



---

**Universidade de São Paulo**  
**Instituto de Química**  
**Programa de Pós-Graduação em Bioquímica**

---

Clonagem, expressão e purificação de um domínio hipotético de ligação a ácidos graxos em *Leishmania amazonensis* para ensaios de biologia estrutural

---

Aluno: Lucas Felipe Almeida Athayde

Versão corrigida da dissertação

São Paulo, 2023

Data de depósito na CPG: 24/04/2023

# **Lucas Felipe Almeida Athayde**

Clonagem, expressão e purificação de um domínio hipotético de ligação a ácidos graxos em *Leishmania amazonensis* para ensaios de biologia estrutural

**Dissertação apresentada ao Instituto de Química da Universidade de São Paulo para obtenção do título de Mestre em Ciências (Bioquímica).**

**Orientador: Prof. Dr. Roberto Kopke Salinas**

**Co-orientadora: Profa. Dra. Maria Fernanda Laranjeira da Silva**

**São Paulo 2023**

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha Catalográfica elaborada eletronicamente pelo autor, utilizando o programa desenvolvido pela Seção Técnica de Informática do ICMC/USP e adaptado para a Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP

Bibliotecária responsável pela orientação de catalogação da publicação:  
Marlene Aparecida Vieira - CRB - 8/5562

A865c	Athayde, Lucas Clonagem, expressão e purificação de um domínio hipotético de ligação a ácidos graxos em Leishmania amazonensis para ensaios de biologia estrutural / Lucas Athayde. - São Paulo, 2023. 129 p.  Dissertação (mestrado) - Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica. Orientador: Salinas, Roberto Coorientador: Laranjeira Silva, Maria  1. Clonagem e expressão heteróloga. . 2. Purificação de proteínas recombinantes. 3. Biologia Estrutural. 4. Leishmania. 5. Fatty acids binding protein. I. T. II. Salinas, Roberto, orientador. III. Laranjeira Silva, Maria, coorientador.
-------	--



Universidade de São Paulo  
**Instituto de Química**

"Clonagem, expressão e purificação de um domínio hipotético de ligação a ácidos graxos em *Leishmania amazonensis* para ensaios de biologia estrutural"

## **LUCAS FELIPE ALMEIDA ATHAYDE**

Dissertação de Mestrado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Mestre em Ciências obtido no Programa Ciências Biológicas (Bioquímica) - Área de Concentração: Bioquímica.

---

**Prof. Dr. Roberto Kopke Salinas**  
(Orientador e Presidente)

**APROVADO(A) POR:**

---

**Prof. Dr. Marcelo Santos da Silva**  
IQ - USP

---

**Profa. Dra. Cristiane Rodrigues Guzzo Carvalho**  
ICB - USP

---

**Profa. Dra. Leda Quercia Vieira**  
UFMG

**SÃO PAULO**  
29 de maio de 2023

## **Dedicatória**

Para Túlio (*in memoriam*) que sonhou este dia comigo. Dedico-lhe este trabalho com amor pelos 11 anos de amizade.

## **Agradecimentos**

Gostaria de agradecer à minha mãe Adriana e à minha irmã Nayara por sempre terem sido minhas maiores figuras de inspiração. Para elas que eu olhava quando era pequeno e desejava ser igual.

Ao Thiago, irmão que a vida me deu, pelos 20 anos de amizade e por ser a família que eu escolhi. Sua escuta compreensiva, suas visitas periódicas a São Paulo e as mensagens aleatórias de carinho no dia-a-dia sempre me fortaleceram.

Aos meus orientadores Professor Roberto Kopke Salinas e Professora Maria Fernanda Laranjeira Silva, que com toda sua expertise, acreditaram em mim e me permitiram conduzir este projeto.

Aos meus amigos do Laboratório de Ressonância Magnética de Proteínas e do Laboratório de Estrutura e Função de Proteínas do IQ-USP: Angy Davalos, Jose Rivera, Made Castro, Jose Edwin, Edgar Llontop, Luís Pizzolato, Daniela Zapata Sifuentes, Santiago Justo Arévalo e Michella Brescia por terem colaborado com este trabalho e principalmente por serem amizades valiosas que fiz aqui.

Ao FeLab do IB-USP: Romário Lopes, Gi Aquino, Marco Gomes, Avery Hong e Ricardo Zampieri por terem sido os primeiros a me acolherem na nova cidade e me ensinarem sobre Biologia Molecular.

Às técnicas Camila Santos Schroeder, Ilda Souza Costa e Camille Ortiz também pela amizade, afago e cafezinhos prazerosos entre uma purificação e outra. E ao secretário de pós-graduação Milton César, sempre solícito.

Aos meus amigos do IQ-USP: Allan Pradelli, Noemi Jacó, Artur Oliveira, Alejandro Herrera, Bianca Scigliano e Doug Ricardo por me incentivarem constantemente.

À CAPES pelo apoio financeiro.

Muito obrigado a todos.

## Resumo

ATHAYDE, L.F.A. Clonagem, expressão e purificação de um domínio hipotético de ligação a ácidos graxos em *Leishmania amazonensis* para ensaios de biologia estrutural.

Número de páginas: 129. Dissertação de Mestrado - Programa de Pós-Graduação em Ciências Biológicas (Bioquímica), Instituto de Química, Universidade de São Paulo, São Paulo.

As leishmanioses são uma classe de doenças tropicais negligenciadas causadas por diferentes espécies do gênero *Leishmania*, da família Trypanosomatidae. Durante o seu ciclo de vida, os parasitas desse gênero multiplicam-se em hospedeiros vertebrados e não-vertebrados, diferenciando-se formas celulares distintas: promastigotas e amastigotas. Esses parasitas capturam ácidos graxos do hospedeiro e os utilizam como fonte de energia ou como esqueletos para a construção de moléculas mais complexas. Dado a importância do metabolismo de ácidos graxos para a sobrevivência e multiplicação dos parasitas no hospedeiro, procurou-se no banco de dados genômicos de tripanosomatídeos, TriTrypDB, proteínas hipotéticas potencialmente semelhantes a proteínas de ligação de ácidos graxos eucarióticas (FABP). A proteína hipotética codificada pelo gene *LmxM.34.3070* foi identificada pela primeira vez em *Leishmania mexicana*. Buscas no banco de dados TriTrypDB demonstraram que o gene *LmxM.34.3070* é conservado em outros tripanosomatídeos. Diferentes métodos de predição foram utilizados para prever a estrutura da proteína *LmxM.34.3070*, revelando a presença de um domínio C-terminal estruturalmente semelhante a proteínas de ligação de ácidos graxos eucarióticas (FABP). Essas predições sugerem que a proteína completa, *LmxM.34.3070*, poderia estar associada ao metabolismo de lipídios e ácidos graxos, essencial para a sobrevivência do parasito no ambiente do hospedeiro. O objetivo principal deste trabalho foi obter amostras do domínio C-terminal de *LmxM.34.3070* adequadas a estudos de biologia estrutural. Para isso, quatro construções foram desenhadas, clonadas e testadas quanto à expressão heteróloga, solubilidade, rendimento e estabilidade. A construção mais solúvel, *LmxM.34.3070*<sub>638-822</sub> ou LeiFABP, foi caracterizada por espectroscopia de dicroísmo circular (CD), espalhamento de luz a múltiplos ângulos (MALS) e espectroscopia de ressonância magnética nuclear (RMN). As medidas de MALS mostraram que o domínio LeiFABP é monomérico. A análise quantitativa do espectro de CD indicou que 32% dos aminoácidos de LeiFABP assumem estrutura do tipo  $\beta$ , enquanto 15% assumem estrutura de  $\alpha$ -hélice. O resultado desta análise é consistente com as predições em Colab Fold. O espectro de RMN de  $^1\text{H}$  apresentou boa dispersão na região dos HNs e linhas relativamente estreitas, sugerindo que a construção LeiFABP é adequada para a análise estrutural por RMN em solução ou cristalografia.

