramboia não apresentou peptídeo sinal nem domínio transmembranar.

De acordo com a análise *in silico*, a sequência de aminoácido de *LpGal* demonstrou ser uma proteína monomérica, apresentando um CRD único que tem resíduos de ligação de açúcar conservados. Trata-se de uma proteína relacionada a galectina do tipo protótipo. A estrutura secundária de *LpGal* exibiu uma configuração semelhante de *LpGal-1*, apresentando dobra em  $\beta$ -sanduíche com folhas antiparalelas, mas monomérica e não homodimérica como a galectina-1 da piramboia. Ratificando a declaração de Liu e Stowell (2023) sobre o equilíbrio monômero-dimérico e que foi mencionado acima.

Como as demais galectinas da piramboia, *LpGal* tanto mostrou capacidade de se ligar a lactose quanto a sacarose, com maior afinidade para esta última. Essa multivalência está associada as substituições que ocorrem nos motivos conservados de ligação ao açúcar. *LpGal* também demonstrou capacidade de interação com LPS e Poly I:C, semelhante a proteína B relacionada a galectina da tainha vermelha (*Liza hematocheila*) (LIM et al., 2023). *LhGal B* demonstrou ser capaz de se ligar a ambos os polissacarídeos de membrana, quanto a infecção induzida por eles foi capaz de modular a expressão dessa proteína na tainha vermelha.

## 6 CONCLUSÃO

- No presente estudo, 12 genes de lectinas foram identificados e caracterizados da piramboia;
- Foram modeladas as estruturas tridimensionais das proteínas traduzidas destes genes, que demonstraram possuir CRDs conservados;
- A capacidade de ligação a carboidratos foi verificada e todas as proteínas exibiram capacidade de ligação ao açúcar;
- As proteínas também demonstraram capacidade de ligação aos polissacarídeos de superfície de patógeno;
- Este estudo sugere que as lectinas descritas podem desempenhar um papel na defesa antimicrobiana.

## REFERÊNCIAS

ABBAS, A. K.; LICHTMAN, A. H.; PILLAI, S. **Basic Immunology E-Book: Functions and Disorders of the Immune System**. [s.l.] Elsevier Health Sciences, 2019.

BAÏSSE, B. et al. The function of P-selectin glycoprotein ligand-1 is conserved from ancestral fishes to mammals. **Journal of Leukocyte Biology**, v. 106, n. 6, p. 1271–1283, 1 dez. 2019.

BERMAN, H. et al. The Protein Data Bank. **Nucleic Acids Research**, v. 28, p. 235–242, 2000.

BONNARDEL, F. et al. LectomeXplore, an update of UniLectin for the discovery of carbohydrate-binding proteins based on a new lectin classification. **Nucleic Acids Research**, v. 49, n. D1, p. D1548–D1554, 8 jan. 2021.

BROWN, G. D.; WILLMENT, J. A.; WHITEHEAD, L. **C-type lectins in immunity and homeostasis**. **Nature Reviews Immunology** Nature Publishing Group, , 1 jun. 2018.

CHEN, L.; LI, J.; YANG, G. **A comparative review of intelectins**. **Scandinavian Journal of Immunology** Blackwell Publishing Ltd, , 1 jul. 2020.

CHOI, M. J. et al. Characterization of a C-Type Lectin Domain-Containing Protein with Antibacterial Activity from Pacific Abalone (*Haliotis discus hannai*). **International Journal of Molecular Sciences**, v. 23, n. 2, 2022.

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

CUI, H. et al. Intelectin enhances the phagocytosis of macrophages via CDC42-WASF2-ARPC2 signaling axis in *Megalobrama amblycephala*. **International Journal of Biological Macromolecules**, v. 236, 1 maio 2023.

EINSENBURG, D.; LUTHY, R.; BOWEI, J. Verify3D: Assessment of protein models with three-dimensional profiles. **Methods in enzymology**, v. 277, p. 396–404, 1997.

GANESHALINGAM, S. et al. Molecular characterization and expression profiling of tandem-repeat galectin-8 from red-spotted grouper (*Epinephelus akaara*): Potential antibacterial, antiviral, and wound healing activities. **Fish and Shellfish Immunology**, v. 121, n. January, p. 86–98, 2022.

HUANG, Y. et al. Characterization and Functional Analysis of Two Transmembrane C-Type Lectins in Obscure Puffer (*Takifugu obscurus*). **Frontiers in Immunology**, v. 11, n. March, p. 1–13, 2020.

JONES, G. et al. Development and validation of a genetic algorithm for flexible docking. **J. Mol. Biol.**, v. 267, p. 727–748, 1997.

KIRON, V. et al. Recognition of purified beta 1,3/1,6 glucan and molecular signalling in the intestine of Atlantic salmon. **Developmental and Comparative Immunology**, v. 56, p. 57–66, 1 mar. 2016.

KO, J. et al. GalaxyWEB server for protein structure prediction and refinement. **Nucleic Acids Research**, v. 40, n. W1, p. 294–297, 2012.

LASKOWSKI, R. A.; THORNTON, M. W.; MACARTHUR, J. M. PROCHECK: validation of protein-structure coordinates. Em: **Crystallography of Biological Macromolecules**. [s.l: s.n.]. p. 722–725.

LEBRETON, A. et al. A comprehensive phylogenetic and bioinformatics survey of lectins in the fungal kingdom. **Journal of Fungi**, v. 7, n. 6, 2021.

LI, J. et al. Characterization of a common carp intelectin gene with bacterial binding and agglutination activity. **Fish and Shellfish Immunology**, v. 108, p. 32–41, 1 jan. 2021.

LIANG, Z. G. et al. Expression and antibacterial analysis of galectin-8 and -9 genes in mandarin fish, *Siniperca chuatsi*. **Fish and Shellfish Immunology**, v. 107, p. 463–468, 1 dez. 2020.

LIM, C. et al. Molecular characterization, immune responses, and functional aspects of atypical prototype galectin from redlip mullet (*Liza haematocheila*) as a

pattern recognition receptor in host immune defense system. **Fish and Shellfish Immunology**, v. 133, 1 fev. 2023.

LINO, M. DA S. et al. Fish Lectins : A Brief Review . n. January 2013, 2015.

LIU, F. T.; STOWELL, S. R. **The role of galectins in immunity and infection. Nature Reviews Immunology** Nature Research, , 2023.