**Palavras-chave:** Biologia Estrutural, Espectroscopia de RMN, tripanosomatídeos, FABP.

## Abstract

ATHAYDE, L.F.A. Cloning, expression, and purification of a hypothetical fatty acid binding domain from *Leishmania amazonensis* for structural biology assays.

Number of pages: 129. M.Sc Dissertation - Graduate Program in Biological Sciences (Biochemistry), Institute of Chemistry, University of São Paulo, São Paulo.

Leishmaniasis is a class of neglected tropical diseases caused by different species of the genus *Leishmania* of the Trypanosomatidae family. During their life cycle, *Leishmania* parasites survive in vertebrate and invertebrate hosts, and differentiate into distinct cellular forms: amastigotes and promastigotes. These parasites scavenge host fatty acids to use as energy sources or as building blocks for synthesis of more complex molecules. Given the importance of lipids and fatty acids to the parasites' survival and multiplication, it was hypothesized that fatty acid binding proteins (FABP) could play a role in *Leishmania* lipids and fatty acids metabolism. The hypothetical protein coded by the *LmxM.34.3070* gene was first identified in *Leishmania mexicana*. Searches in the TriTrypDB database showed that *LmxM.34.3070* was conserved in other trypanosomatids. Predictions of the *LmxM.34.3070* protein structure by different methods highlighted the presence of a C-terminal domain structurally similar to eukaryotic fatty acid binding proteins (FABP). These predictions suggested that the full-length *LmxM.34.3070* protein could be associated with the metabolism of lipids and fatty acids, which is essential for *Leishmania* survival in the host. The main objective of this work was to obtain samples from the C-terminal domain of *LmxM.34.3070* suitable for structural biology studies. A total of four constructs from the *LmxM.34.3070* C-terminal domain were designed, cloned and tested regarding their heterologous expression level, solubility, and stability. The most soluble construct, *LmxM.34.3070*<sub>638-822</sub> or LeiFABP, was characterized by Circular Dichroism (CD) spectroscopy, Multi-Angle Light Scattering (MALS) and Nuclear Magnetic Resonance (NMR) spectroscopy. MALS experiments showed that this construct was monomeric under the buffer conditions tested. Quantitative analysis of the CD spectra indicated that 32% of LeiFABP amino acids were in  $\beta$ -type structure, while 15 % were in  $\alpha$ -helix, in agreement with the ColabFold prediction. The <sup>1</sup>H NMR spectrum of LeiFABP showed good dispersion and sharp signals at the HN region, suggesting that it is a suitable construct for structural analysis by solution NMR spectroscopy or crystallography.

**Key words:** Structural Biology, NMR spectroscopy, trypanosomatids, FABP.



## Lista de abreviaturas e siglas

Anti-His	anticorpo monoclonal para <i>tag</i> de histidinas
BCIP	5-bromo-4-cloro-3-indolil fosfato
BSA	albumina de soro bovino
CD	dicroísmo circular
DNA	ácido desoxirribonucleico
dNTP	desoxirribonucleotídeos trifosfato
EDTA	ácido etileno diamino tetracético
FABP	proteína de ligação a ácidos graxos
DNA <sub>g</sub>	DNA genômico
HCl	ácido clorídrico
IPTG	isopropil β-D-1-tiogalactosídeo
LB	Luria-Bertani
MALS	<i>Multiangle light scattering</i>
mRNA	RNA mensageiro
NaPi	tampão fosfato de sódio (monobásico/dibásico)
NBT	<i>Nitroblue tetrazolium</i>
PBS	tampão fosfato salino
PCR	reação em cadeia da polimerase
PMSF	fluoreto de fenil metil sulfonila
RMN	ressonância magnética nuclear
RNA	ácido ribonucleico
RNase	ribonuclease
SDS	dodecil sulfato de sódio
SDS-PAGE	eletroforese em gel de poliacrilamida contendo SDS
SEC	<i>size exclusion chromatography</i>
TAE	tampão Tris-acetato-EDTA
TBS	tampão Tris-salino
T <sub>m</sub>	Temperatura de <i>melting</i>
Tris	tris (hidroximetil) aminometano
tRNA	RNA transportador
U	Unidade
UV	ultravioleta
V	volume
VF	volume final
V/V	relação volume para volume
X-GAL	5-bromo-4-cloro-3-indolil-D-galactopiranosídeo

## Lista de aminoácidos

A	alanina
C	cisteína
D	aspartato
E	glutamato
F	fenilalanina
G	glicina
H	histidina
I	isoleucina
K	lisina
L	leucina
M	metionina
N	asparagina
P	prolina
Q	glutamina
R	arginina
S	serina
T	treonina
V	valina
Y	tirosina
W	triptofano

## Unidades de medidas

°C	grau Celsius
g	grama
g/L	grama por litro
h/hrs	hora/ horas
kb	kilobases
kDa	kilodaltons
L	litro
m	metro
M	molar
mA	miliamper
mDeg	miligras
mF	milifarad
min	minuto
ms	milissegundos
pb	par de bases
ppm	partes por milhão
rpm	rotações por minuto
V	volt
x g	unidade de aceleração normal da gravidade

# Sumário

<b>1.Introdução</b> .....	20
1.1 Leishmanioses .....	21
1.2 Aspectos clínicos e epidemiológicos das leishmanioses .....	22
1.3 Mecanismos de captura, síntese e armazenamento de ácidos graxos e lipídeos em <i>Leishmania</i> spp. ....	24
1.4. Proteínas de ligação a ácidos graxos (FABPs).....	28
1.5 <i>LmxM.34.3070</i> é conservado no gênero <i>Leishmania</i> .....	29
1.6. Modelagem preliminar da proteína <i>LmxM.34.3070</i> .....	31
<b>2. Objetivos</b> .....	34
<b>3. Materiais e métodos</b> .....	36
3.1. Predição de estrutura secundária e terciária .....	37
3.2. Células: .....	37
3.3. Meios de cultura: .....	38
3.4. Extração de DNA genômico de <i>L. amazonensis</i> .....	38
3.5. Reação em cadeia da polimerase (PCR).....	38
3.6. Tecnologia de DNA recombinante .....	41
3.7. Subclonagem de fragmentos codificantes obtidos por PCR .....	42
3.7.1. Subclonagem em vetor pCR2.1 com o kit TOPO-TA®: .....	42
3.7.2. Subclonagem em pGEM®-T Easy Vector Systems: .....	42
3.8. Transformação de bactérias competentes.....	43
3.9. Blue-White colony screening .....	44
3.10. Mini-preparação de plasmídeo bacteriano (MINIPREP) .....	44
3.11. Clonagem por restrição em vetor de expressão pET24a e pET28a.....	45
3.11.1. Digestões dupla.....	45
3.11.2. Purificação de fragmentos de DNA de gel de agarose .....	45
3.11.3. Reação de ligação com a T4 Ligase.....	46
3.12. Clonagem direta por Gibson em vetor de expressão pET28a.....	46
3.13. PCR de colônia .....	47
3.14. Sequenciamento de clones (confirmados por PCR de colônia) .....	48
pelo método de Sanger .....	48
3.15. Teste de expressão heteróloga dos diferentes fragmentos recombinantes ....	48
3.16. Eletroforese unidimensional em gel de poliacrilamida (SDS-PAGE).....	50

3.17. Western blot Anti-His.....	53
3.18. Cromatografia de Afinidade.....	53
3.19. Cromatografia por exclusão molecular.....	58
3.20. Método do coeficiente de extinção molar.....	61
3.21. Espalhamento de luz a múltiplos ângulos acoplado à cromatografia por exclusão molecular (SEC-MALS).....	62
3.22. Espectroscopia de dicroísmo circular (CD).....	63
3.23. Desnaturação térmica.....	64
3.24. Espectroscopia de Ressonância Magnética Nuclear.....	65
<b>4. Resultados e Discussão.....</b>	<b>66</b>
4.1. Os modelos preditos por Raptor-X e ColabFold são semelhantes aos modelos gerados por PHYRE2.....	67
4.2. Conteúdo de estrutura secundária do domínio FABP-like foi estimado por PSIPRED.....	74
4.3. Amplificação por PCR dos fragmentos de interesse.....	78
4.4. Confirmação da subclonagem por ensaios de dupla de digestão: pCR2.1-TOPO-LmxM.34.3070 <sub>652-822</sub> e pGEM-LmxM.34.3070 <sub>669-822</sub> .....	79
4.5. Confirmação da clonagem por PCR de colônia.....	81
4.6. Sequenciamento dos clones de expressão.....	82
4.7. LmxM.34.3070 <sub>652-822</sub> é mais expresso quando fusionado a uma <i>his-tag</i> N-terminal.....	90
4.8. Expressão em larga escala e purificação de LmxM.34.3070 <sub>652-822</sub> .....	94
4.9. A temperatura de indução e a concentração de IPTG influenciaram na solubilidade de LmxM.34.3070 <sub>652-822</sub> recombinante.....	95
4.10. Teste de aditivos no tampão de lise e otimização da cromatografia de afinidade.....	97
4.11. As hélices N-terminais adicionais preditas por PSIPRED e ColabFold são essenciais para expressão do domínio C-terminal de LmxM34.3070 solúvel.....	102
4.12. LeiFABP tende a formar oligômeros.....	107
4.13. Obtendo amostras monoméricas de LeiFABP.....	111
4.14. LeiFABP possui estruturas do tipo Beta.....	114
4.15. Influência do tampão na estabilidade térmica de LeiFABP.....	115
4.16. Espectro de RMN Unidimensional de LeiFABP.....	118
5.1. CONCLUSÕES.....	122
5.2. PERSPECTIVAS.....	123
<b>6. Referências.....</b>	<b>124</b>

## Lista de figuras

**Figura 1:** Ciclo de vida de *Leishmania*.

**Figura 2:** Representação esquemática do metabolismo de ácidos graxos e lipídios em *Leishmania* spp.

**Figura 3:** Microscopia eletrônica de transmissão de amastigota de *Leishmania amazonensis*.

**Figura 4:** Predição de domínios da proteína LmxM.34.3070 por PHYRE2.

**Figura 5:** Mecanismo de captura em Hitrap.

**Figura 6:** Mecanismo de separação por exclusão molecular.

**Figura 7:** Esquema geral do protocolo de purificação dos diferentes fragmentos recombinantes da região C-terminal de LmxM.34.3070.

**Figura 8:** Sobreposição par a par das estruturas geradas para a região 52-488 correspondente ao domínio N-terminal de LmxM.34.3070 por PHYRE2, Raptor-X e ColabFold.

**Figura 9:** Sobreposição par a par das estruturas geradas para a região C-terminal (aminoácidos 665-809) de LmxM.34.3070 por PHYRE2, Raptor-X e ColabFold.

**Figura 10:** Predição de estrutura da proteína LmxM.34.3070 em Raptor-X.

**Figura 11:** Predição de estrutura da proteína LmxM.34.3070 em ColabFold.

**Figura 12:** Resultados da busca comparativa no servidor DALI entre a estrutura predita do domínio LeiFABP por ColabFold.

**Figura 13:** **Acima:** Alinhamento entre a sequência de amino ácidos do domínio LeiFABP e de As-p18 (6I9F). **Abaixo:** Alinhamento da sequência de aminoácidos de *LmxM.34.3070* e potenciais ortólogos em outros tripanosomatídeos.

**Figura 14:** Representação gráfica em WebLogo do alinhamento múltiplo obtido entre a região C-terminal de LmxM.34.3070 em diferentes espécies de *Leishmania*.

**Figura 15:** A) Conteúdo de estrutura secundária predita para a região C-terminal (aminoácidos 665-809) de LmxM.34.3070 por PSIPRED. B) Topologia clássica das FABPs.

**Figura 16:** Conteúdo de estrutura secundária predita para o fragmento 638-822 de LmxM.34.3070 por PSIPRED.

**Figura 17:** Modelos ColabFold para os fragmentos C-terminais de LmxM.34.3070 testados. A) Fragmento LmxM.34.3070<sub>638-822</sub>. B) Fragmento LmxM.34.3070<sub>652-822</sub>. C) Fragmento LmxM.34.3070<sub>669-822</sub>.

**Figura 18:** Figura 16: A) Gradiente de temperatura (58, 62 e 65°C) para amplificação do fragmento de DNA codificante de LmxM.34.3070<sub>652-822</sub> com os sítios de restrição de *Bam*HI

e *HindIII*. B) Gradiente de temperatura (58, 60, 62 e 64°C) para amplificação do fragmento de DNA codificante LmxM.34.3070<sub>652-822</sub> com os sítios de restrição de *NdeI* e *BamHI*.

**Figura 19:** A) Gradiente de temperatura (58, 60, 62 e 64°C) para amplificação do fragmento de DNA codificante de LmxM.34.3070<sub>638-822</sub> com regiões de sobreposição para clonagem pela metodologia de Gibson. B) Gradiente de temperatura (58, 60, 62 e 64°C) para amplificação do fragmento de DNA codificante LmxM.34.3070<sub>669-822</sub> com os sítios de restrição de *NdeI* e *HindIII*.