LIU, S. et al. Anti-viral activity of galectin-1 from flounder *Paralichthys olivaceus*. **Fish and Shellfish Immunology**, v. 34, n. 6, p. 1463–1469, 2013.

LUDWIG, R. J.; SCHÖN, M. P.; BOEHNCKE, W. H. **P-selectin: A common therapeutic target for cardiovascular disorders, inflammation and tumour metastasis. Expert Opinion on Therapeutic Targets**, 2007.

MACIEL CARVALHO, E. V. M. et al. Detection of the first lectin with antimicrobial activity present in serum of the Amazonian fish tambaqui *Colossom macropomum*. p. 879–887, 2012.

MATSUSHITA, N. et al. Requirement of divalent galactoside-binding activity of ecalectin/galectin-9 for eosinophil chemoattraction. **Journal of Biological Chemistry**, v. 275, n. 12, p. 8355–8360, 24 mar. 2000.

NISHI, N. et al. Galectin-8 modulates neutrophil function via interaction with integrin  $\alpha$ M. **Glycobiology**, v. 13, n. 11, p. 755–763, nov. 2003.

NIU, J. et al. Characterization and functional analysis of a galectin-related protein B from Nile tilapia involved in the immune response to bacterial infection. **Journal of Fish Diseases**, v. 44, n. 2, p. 171–180, 1 fev. 2021.

NOGUEIRA, A. F. et al. Tetrapod limb and sarcopterygian fin regeneration share a core genetic programme. **Nature Communications**, v. 7, 2 nov. 2016.

NONNECKE, E. B. et al. Human intelectin-2 (ITLN2) is selectively expressed by secretory Paneth cells. **FASEB Journal**, v. 36, n. 3, 1 mar. 2022.

PETIT, J. et al. Studies into  $\beta$ -glucan recognition in fish suggests a key role for the C-Type lectin pathway. **Frontiers in Immunology**, v. 10, n. FEB, 2019.

RAPOSO, C. D.; CANELAS, A. B.; BARROS, M. T. Human lectins, their carbohydrate affinities and where to find them. **Biomolecules**, v. 11, n. 2, p. 1–22, 1 fev. 2021.

RIZZATO, G. A. DETECÇÃO POR ABORDAGEM MOLECULAR DE UMA LECTINA TIPO-C NA VIEIRA *Nodipectenodosus* ( BIVALVIA: PECTINIDAE ) DETECÇÃO POR ABORDAGEM MOLECULAR DE UMA LECTINA TIPO-C NA VIEIRA *Nodipectenodosus* ( BIVALVIA: PECTINIDAE ). 2009.

SACO, A. et al. A Genomic and Transcriptomic Analysis of the C-Type Lectin Gene Family Reveals Highly Expanded and Diversified Repertoires in Bivalves. **Marine Drugs**, v. 21, n. 4, 1 abr. 2023.

SINGRANG, N. et al. Molecular properties and ligand specificity of zebrafish intelectin-2. **Fish and Shellfish Immunology**, v. 123, p. 528–536, 1 abr. 2022.

SMITH, N. C.; RISE, M. L.; CHRISTIAN, S. L. A Comparison of the Innate and Adaptive Immune Systems in Cartilaginous Fish, Ray-Finned Fish, and Lobe-Finned Fish. **Frontiers in Immunology**, v. 10, n. October, 2019.

SOMPAYRAC, L. M. **How the Immune System Works**. [s.l: s.n.].

STOWELL, S. R. et al. Dimeric galectin-8 induces phosphatidylserine exposure in leukocytes through polylectosamine recognition by the C-terminal domain. **Journal of Biological Chemistry**, v. 283, n. 29, p. 20547–20559, 18 jul. 2008.

SUN, G. et al. Molecular cloning and expression analysis of P-selectin from zebrafish (*Danio rerio*). **International Journal of Molecular Sciences**, v. 11, n. 11, p. 4618–4630, nov. 2010.

TANG, X. H. et al. Molecular characterization and expression analysis of a novel C-type lectin (CTL) gene in yellow catfish *Pelteobagrus fulvidraco*. **Aquaculture Reports**, v. 20, n. December 2020, p. 100640, 2021.

TIAN, M. et al. Galectins in turbot (*Scophthalmus maximus* L.): Characterization and expression profiling in mucosal tissues. **Fish and Shellfish Immunology**, v. 109, p. 71–81, 1 fev. 2021.

VERDONK, M. L. et al. Improved Protein–Ligand Docking Using GOLD Marcel. **Proteins**, v. 52, n. November 2002, p. 609–623, 2003.

WANG, L. et al. The N- and C-terminal carbohydrate recognition domains of galectin-9 from *Carassius auratus* contribute differently to its immunity functions to *Aeromonas hydrophila* and *Staphylococcus aureus*. **Journal of Fish Diseases**, v. 44, n. 11, p. 1865–1873, 1 nov. 2021.

WANGKANONT, K. et al. Structures of *Xenopus* embryonic epidermal lectin reveal a conserved mechanism of microbial glycan recognition. **Journal of Biological Chemistry**, v. 291, n. 11, p. 5596–5610, 11 mar. 2016.

WEBB, B.; SALI, A. Comparative protein structure modeling using MODELLER. **Current Protocols in Bioinformatics**, v. 2016, n. June, p. 5.6.1-5.6.37, 2016.

WU, B. et al. Characterization and functional study of a chimera galectin from yellow drum *Nibeaalbiflora*. **International Journal of Biological Macromolecules**, v. 187, n. April, p. 361–372, 2021.

WU, B. et al. Molecular Cloning and Functional Characterization of Galectin-1 in Yellow Drum (*Nibeaalbiflora*). **International Journal of Molecular Sciences**, v. 24, n. 4, 1 fev. 2023.

XU, H. et al. Molecular characterization and antibacterial ability of galectin-3 and galectin-9 in *Onychostomamacrolepis*. **Developmental and Comparative Immunology**, v. 128, 1 mar. 2022.

ZHANG, T.; JIANG, S.; SUN, L. A fish galectin-8 possesses direct bactericidal activity. **International Journal of Molecular Sciences**, v. 22, n. 1, p. 1–15, 1 jan. 2021.

ZHANG, X. W. et al. A C-type lectin with antibacterial activity in weather loach, *Misgurnusanguillicaudatus*. **Journal of Fish Diseases**, v. 43, n. 12, p. 1531–1539, 2020.