**Figura 20:** A) Eletroforese em gel dos produtos de digestão de LmxM.34.3070<sub>652-822</sub> subclonado em pCR2.1. B) Eletroforese em gel dos produtos de digestão de LmxM.34.3070<sub>669-822</sub> subclonado em pGEM.

**Figura 21:** PCRs de colônias dos clones construídos para expressão de diferentes fragmentos da região C-terminal de LmxM.34.3070.

**Figura 22:** Alinhamento por BLAST entre as sequências experimentais e as sequências de aminoácidos do fragmento LmxM34.3070<sub>652-822</sub>(pET24a)

**Figura 23:** Análise de substituição no Cromatograma de sequenciamento do clone LmxM34.3070<sub>652-822</sub>(pET24a).

**Figura 24:** Análise de substituição no Cromatograma de sequenciamento do clone LmxM34.3070<sub>652-822</sub>(pET24a).

**Figura 25:** Alinhamento por BLAST entre as sequências experimental e esperada do clone LmxM34.3070<sub>652-822</sub>(pET28a).

**Figura 26:** Alinhamento por BLAST entre as sequências experimental e esperada dos clones A) LmxM34.3070<sub>652-822</sub>(pET28a) B) LmxM34.3070<sub>669-822</sub>(pET28a).

**Figura 27:** A) Perfil eletroforético de expressão do clone LmxM.34.3070<sub>652-822</sub>(pET24a), com uma *tag* de histidinas fusionada na região C-terminal. B) Detecção da expressão do fragmento LmxM.34.3070<sub>652-822</sub> por western blot em culturas induzidas de diferentes cepas de BL21.

**Figura 28:** A) Perfis eletroforéticos de expressão do clone LmxM.34.3070<sub>652-822</sub>(pET28a), com uma *tag* de histidinas fusionada na região N-terminal (~21 kDa) em BL21 e RIP-L avaliados por SDS-PAGE. B) Detecção da expressão por western blot em amostras de cultura induzida de BL21 transformadas com o clone LmxM.34.3070<sub>652-822</sub>(pET28a).

**Figura 29:** A) Bandas de expressão do clone LmxM.34.3070<sub>669-822</sub>(pET28a) com uma *tag* de histidinas na região N-terminal (~19 kDa) em RIL e C43 detectadas por western blot. B)

Bandas de expressão do fragmento LmxM.34.3070<sub>669-822</sub> com uma *tag* de histidinas fusionada na região N-terminal (~19 kDa) em BL21 e RIP-L detectadas por western blot.

**Figura 30:** Perfil eletroforético de expressão do clone LmxM.34.3070<sub>638-822</sub>(pET28a) com uma *tag* de histidinas fusionada na região N-terminal (~22 kDa) em RIP-L avaliado por SDS-PAGE.

**Figura 31:** Perfil eletroforético do primeiro teste de lise de células RIP-L transformadas com LmxM.34.3070<sub>652-822</sub> (pET28a).

**Figura 32:** À esquerda: Perfil eletroforético da primeira tentativa (Teste 2 Tabela 4.1) de purificação do fragmento LmxM.34.3070<sub>652-822</sub> (~21 kDa). À direita: Cromatograma de purificação do fragmento LmxM.34.3070<sub>652-822</sub> através de cromatografia de afinidade em coluna HiTrap carregada com Ni<sup>2+</sup> no teste 2.

**Figura 33:** A: Perfil eletroforético (à esquerda) e cromatograma da purificação (à direita) do fragmento LmxM.34.3070<sub>652-822</sub> recombinante (~21 kDa) através de cromatografia de afinidade no Teste 3 (tabela 4.1) em coluna HiTrap carregada com Ni<sup>2+</sup>. B) Perfil eletroforético e cromatograma da etapa de purificação do fragmento LmxM.34.3070<sub>652-822</sub> recombinante (~21 kDa) por exclusão molecular (experimento GF1, tabela 5) em coluna Superdex S75 16/600.

**Figura 34:** Acima: Perfil eletroforético e cromatograma da quarta tentativa de purificação do fragmento LmxM.34.3070<sub>652-822</sub> recombinante (~21 kDa) através de cromatografia de afinidade (Teste 4, tabela 4.1) em coluna HiTrap carregada com Ni<sup>2+</sup>. Abaixo: INS= amostra da fração insolúvel avaliada por SDS-PAGE.

**Figura 35:** Perfil eletroforético (à esquerda) e cromatograma (à direita) da etapa de purificação do fragmento LmxM.34.3070<sub>652-822</sub> recombinante (~21 kDa) através de cromatografia de exclusão molecular (experimento GF2, tabela 5) na coluna Superdex S75 16/600.

**Figura 36:** Modelo predito para a região C-terminal de LmxM.34.3070 demonstrando em amarelo a presença do motivo hélice-alça-hélice e fitas beta antiparalelas que formam um barril (topologia FABP-like) e todas as hélices N-terminais adicionais previstas por ColabFold e PSIPRED.

**Figura 37:** Cromatograma (à esquerda) e perfil eletroforético (à direita) da purificação do fragmento LmxM.34.3070<sub>638-822</sub> recombinante (~22 kDa) através de cromatografia de afinidade em coluna HiTrap carregada com Ni<sup>2+</sup>.



**Figura 38:** SDS-PAGE do teste de solubilidade de células RIP-L e C43 transformadas com LmxM.34.3070<sub>669-822</sub>(pET28a) (~19 kDa).

**Figura 39:** Cromatogramas das três tentativas de purificação do fragmento LmxM.34.3070<sub>669-822</sub> recombinante (~19 kDa) através de cromatografia de afinidade em coluna HiTrap carregada com Ni<sup>2+</sup>. Estes experimentos correspondem aos Testes 6, 7 e 8, (Tabela 4.2).

**Figura 40:** Predição de desordem por fIDPnn para toda a sequência de resíduos de aminoácidos do fragmento recombinante expressa pelo clone LmxM34.3070<sub>638-822</sub>(pET28a).

**Figura 41:** Cromatograma de purificação de LeiFABP por exclusão molecular (à esquerda) em coluna Superdex S75 16/600 no experimento GF3 (tabela 5).

**Figura 42:** Análise do estado de agregação das frações 16,19 e 23 por SEC-MALS.

**Figura 43:** Cromatograma e perfil eletroforético da etapa de purificação do domínio LeiFABP por exclusão molecular em coluna Superdex S75 16/600 em GF4 (tabela 5).