ZHANG, Z. et al. Molecular characterization of a novel C-type lectin receptors (CD302) in Nile tilapia (*Oreochromis niloticus*) and its functional analysis in host defense against bacterial infection. **Aquaculture Reports**, v. 27, n. July, 2022.

ZHU, D. et al. Molecular characterization, tissue distribution and functional analysis of galectin 1-like 2 in grass carp (*Ctenopharyngodonidella*). *Fish and Shellfish Immunology*, v. 94, p. 455–463, 1 nov. 2019.

## APÊNDICE 1 – SEQUÊNCIAS NUCLEOTÍDEAS DAS LECTINAS

### Sequência de nucleotídeos e aminoácidos de *LpCTL-1*

```

1 ATGAATTACCTGATACTGATGTCTGTACTCCTGGCTGCCTTTATT
  M N Y L I L M S V L L A A F I
46 CCTGGTGGTGACACTGCAGATGGCTGTCTCTGTTATCAGGGGGAC
  P G G D T A D G C L C Y Q G D
91 TGCTCATTGGATGGTTCCAGTACAAGAATTCATGCTTTCGTCCC
  C S F G W F Q Y K N S C F R P
136 ATACAAATAAAAAATGACATGGGTTAATGCTGAGGCCTACTGCCAA
  I Q I K M T W V N A E A Y C Q
181 AGAAACTACTACGGAGCTCATTGGCATCCATCCACAGTCTTGAA
  R N Y Y G A H L A S I H S L E
226 GAAAATAACTACATTTATGAACTAATGGGTAGACCCAATGATTAC
  E N N Y I Y E L M G R P N D Y
271 ATAAAAGGACAGGCTTACTGGATTGGATTACATGACACTTTTGTA
  I K G Q A Y W I G L H D T F V
316 GAGGGTGTTTTTATGTGGACTGATGGAAGTCCAGTGAACCTCCTA
  E G V F M W T D G S P V N F L
361 CGTTTTGGCAAATCCCAACCAGACAATCTGGCAATGAAGATTAT
  R F G K S Q P D N L G N E D Y
406 GTTGGTGCATGGTATTTCCAGAATGGTACTATTACCTGGAATGAC
  V G S W Y F Q N G T I T W N D
451 TACAGTAGCAGATGGTCTTCCCTTTTGTCTGCAAATATGACCTT
  Y S S R W S F P F V C K Y D L
496 CGAAGATGCTCTGTATACGCACCAATATTCTAA
  R R C S V Y A P I F *

```

### Sequência de nucleotídeos e aminoácidos de *LpCD302*.

```

1 1 ATGCAGAAAGTGAAAAAGGCATCTGCACTGCTGTTTTGTTGGGC
2   M Q K V K K A S A L L F C W G
3
4 46 TGTGTTTTATTTGTATTCCATTGCCCTTCCGTTATGTGTTTAGAA
5   C F Y F V F H C P S V M C L E
6
7 91 TGTCCATCTGGGACTAAGATTTTCATGGATATCGTATAAGAACAGA
8   C P S G T K I S W I S Y K N R
9
10 136 TGTATACTGTTTTGGATGAATTTTCAAACATTCCCTTGGTATA
11   C Y T V L D E F S K H S L G I
12
13 181 GAAGATGCAAGGGAAGCTCTGAGAGCTTCCAGTTCCAGTGTGAC
14   E D A R E L C R A S A S S A D
15
16 226 CTTGTGAGTGTGAACAGTGAAGAAAGAAATGTGTTTTAGTAAAC
17   L V S V N S E E E N V F L V N
18
19 271 ACTTACAGGACAAATTGGCAGGACCTGAAGAAGTTTGGCTTGGC
20   T Y R T N W Q G P E E V W L G
21
22 316 ATGTTTTATGATACAGACACGGATAGTTTGAATGGTTTGACAAC
23   M F Y D T D T D S L K W F D N
24
25 361 TCAGAAGTGCATATTCGAAGTGGACTAAACAATACAGTGAAGAT
26   S E V S Y S N W T K Q Y S E D
27
28 406 TTGCCTCTTGGCAGATATGTGCCGTCATTCATATCAAATCTGGA
29   L P L A E I C A V I H I K S G
30
31 451 CACTGGAGCAAGGTGACAAGAAGGCTTTGA
32   H W S K V T R R L *

```

### Sequência de nucleotídeos e de aminoácido *LpSELP*

```

1 ATGATGAAGATTGTGTCGAGATTTACATTAATAAGACAGAAGCTCTA
   M M K I V S R F T L K D R T L
361 TCTGGGACCTGGGATGGTGAACAACCCACATGCAAAGCTGTAAC
   S G T W D G E Q P T C K A V N
46 GTGGACAGTGAATGATGAACCTTGACAAAAAAGAAGGCTGCT
   V D S G M M N L A Q K K K A A
406 TGTGGGGAAGCTTAAGCAGCCTGAAAATAGCTCTATGAATTGCTCC
   C G E L K Q P E N S S M N C S
91 CTGTGTTACACAGCTTCATGCAATTCAGCTCTTGCAGTGGGCAT
   L C Y T A S C N S S S C S G H
451 CATCCTAATGGAAACTTTTTTTATAATCCACATGTTTATTCAAT
   H P N G N F F Y N S T C L F N
136 GGAGAGTGTGTGAACACATAAATAACCACAGCTGCATCTGCCAC
   G E C V E H I N N H S C I C H
496 TGCCCTGAGGGCTTCCAGCTCAGAGGATCAGAGAACTCCATTGC
   C P E G F Q L R G S E K L H C
181 GAAGGGTTTTATGGTCCAAATGCGAATGCTGTGGTCTGTAAC
   E G F Y G P K C E H A V V C N
541 CTGTCTGAAGGAAATTGGAGCAATGAAATGCCTAAATGTGAAGGT
   L S E G N W S N E M P K C E G
226 GCATTGGTGAACAGATAATCTGAATATGCAATGCTCCTCTCCC
   A L V E P D N L N M Q C S S P
586 AAGAAACATTCTCCAGTAGAAGTGTCTAAATATCGGGACCAGGC
   K K H S P V E V S K L S G P G
271 CTTGGTATTCCACTATGGCTCGGTCTGCAATTTTCAGCTGTAAG
   L G D F H Y G S V C N F S C K
631 CATATACTGTGA
   H I L *
316 GAAGGATTCCTCTGCATGGAGCAAATGAGATGCGATGCAAGTAA
   E G F L L H G A N E M R C S E

```

### Sequência de nucleotídeos e de aminoácido *LpGal-1*.

```

1 ATGGCAGAGATAAAGCATGCTACCTTAACTGACCTGGATCTGCGT
  M A E I K H A T L T D L D L R

46 CCAGGCACAGAGATTGAAGTTGAAGGATTTATAAAGCCTGACGCT
  P G T E I E V E G F I K P D A

91 CAAAGCTTTTCTGTAGAACTGGGAAAAAATGATGACACTGTTCTC
  Q S F S V E L G K N D D T V L

136 CTGAAATTCATCCTTGTTTTGATTACAAGGGTGTTAAAAAGACA
  L K F N P C F D Y K G V K K T

181 ACTGTGTGCAGCGCTATAATTAATGGAAATTCGTCCGAGAACAG
  T V C S A I I N G N S V A E Q

226 AACATCCAAAACCTTTCCCTTTCAGCAAGGAATGTCAACCAAGATG
  N I Q N F P F Q Q G M S T K M

271 CTGTTCAAATTCATGGTGCCGAAGGCGTCACTATAACATTTCCA
  L F K F N G A E G V T I T F P

316 AATAAAGATAAAAATAACCTTCGCTAACCCACACAACCATTTGGACA
  N K D K I T F A N H T T I W T

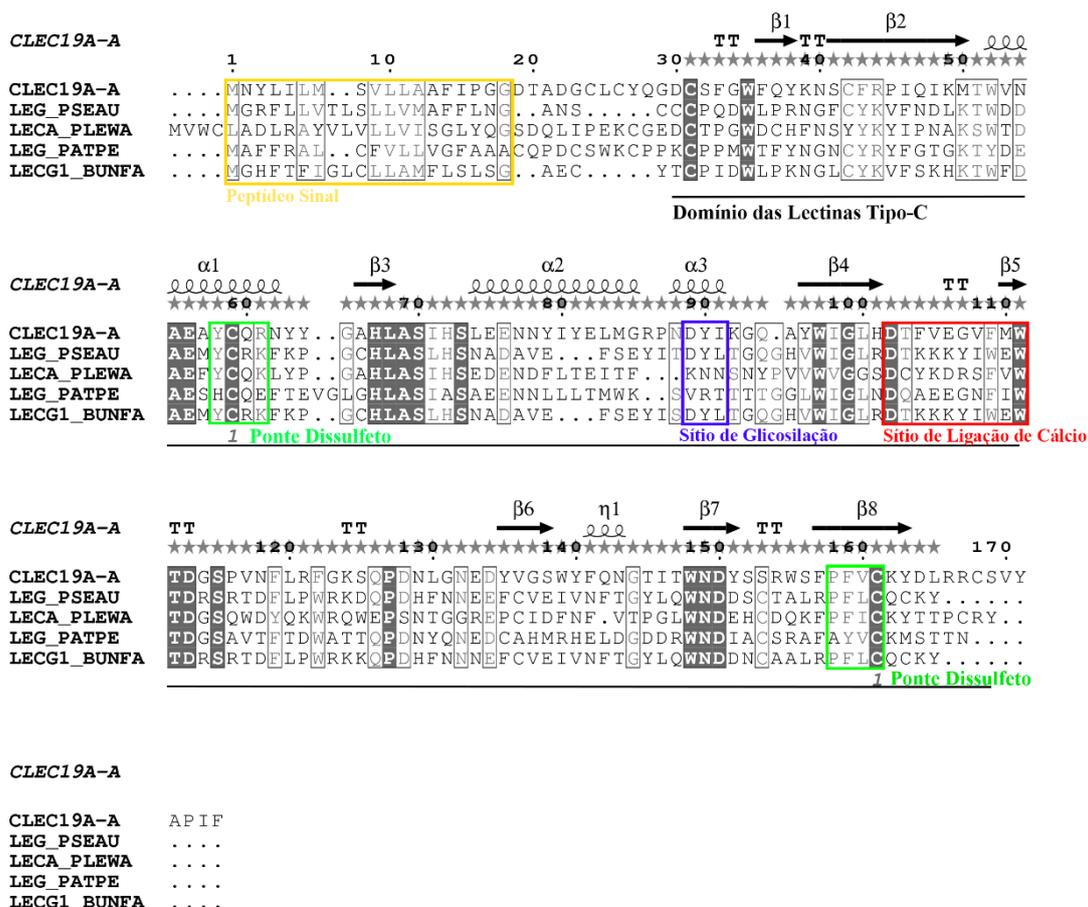
361 ATAACCTATATTACCATCCATGGTAACATCGATGTGAAATGCATC
  I T Y I T I H G N I D V K C I

406 AGGATTCAGGACCAAGTGA
  R I Q D Q *