**Figura 44:** Medidas de SEC-MALS da fração 23 que eluiu da coluna Superdex S75 no experimento GF4 (tabela 5).

**Figura 45:** Deconvolução do espectro de CD de LeiFABP a 25 °C através do DichroWeb.

**Figura 46:** Sobreposição e curvas de desnaturação de LeiFABP em diferentes pHs e composições de tampão.

**Figura 47:** Espectro de Ressonância Magnética Nuclear 1D (1H) de LeiFABP.

## Lista de tabelas

**Tabela 1:** Potenciais ortólogos de *LmxM.34.3070* no gênero *Leishmania*.

**Tabela 2:** Características físico-químicas dos oligonucleotídeos.

**Tabela 3:** Genótipos das cepas utilizadas nos testes de expressão dos diferentes clones obtidos.

**Tabela 4.1:** Testes de condições de expressão, lise e cromatografia de afinidade.

**Tabela 4.2:** Testes de condições de expressão, lise e cromatografia de afinidade.

**Tabelas 4.3:** Testes de condições de expressão, lise e cromatografia de afinidade.

**Tabela 5:** Testes de condições de purificação de diferentes construções por exclusão molecular.

**Tabela 6:** Características dos fragmentos testados quanto ao número de resíduos de aminoácidos, posição da *tag* de histidinas, vetor de expressão utilizado, massa molecular, número de hélices N-terminais adicionais e método de clonagem.

**Tabela 7:** Resultados dos testes de expressão dos diferentes fragmentos das região C-terminal de *LmxM.34.3070*.

**Tabela 8:** Resumo dos resultados envolvendo testes de solubilidade, perfil cromatográfico da exclusão molecular, solubilidade e rendimento final com os três fragmentos construídos para estudos estruturais de LeiFABP: *LmxM.34.3070*<sub>652-822</sub>, *LmxM.34.3070*<sub>638-822</sub> e *LmxM.34.3070*<sub>669-822</sub>.

## **Súmula Curricular**

### **DADOS PESSOAIS**

Nome: Lucas Felipe Almeida Athayde

Data de nascimento: 12/02/1996

Local de nascimento: Belo Horizonte, Minas Gerais, Brasil.

### **EDUCAÇÃO**

- Ensino Técnico Integrado em Equipamentos Biomédicos (2011-2014). Centro Federal de Educação Tecnológicas de Minas Gerais (CEFET-MG).
- Graduação em Ciências Biológicas (Licenciatura) (2016-2021). Universidade Federal de Minas Gerais (UFMG).

### **FORMAÇÃO COMPLEMENTAR**

- 2023: Aceito para participar da São Paulo School of Advanced Science in CryoEM (FAPESP)
- 2022: Capacitação no uso e manejo de animais de laboratório (ICB/USP)
- 2022: Fundamentos e aplicações da CRISPR- Cas9 (Biotech Brasil)
- 2021: Advanced Mass Spectrometry (Instituto de Química da USP)
- 2021: Difractometria de Raios-X (Instituto de Química USP)
- 2021: Técnicas de eletroforese (Instituto de Química USP)
- 2021: PCR em tempo real no diagnóstico de doenças (Biotech Brasil)
- 2020: Expressão heteróloga e produção de proteínas recombinantes (UFV)

### **OCUPAÇÃO ATUAL**

Bolsista de Mestrado CAPES PROEX: 88887.613399/2021-00. IQ-USP (2021-2023)

### **AUXÍLIOS ANTERIORES**

- Bolsa do CNPq na modalidade IC, para o período de 01/09/2020 a 31/03/2021 ICB/UFMG;
- Bolsa PIC/IRR na modalidade IC, período 01/03/2019 a 30/04/2020.
- Bolsa INOVA FIOCRUZ na modalidade IC, período 01/05/2020 a 31/07/2020

### **PUBLICAÇÕES**

Hickson, J., Athayde, L.F.A., Miranda, T.G. et al. (2022) "*Trypanosoma cruzi* iron superoxide dismutases: insights from phylogenetics to chemotherapeutic target assessment". Parasites Vectors. 15: 194. (<http://dx.doi.org/10.1186/s13071-022-05319-2>)

### **PARTICIPAÇÃO EM EVENTOS**

II Simpósio de Pós-Graduação em Bioquímica do Instituto de Química da USP (2023)

52º Congresso da Sociedade Brasileira de Bioquímica de Biologia Molecular (2023)

---

# **1.Introdução**

---

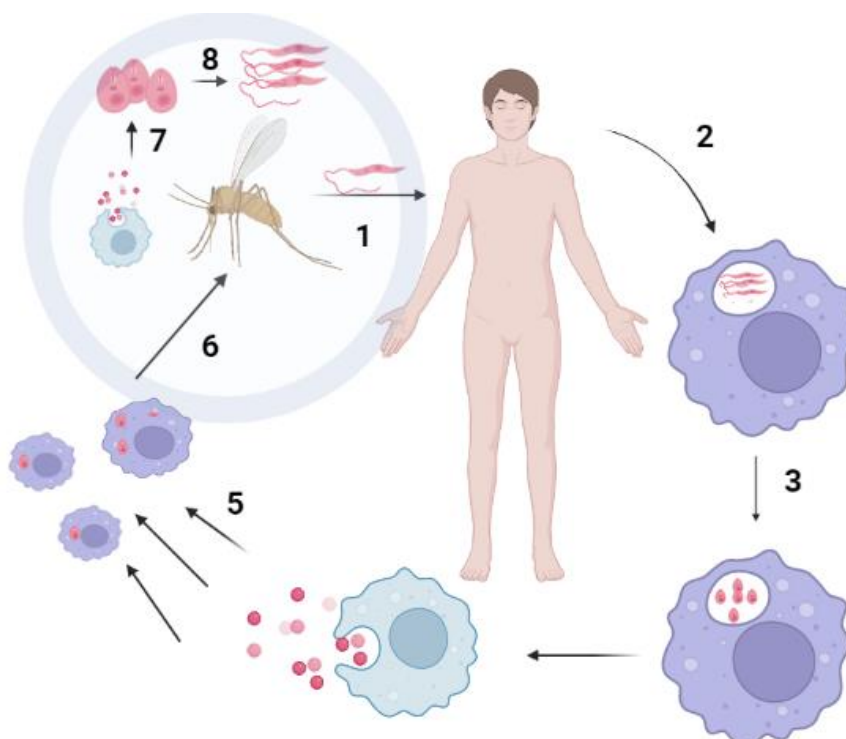
## 1.1 Leishmanioses

Leishmanioses são doenças causadas por protozoários do gênero *Leishmania*, da ordem Kinetoplastida, família Tripanosomatidae (Handman, 1999) (Ponte-Sucre et al., 2017). Os protozoários deste gênero possuem um ciclo de vida heteroxeno, com um estágio em hospedeiros invertebrados, sendo estes as fêmeas de insetos flebotomíneos (família Psychodidae), e um estágio em hospedeiros vertebrados, dentre eles os seres humanos (Handman, 1999) (Kaye & Scott, 2011).

Os parasitos do gênero *Leishmania* diferenciam-se em formas celulares morfológica e metabolicamente distintas durante o seu ciclo de vida. As formas promastigotas são fusiformes e apresentam alta motilidade devido à presença de um flagelo terminal (Waller & McConville, 2002), podendo ser classificadas como promastigotas procíclicas ou metacíclicas (Alcolea et al., 2019). No inseto, as formas promastigotas procíclicas replicam-se por divisão binária e por um processo denominado metaciclogênese, diferenciam-se em promastigotas metacíclicas não-replicativas, infectivas para o hospedeiro vertebrado (Cunningham, 2002). Após o repasto sanguíneo do inseto, as formas promastigotas metacíclicas são regurgitadas no hospedeiro vertebrado, onde ocorre a infecção de macrófagos e outras células pelo parasito (Waller & McConville, 2002) (Figura 1).

Ao se estabelecerem em compartimentos denominados vacúolos parasitóforos no interior dos macrófagos, as formas promastigotas se diferenciam nas formas replicativas, conhecidas como amastigotas, de forma arredondada e sem flagelo aparente (Cunningham, 2002) (Figura 1). As amastigotas exibem uma notável capacidade de sobrevivência no interior de macrófagos, e conseguem evadir-se às agressões do sistema imune do hospedeiro, como acidificação do microambiente, estresse oxidativo, enzimas proteolíticas no fagolisossomo, superando ainda oscilações de temperatura e privação

denutrientes (Cunningham, 2002). A replicação de amastigotas leva ao rompimento da célula hospedeira infectada, (Kaye & Scott, 2011) (Figura 1). Macrófagos infectados do hospedeiro são captados por fêmeas de flebotomíneos durante o repasto sanguíneo no hospedeiro mamífero infectado. Uma vez no intestino médio do inseto, as amastigotas são liberadas dos macrófagos, e se diferenciam em promastigotas procíclicas, reiniciando assim o ciclo das leishmanioses (Bogitsh, 2005).



**Figura 1:** Ciclo de vida simplificado de *Leishmania*. Na fêmea de flebotomíneo, as formas amastigotas diferenciam-se em promastigotas infectivas para o hospedeiro vertebrado, conhecidas como promastigotas metacíclicas (etapa 8). As promastigotas metacíclicas são fagocitadas (etapa 2) e, uma vez no interior do macrófago, se diferenciam em amastigotas intracelulares replicativas (etapa 3). A proliferação das amastigotas intracelulares leva à lise celular e à liberação dos parasitos (etapa 4). As amastigotas, infectam outras células do hospedeiro (etapa 5) que são captadas pelas fêmeas de flebotomíneos durante o repasto sanguíneo (etapa 6), reiniciando o ciclo de *Leishmania* (etapas 1 a 8).

## 1.2 Aspectos clínicos e epidemiológicos das leishmanioses

As manifestações clínicas das leishmanioses podem ser divididas em leishmaniose tegumentar e leishmaniose visceral (Machado et al., 2019). A forma tegumentar é caracterizada pelo aparecimento de diversas lesões nodulares e granulomas na pele e na mucosa (Machado et al., 2019). Já a forma visceral é potencialmente letal,

sendo causada principalmente pelas espécies *L. infantum* e *L. donovani* (Ghorbani & Farhoudi, 2018). No entanto, Murray e colaboradores descreveram, em 2005, que, na América do Sul, a forma visceral da doença também é ocasionalmente associada à *L. amazonensis* (Murray et al., 2005). Em sua forma visceral, a doença pode ser assintomática, subaguda, aguda ou crônica, sendo os principais sintomas febre, hepatoesplenomegalia, pancitopenia e hipergamaglobulinemia (Desjeux, 1996).

O tratamento das leishmanioses baseia-se em três classes de medicamentos: Antimoniais, Anfotericina B e Miltefosina. O Antimonial Pentavalente é uma pró-droga, cuja forma pentavalente é processada nos macrófagos na forma trivalente altamente citotóxica (Tessarollo et al., 2015). O mecanismo de ação da Anfotericina B consiste na sua ligação ao ergosterol de membrana do parasito, formando poros que afetam o fluxo de íons e homeostase (Lemke et al., 2005) (Soares-Bezerra et al., 2004). Já a Miltefosina está associada ao bloqueio da cadeia respiratória através da inibição da enzima citocromo c oxidase (complexo IV), e de vias de sinalização por inibição da proteína quinase B (PKB). Estes tratamentos apresentam limitadas taxas de cura e alta toxicidade para o paciente (Ghorbani & Farhoudi, 2018) (Ponte-Sucre et al., 2017). Somado a esses fatores, também pode-se mencionar a existência de cepas resistentes (Ghorbani & Farhoudi, 2018).

Mais de 1 bilhão de pessoas vivem atualmente em zonas de risco e 30 mil novos casos de leishmaniose visceral são registrados anualmente nas zonas endêmicas (WHO, 2021). No caso da leishmaniose tegumentar, esses números são ainda maiores, ocorrendo a notificação de mais de 1 milhão de casos por ano (WHO, 2021). Fatores ambientais (aumento do desmatamento), fatores populacionais (ampliação dos processos migratórios) e a seleção de cepas resistentes contribuem para a emergência e reemergência das leishmanioses (Hong et al., 2020). Portanto, as leishmanioses ainda

representam um grande desafio do ponto de vista de saúde pública, tornando-se necessário ampliar o entendimento sobre o parasito para o desenvolvimento de novas abordagens terapêuticas mais efetivas.

### **1.3 Mecanismos de captura, síntese e armazenamento de ácidos graxos e lipídeos em *Leishmania* spp.**

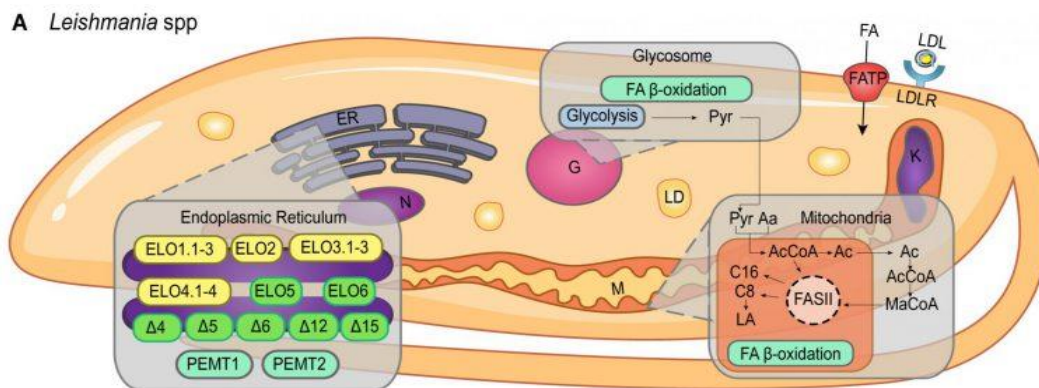
Descritos como organismos com uma capacidade limitada de biossíntese *de novo* (Naderer et al., 2006), os parasitos do gênero *Leishmania* dispõem de mecanismos para capturar uma diversidade de nutrientes do ambiente hospedeiro, entre eles os ácidos graxos (Naderer et al., 2006) (Uttaro, 2014) (Figura 2). Estudos com amastigotas de diferentes espécies de *Leishmania* demonstraram a presença de derivados de esfingolipídeos e fosfolipídeos do hospedeiro em seu lipidoma (Henriques et al., 2003). Uma vez internalizados, os lipídios exógenos são frequentemente remodelados em lipídeos de importância para a nutrição e sobrevivência do parasito (Uttaro, 2014). O sequestro de colesterol proveniente do hospedeiro é amplamente descrito em *L. amazonensis*, visto que estes não estão aptos a realizar a síntese *de novo* dessa molécula (de Aquino et al., 2021).