```

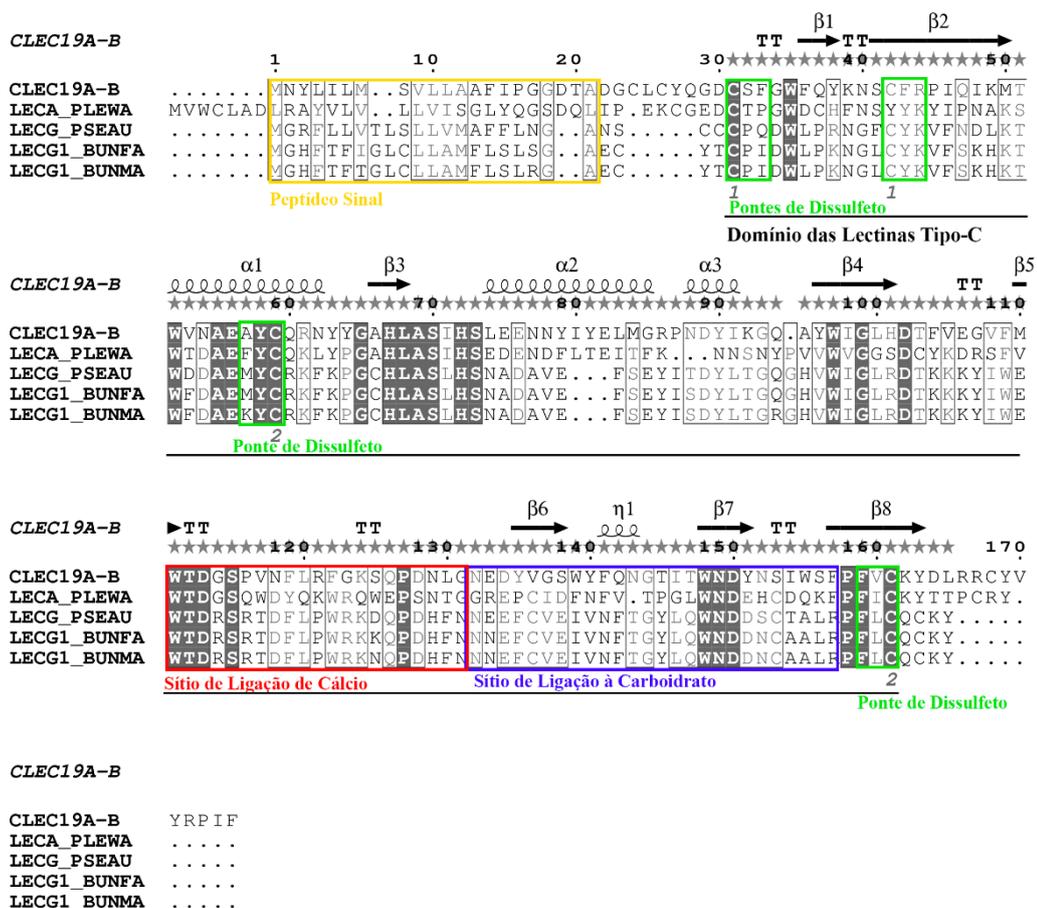
## APÊNDICE 2 – ALINHAMENTOS DAS LECTINAS DA PIRAMBOIA

**Alinhamentos de sequências da lectina do tipo C de piramboia com outros organismos.** Os resíduos homólogos estão marcados em cinza escuro. O peptídeo sinal está destacado de amarelo, os resíduos responsáveis pela formação das pontes dissulfeto, em verde e os sítios de glicosilação e de ligação de cálcio marcados de azul e vermelho, respectivamente. O domínio de reconhecimento de carboidratos está sublinhado de preto e as cadeias  $\alpha$ -hélices e folhas  $\beta$  sinalizadas acima da sequência.

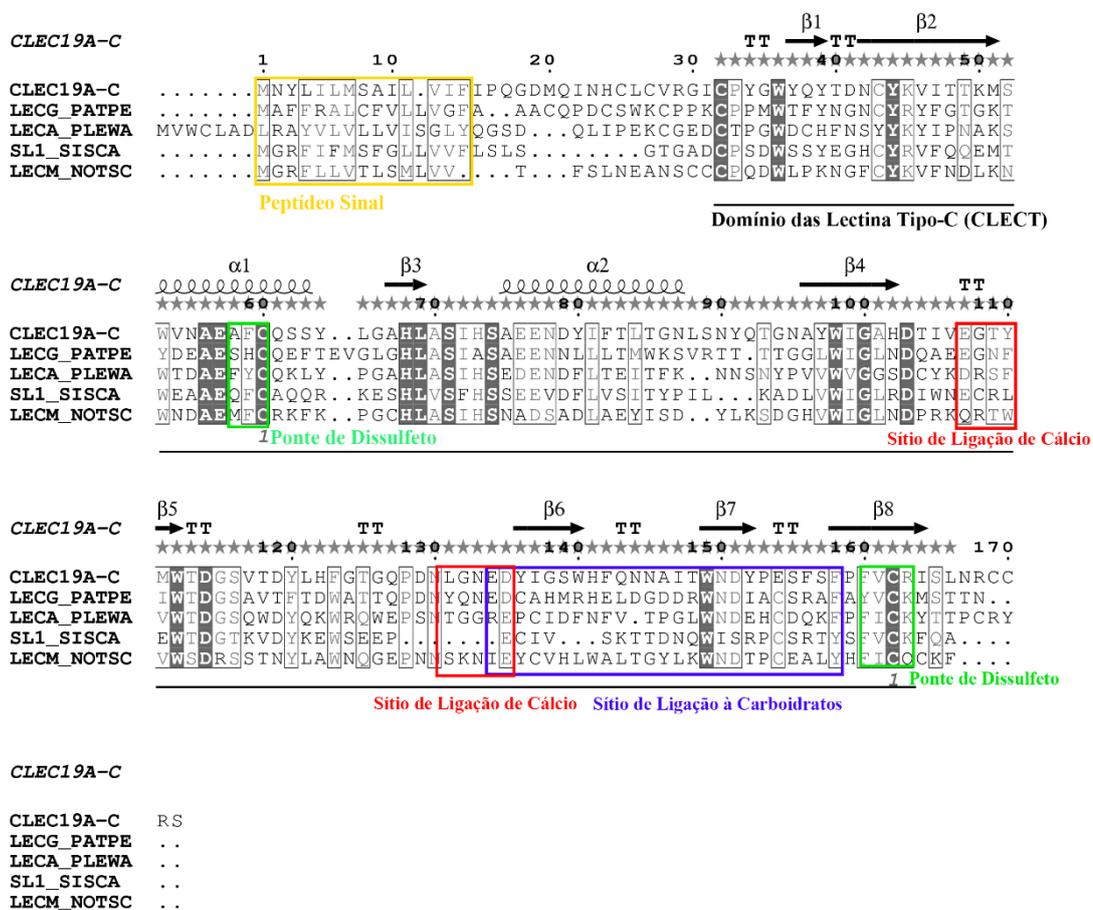


Alinhamento de seqüências de LpCTL-1 com serpentes, salamandra e estrela-do-mar. *Lepidosirenparadoxa*, *Pseudechisaustralis*, *Pleurodeleswalti*, *Patiriapectinifera* e *Bungarusfasciatus*

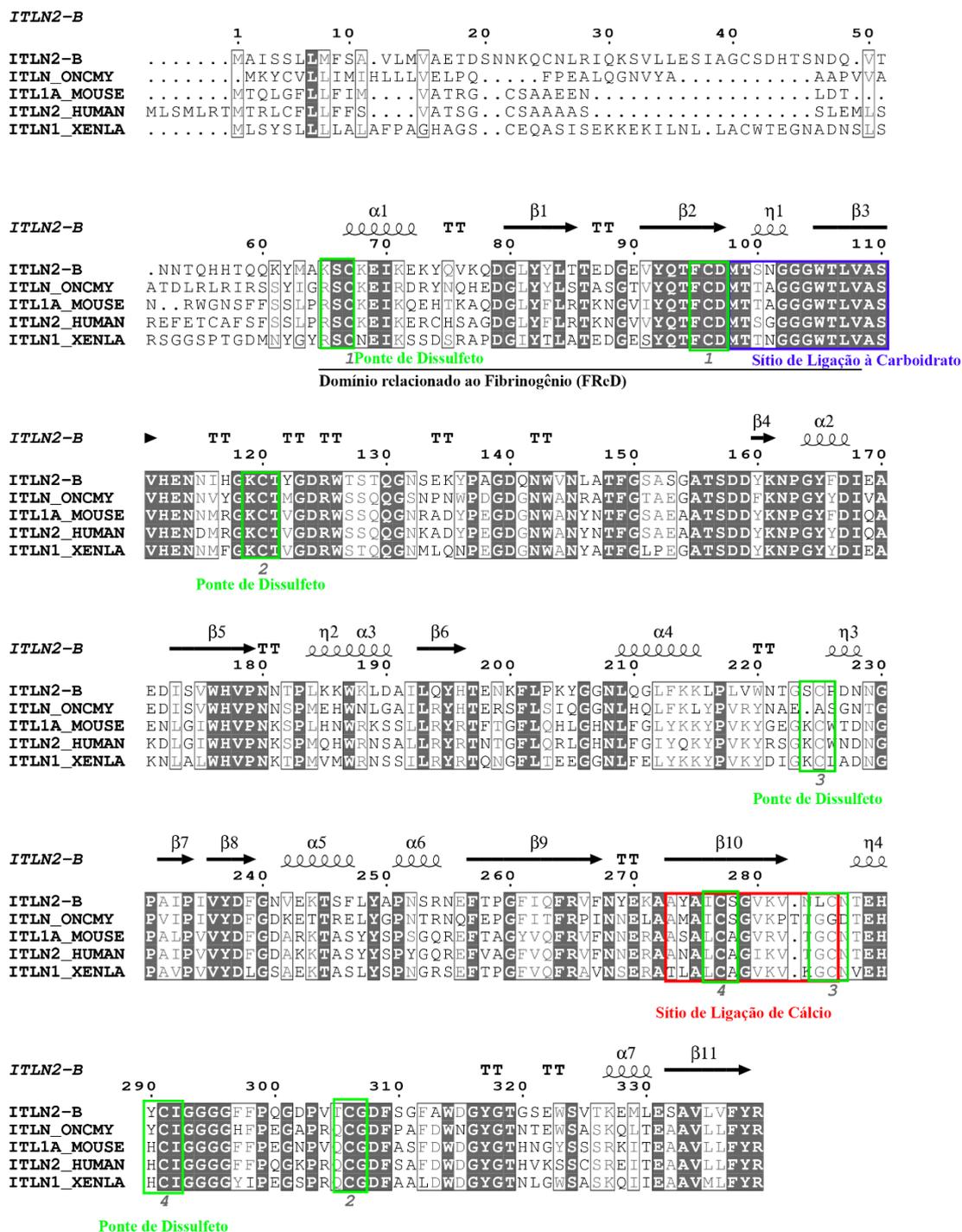
Alinhamento de seqüências de LpCTL-2 com salamandra e serpentes. *Lepidosirenparadoxa*, *Pleurodeleswalti*, *Pseudechisaustralis*, *Bungarusfasciatus* e *Bungarusmulticinctus*.



Alinhamento de sequências de LpCTL-3 com estrela-do-mar, salamandrae serpentes. *Lepidosirenparadoxa*, *Patiriapectinifera*, *Pleurodeleswalti*, *Sistruruscatenatusedwardsii* e *Notechisscutatus*.