A captura de ácidos graxos, como ácido mirístico, ácido palmítico e ácido palmitoleico, também já foi descrita em *Leishmania*. Entretanto, o processo de internalização dessas moléculas ainda não é bem conhecido (de Aquino et al., 2021). Lipoproteínas endocitadas são degradadas em lisossomos e a hidrólise dos lipídeos fornece grupos hidrofílicos, ácidos graxos e esteróides para o parasito (Uttaro, 2014). A captação de ácidos graxos mediada por translocação associada a transportadores de membrana também é um mecanismo citado na literatura (Glatz et al., 2010) (K. Zhang & Beverley, 2010). Alternativamente, o conteúdo lipídico do lúmen e da membrana do vacúolo parasitóforo representa uma fonte de nutrientes para o parasito, onde a



disponibilidade glicosfingolípídeos e esfingomielina estimula a replicação das formas amastigotas no interior dos macrófagos (Naderer et al., 2006) (K. Zhang & Beverley, 2010).

A biossíntese de ácidos graxos com cadeias carbônicas longas a partir dos ácidos graxos de cadeias menores sequestrados do ambiente hospedeiro, requer ativação dos substratos por esterificação com coenzima A, originando Acil-Coa, que se liga às proteínas carreadoras de grupo acila (ACPs). As ACPs ligadas a ácidos graxos então seguem para a via de síntese de ácidos graxos dirigida pelo complexo enzimático de síntese de ácidos graxos FAS. Análises *in silico* demonstraram a presença de genes da via FAS II mitocondrial no genoma de *Leishmania* (Berriman et al., 2005) (Uttaro, 2014) (Figura 2), mostrando que estes organismos provavelmente sintetizam ácidos graxos na mitocôndria por essa via. Alternativamente, outro sistema envolvido na biossíntese de ácidos graxos de cadeia longa em *Leishmania* é o sistema ELO, composto por elongases (Livore et al., 2007) (Figura 2). As elongases são proteínas transmembranares que catalisam as reações de extensão das cadeias de ácidos graxos a partir de C4 formando ácidos graxos de até 18 carbonos (Lee et al., 2006). Variantes das elongases (ELO 5 e ELO 6) (Uttaro, 2014), em conjunto com as desaturases, atuam em vias de produção de ácidos graxos poli-insaturados (Paloque et al., 2019), adicionando ligações duplas às cadeias acila (Uttaro, 2014). O sequenciamento do genoma de *L. major* mostrou a presença de doze genes codificantes para as elongases, além de dois pseudogenes (Lee et al., 2006) (Uttaro, 2014). Resultados semelhantes foram obtidos para as desaturases (Uttaro, 2014). Isso demonstra a importância de tais maquinarias enzimáticas, uma vez que elas produzem diferentes tipos de ácidos graxos necessários para o desenvolvimento das *Leishmania* spp. durante o seu ciclo de vida.

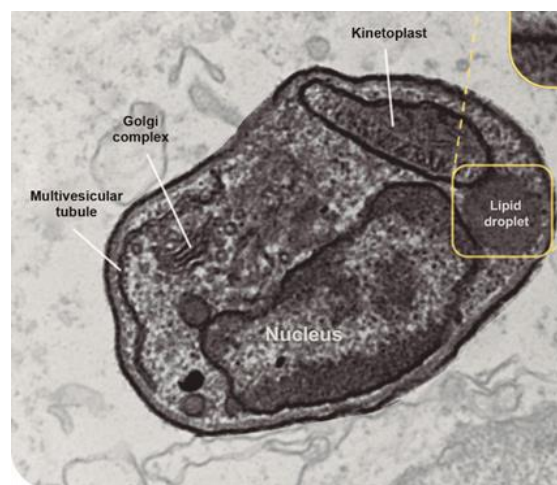


**Figura 2:** Representação esquemática do metabolismo de ácidos graxos e lipídios em *Leishmania* spp. A captura de lipoproteínas é mediada por receptores de membrana (LDLR-LDL), enquanto ácidos graxos podem ser capturados diretamente por transportadores de membrana (FA-FATP). O sistema FAS II dirige a síntese de ácidos graxos na mitocôndria. Alternativamente, o sistema ELO, presente no retículo endoplasmático, é capaz de alongar cadeias de ácidos graxos através de elongases e adicionar insaturações através de desaturases. O armazenamento de ácidos graxos e lipídios complexos ocorre nos corpúsculos lipídicos (LD). A degradação de ácidos graxos ocorre por beta-oxidação na mitocôndria e nos glicosomos (G). FAS: complexo de síntese de ácidos graxos, ELO: elongases, Delta: desaturase, LD: lipid droplets (corpúsculo lipídico), G: glicosomo, M: mitocôndria, K: cinetoplasto, LDLR: receptor de lipoproteína de baixa densidade, LDL: lipoproteína de baixa densidade, FA: ácidos graxos, FATP: proteínas transportadoras de ácidos graxos, PEMT (fosfatidiletanolamina N-metiltransferase) (extraído de de Aquino et al., 2021).

Análises funcionais revelaram a regulação positiva de enzimas envolvidas na β-oxidação de ácidos graxos em amastigotas de *L. donovani* e *L. major* em relação às promastigotas (Naderer et al., 2006). De fato, a forma amastigota parece não depender tanto da mobilização de glicose em comparação às promastigotas, compensando a menor atividade glicolítica através do aumento da β-oxidação de lipídios (Saunders et al., 2014). Em contraste, as formas promastigotas parecem depender mais da oxidação da glicose para produção de ATP em relação às amastigotas, em que a reposição dos níveis de ATP e NAD<sup>+</sup> é feita através da fermentação do succinato no glicosomo e a produção de glutamato e glutamina ocorre através do aproveitamento de intermediários do ciclo dos ácidos tricarboxílicos (Saunders et al., 2014).

O armazenamento de lipídeos já foi descrito em tripanosomatídeos (Pereira et al., 2011). Em *Leishmania*, o armazenamento de lipídios se dá através da formação de corpúsculos lipídicos no citosol como foi verificado em amastigotas de *L. major* (Rabhi

et al., 2016) (Figura 3). Corpúsculos lipídicos são estruturas delimitadas por uma monocamada de fosfolipídeos, e acumulam colesterol exógeno em seu interior (Rabhi et al., 2016). Além disso, esses corpúsculos podem atuar como sítios de produção de prostaglandinas E2 (PGE2) como verificado em *L. infantum* (Tavares et al., 2021). As PGE2 são essenciais para promover a infecção dos macrófagos, mostrando a importância dos corpúsculos lipídicos na evolução da doença (de Aquino et al., 2021).



**Figura 3:** Microscopia eletrônica de transmissão de amastigota de *Leishmania*. Destacado em amarelo, nota-se o corpúsculo lipídico (*lipid droplet*), uma organela de armazenamento de ácidos graxos e lipídios definida por uma monocamada de fosfolipídios. Essas organelas são responsáveis pelo armazenamento de ácidos graxos e ésteres de colesterol, bem como pela metabolização de precursores de PGE2 (extraído de Tavares et al., 2021).