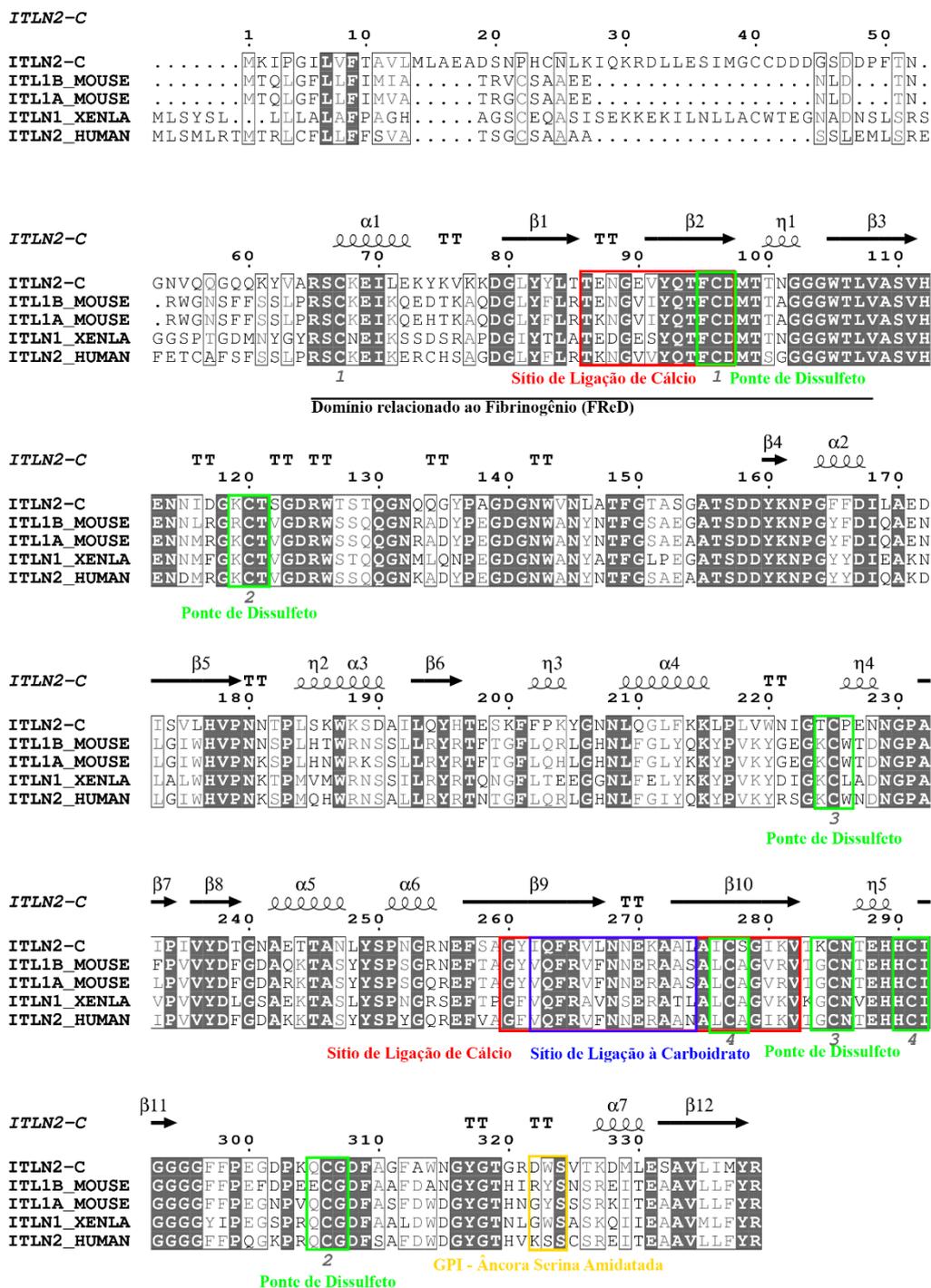


Alinhamento da sequência de piramboa, camundongo, humano e anuro de *LpITLN2-B*.



Os resíduos homólogos estão marcados em cinza escuro. O domínio semelhante ao fibrinogênio (FRcD) está sublinhado de preto e as cadeias  $\alpha$ -hélices e folhas  $\beta$  sinalizadas acima da sequência. Os resíduos responsáveis pela formação das pontes dissulfeto estão marcados em amarelo e os sítios de ligação a carboidrato e de ligação ao cálcio estão marcados de azul e vermelho, respectivamente. *Lepidosiren paradoxa*, *Oncorhynchus mykiss*, *Mus musculus*, *Homo sapiens* e *Xenopus laevis*.

**Alinhamento de sequências de LpITLN2-C com camundongos, sapo e humana. *Lepidosiren paradoxa*, *Mus musculus*, *Xenopus laevis* e *Homo sapiens*.**



Alinhamento de seqüências de pirâmboia *LpITLN2-D* com sapo, humana e camundongo. *Lepidosiren paradoxa*, *Xenopus laevis*, *Homo sapiens* e *Mus musculus*.

*ITLN2-D*

	1	10	20	30	40	50
<i>ITLN2-D</i>	.....	MTISGLFFTTL	..LMLAQAGNDMECN	LKLPKMFMLMASLANCNGN	.FTSEPPSS	
<i>ITLN_XENLA</i>	.....	MLSYSLLLLLAL	AFFPAGHAGSC	EQASISEKKEKIL	...NLLACWTEGNADNSLS	
<i>ITLN2_HUMAN</i>	MLSMRLRT	MTRLCFLFLFFS	...VATSGCSAAAAA	S.....	SLEMLS	
<i>ITL1A_MOUSE</i>	.....	MTQLGFLFLFIM	...VATRGC	SAAEEN	.....LDT...	
<i>ITL1B_MOUSE</i>	.....	MTQLGFLFLFIM	...IATRVC	SAAEEN	.....LDT...	

*ITLN2-D*

	60	70	80	90	100	110
		$\alpha 1$	TT	$\beta 1$	TT	$\beta 2$
		$\eta 1$				$\beta 3$
<i>ITLN2-D</i>	NNFYSWD	GKRRY	TGHSCREIK	EKEYNVR	ODGLIY	LTTEDEGEVY
<i>ITLN_XENLA</i>	RSGGSPT	GDMNY	GYRSCNEIK	SSDSRAP	DGIYTL	ATEDGESY
<i>ITLN2_HUMAN</i>	REFETCA	FSFSL	PRSCREIK	ERCHSAG	DGLYFL	RTKNGV
<i>ITL1A_MOUSE</i>	N..RWGNS	FFSS	LPSCREIK	QEHTKA	ODGLYFL	RTKNGV
<i>ITL1B_MOUSE</i>	N..RWGNS	FFSS	LPSCREIK	QEDTKA	ODGLYFL	RTENGV

Ponte de Dissulfeto 1      Sítio de Ligação de Cálcio 1

**Domínio relacionado ao Fibrinogênio (FRcD)**

*ITLN2-D*

	120	130	140	150	160	170
	TT	TT TT	TT	TT	$\beta 4$	$\alpha 2$
<i>ITLN2-D</i>	VHENNIY	GRCTV	GDRWTS	TQGNKQNY	PAGDGNW	VNLA
<i>ITLN_XENLA</i>	VHENNMF	GRCTV	GDRWST	QGNMLO	NPEGDGN	WANY
<i>ITLN2_HUMAN</i>	VHENDMR	GRCTV	GDRWSS	QGNKADY	PEGDGN	WANY
<i>ITL1A_MOUSE</i>	VHENNMR	GRCTV	GDRWSS	QGNRADY	PEGDGN	WANY
<i>ITL1B_MOUSE</i>	VHENNLR	GRCTV	GDRWSS	QGNRADY	PEGDGN	WANY

Ponte de Dissulfeto 2

*ITLN2-D*

	180	190	200	210	220	230
	$\beta 5$	TT	$\eta 2$	$\alpha 3$	$\beta 6$	$\eta 3$
						$\alpha 4$
<i>ITLN2-D</i>	DDVSV	WHVPNN	TPLKK	WRNDA	ILOYHT	ESKFL
<i>ITLN_XENLA</i>	KNLAL	LWHVPNK	TPMVM	WRNNS	ILRYRT	ONGFL
<i>ITLN2_HUMAN</i>	KDLGI	IWHVPNK	SPMQH	WRNSA	LLRYRT	NTGFL
<i>ITL1A_MOUSE</i>	ENLGI	IWHVPNK	SPLHN	WRKSS	LLRYRT	FTGFL
<i>ITL1B_MOUSE</i>	ENLGI	IWHVPNN	SPLHT	WRNSS	LLRYRT	FTGFL

Ponte de Dissulfeto 3

*ITLN2-D*

	240	250	260	270	280	290
	$\beta 7$	$\beta 8$	$\alpha 5$	$\eta 4$	$\beta 9$	TT
						$\beta 10$
						$\eta 5$
<i>ITLN2-D</i>	PAIPVI	YDFG	DAEKTAN	LYSPN	GRREF	NFC
<i>ITLN_XENLA</i>	PAPV	VVYD	LGS	AEKTAS	LYSPN	GRSE
<i>ITLN2_HUMAN</i>	PAIP	VVYD	FGDA	AKKTAS	YSPY	QREF
<i>ITL1A_MOUSE</i>	PALP	VVYD	FGDA	ARKTAS	YSPS	GRQ
<i>ITL1B_MOUSE</i>	PAFP	VVYD	FGDA	OKTAS	YSPS	GRNE

Sítio de Ligação de Cálcio      Sítio de Ligação à Carboidrato      Ponte de Dissulfeto 4

*ITLN2-D*

	300	310	320	330
		TT	TT	$\alpha 6$
				$\beta 11$
<i>ITLN2-D</i>	CIGGGG	FPE	EGDPR	CCGDF
<i>ITLN_XENLA</i>	CIGGGG	YIPE	GS	PRDCC
<i>ITLN2_HUMAN</i>	CIGGGG	FPE	EGKPR	CCGDF
<i>ITL1A_MOUSE</i>	CIGGGG	FPE	EGNPV	CCGDF
<i>ITL1B_MOUSE</i>	CIGGGG	FPE	EPDPE	CCGDF

Ponte de Dissulfeto 4      GPI - Âncora Serina Amidatada

Alinhamento de seqüências de LpITLN2-F com peixe, camundongos, humana e anuro. *Lepidosiren paradoxa*, *Onchorhynchus mykiss*, *Mus musculus*, *Homo sapiens* e *Xenopus laevis*.

ITLN2-F

```

ITLN2-F      .....
ITLN_ONCMY  .MKYCVLLIMIHLLLVLELPQFPEALQGNVYAAAP.....VVAAT
ITL1B_MOUSE .....MTQLGF...LLFIMIATRVCSA...NLD...TN.
ITLN2_HUMAN MLSM..LRTMTRL...LLFFSVATSGCSAAA...SSLEMLSRE
ITLN1_XENLA  MLSY..SL...L...L...LA...L...PAGHAGSCEQASISEKKEKILNLLACWTEGNADNSLSRS
  
```

ITLN2-F

```

ITLN2-F      .....
ITLN_ONCMY  DLRLRIRSSYIGRSCKEIRDRYNQHEI...GLYYLSTASGTVYQIFCDMTAGGGWTLVASVH
ITL1B_MOUSE .RWGNSFFSSLP...RSCKEIKQEDTKAQI...GLYFLRRTENGVIYQIFCDMTAGGGWTLVASVH
ITLN2_HUMAN FETCAFSSFL...RSCKEIKERCHSAGI...GLYFLRRTKNGVYQIFCDMTAGGGWTLVASVH
ITLN1_XENLA  GGSPGTGMNY...YRSCNEIKSSDSRAP...GTYTLATEDGESYQIFCDMTAGGGWTLVASVH
  
```

Peptídeo Sinal      Ponte de Dissulfeto

Domínio relacionado ao Fibrinogênio (FReD)

ITLN2-F

```

ITLN2-F      .....
ITLN_ONCMY  ENNIHGRKCTYGDRWSTSTQGNSEKYPAGDONWVNLA...TFGSASGATSDDYKNEGYFDI...EADI
ITL1B_MOUSE ENNVYGRKCTMGDRWSSQGSNPNWPDGDNWANRAT...FTAEGATSDDYKNEGYD...I...VAEI
ITLN2_HUMAN ENNLRGRCTV...GDRWSSQGNR...ADY...PEGDGNWANYNT...TFGSAEGATSDDYKNEGYD...I...QAEI
ITLN1_XENLA  ENNMFGRKCTV...GDRWSTQGNM...LQNP...EGDNWANYNT...TFGLPEGATSDDYKNEGYD...I...BAKI
  
```

Sítio de Ligação de Cálcio      Sítio de Ligação de Cálcio      Sítio de Ligação à Carboidrato

ITLN2-F

```

ITLN2-F      .....
ITLN_ONCMY  LISVWHIPNNTPELKKWRLDALTQYHTE...NKFTRYGGNLQGLFKKLP...LMFNIGSCPNNSNGPA
ITL1B_MOUSE LISVWHVPNNSPMEHWNLGA...LLRYHTERS...F...LSIQGNLHQLFKLYPVRYNAEA...SGNTG...FV
ITLN2_HUMAN LGIWHVPNKS...PMQHWRNS...ALLRYR...T...T...G...L...Q...R...L...G...H...N...L...F...G...L...Y...Q...K...Y...P...V...K...Y...R...S...G...K...C...W...N...D...N...G...P...A
ITLN1_XENLA  LALWHVVPNKTPMVMWRNS...S...LLRYR...T...Q...N...G...F...L...T...E...E...G...G...N...L...F...E...L...Y...K...K...Y...P...V...K...Y...D...I...G...K...C...L...A...D...N...G...P...A
  
```

ITLN2-F

```

ITLN2-F      .....
ITLN_ONCMY  IPIVYDFGN...AEK...TINLYSELSKTL...YRRW...FFFRR...
ITL1B_MOUSE IPIVYDFGDKETTRELYGPNTRNQFEPGFITFRPINNELAAMAICSGVKPPTGGDTEHYC
ITLN2_HUMAN FPVVYDFGDAQKTASYSPSGRNEFTAGYVQFRVFNNE...RAASALCAGVRV...TGCNTEHHC
ITLN1_XENLA  IPVVYDFGDAKKTASYSPYQREFVAGFVQFRVFNNE...RAANALCAGIKV...TGCNTEHHC
ITLN1_XENLA  VPVVYDLS...AEK...TAS...L...Y...S...P...N...G...R...S...E...F...T...P...G...F...V...Q...F...R...A...V...N...S...E...R...A...T...L...A...L...C...A...G...V...K...V...K...G...C...N...V...E...H...H...C
  
```

ITLN2-F

```

ITLN2-F      .....
ITLN_ONCMY  IGGGGHFP...EGAP...RQCGDF...PA...FDWN...GYGT...NTEWS...ASK...QLTEAAVLLFYR
ITL1B_MOUSE IGGGGFFPEFDFPEECGDF...A...FDANGYGT...HIRYSNSREITEAAVLLFYR
ITLN2_HUMAN IGGGGFFPQGKFRQCGDFS...AFDWD...GYGTHVKS...SCSREITEAAVLLFYR
ITLN1_XENLA  IGGGGYTFEGSPRQCGDFA...ALDWD...GYGT...NLGWS...ASK...QLTEAAVMLFYR
  
```

**APÊNDICE 3 – *Lp*ITLN2 E O LIGANTE POLI (I:C)**