Embora a presença de ácidos graxos e lipídeos exógenos entre os parasitos do gênero *Leishmania* tenha sido amplamente detectada, **mecanismos de incorporação e transporte dessas moléculas ainda não são bem conhecidos**. Além disso, **o metabolismo de ácidos graxos e lipídeos em *Leishmania* spp. ainda é pouco compreendido** em comparação a outros tripanosomatídeos como *T. brucei* e *T. cruzi*. Portanto, o estudo do metabolismo de ácidos graxos e lipídios em *Leishmania* spp. e a sua relação com a sobrevivência do parasito em ambiente hospedeiro é uma área de grande interesse.

#### 1.4. Proteínas de ligação a ácidos graxos (FABPs)

As proteínas de ligação a ácidos graxos (FABPs) fazem parte da superfamília das calicinas composta por proteínas associadas ao metabolismo e transporte de moléculas hidrofóbicas (Chmurzyńska, 2006). A superfamília das calicinas inclui também as lipocalinas, a avidina e a triabina (Flower et al., 2000). Os membros da família das calicinas são formados por um barril composto por fitas-beta antiparalelas que envolve uma cavidade hidrofóbica capaz de acomodar o ligante (Chiu et al., 2010) (Flower, 1996). No caso de FABPs, existe uma grande variedade de substratos, não apenas ácidos graxos, mas também ácido retinóico, heme, esteróis e ácidos biliares (Chmurzyńska, 2006) (Flower et al., 2000).

FABPs já foram identificadas em diversos organismos patogênicos como *Echinococcus granulos*, *Schistosoma mansoni*, *Ascaris suum*, *Mycobacterium tuberculosis*, *Staphylococcus aureus* e *Streptococcus pneumoniae*. As FABPs dos helmintos parasitos *E. granulos*, *S. mansoni* e *A. suum* possuem uma estrutura similar às FABPs de mamíferos, assumindo a forma de um barril-beta contendo uma cavidade acessível ao solvente que acomoda o ligante hidrofóbico (de Aquino et al., 2021). Um motivo hélice-alça-hélice em um dos lados do barril recobre esta cavidade, contribuindo para limitar a entrada e saída desses ligantes ao citosol (de Aquino et al., 2021). Em mamíferos, diferentes isoformas da FABPs estão envolvidas no metabolismo de ácidos graxos e lipídeos, participando do tráfego de ácidos graxos para a mitocôndria e perioxissomo, síntese de membrana, da formação de corpúsculos lipídicos em adipócitos e na regulação de respostas inflamatórias devido ao seu envolvimento na conversão de ácidos graxos em eicoisanóides (Furuhashi & Hotamisligil, 2008).

As FABPs são conservadas em organismos eucarióticos, e desempenham um papel central no metabolismo de ácidos graxos e lipídeos em diversas espécies

(Furuhashi & Hotamisligil, 2008). Considerando-se que o metabolismo de ácidos graxos e lipídeos é essencial para *Leishmania*, nossa hipótese é que *L. amazonensis* expressa uma proteína de ligação a ácidos graxos (FABP-like). Buscas na base de dados TriTrypDB (Aslett et al., 2009) por proteínas homólogas à FABP humana realizadas pela Profa. Maria Fernanda Laranjeira da Silva em colaboração com o Prof. Roberto K. Salinas e a Dra. Simona Tomaselli (CNR, Itália), retornaram um gene candidato anotado no genoma de *L. mexicana* como *LmxM.34.3070* que codifica uma proteína putativa altamente conservada no gênero *Leishmania*, com 822 aminoácidos e massa molecular de aproximadamente 92 kDa (Rogers et al., 2011). A anotação deste gene pode ser encontrada nos seguintes bancos de dados: OrthoMCLDB (ID: [OG6\\_156954](#)), TriTrypDB (ID: [Gene - LmxM.34.3070](#)).

### **1.5. *LmxM.34.3070* é conservado no gênero *Leishmania***

Para verificar a existência de sequências de DNA com alta similaridade em outros tripanossomatídeos, a sequência de nucleotídeos de *LmxM.34.3070* foi analisada por BLAST (Altschup et al., 1990) no TriTrypDB. Na Tabela 1 estão mostradas as sequências mais bem ranqueadas pela função de pontuação do BLAST. Nesta busca, o BLAST retornou sequências gênicas bem pontuadas, com valores nulos de *E-value* e altos percentuais de similaridade em outras espécies de *Leishmania*, demonstrando que o gene *LmxM.34.3070* não é espécie-específico. Esse resultado sugere a existência de potenciais ortólogos de *LmxM.34.3070* em outras espécies de *Leishmania*.

Tabela 1. Potenciais homólogos a *LmxM.34.3070* em outras espécies de *Leishmania* (Score > 2000)

Acesso	Tamanho (pb)	E-value	Organismo	Cromossomo	%ID	Score
LAEL147_000782000	2469	0	<i>Leishmania aethiopica</i>	35	94.94	3889
LAMA_000797300	2469	0	<i>Leishmania amazonensis</i>	34	99.23	4367
LARLEM1108_350037900	2469	0	<i>Leishmania arabica</i>	35	94.17	3804
LBRM2903_340038400	2469	0	<i>Leishmania braziliensis</i>	34	83.48	2587
LbrM.34.2980	2469	0	<i>Leishmania braziliensis</i>	34	83.56	2596
LbrM.34.2.002980	2469	0	<i>Leishmania braziliensis</i>	34	83.56	2596
LdBPK_353120.1	2469	0	<i>Leishmania donovani</i>	35	94.45	3831
LdBPK.35.2.003120	2469	0	<i>Leishmania donovani</i>	35	94.45	3831
LENLEM3045_350037100	2472	0	<i>Leishmania enriettii</i>	35	82.50	2491
LGELEM452_350037500	2469	0	<i>Leishmania gerbilli</i>	35	94.49	3840
LINF_350036000	2469	0	<i>Leishmania infantum</i>	35	94.45	3831
LmjF.35.3070	2469	0	<i>Leishmania major</i>	35	93.68	3750
LMJLV39_350038000	2469	0	<i>Leishmania major</i>	35	93.60	3741
LPAL13_340035500	2469	0	<i>Leishmania panamensis</i>	34	83.48	2587
LPMP_342920	2469	0	<i>Leishmania panamensis</i>	34	83.48	2587
LMARLEM2494_350036500	2472	0	<i>unclassified Leishmania</i>	35	80.80	2266
LTRL590_350036700	2469	0	<i>Leishmania tropica</i>	35	94.90	3885
LTULEM423_350037900	2469	0	<i>Leishmania turanica</i>	35	94.73	3867

## 1.6. Modelagem preliminar da proteína *LmxM.34.3070*

Modelagens preliminares com o programa PHYRE2 (Kelley et al., 2015) e InterPRO sugeriram a presença de um domínio AttH *lipocalin-like* ([IPR010791](#)) na porção N-terminal de *LmxM.34.3070* (Figura 4). O domínio N-terminal (*LmxM.34.3070*<sub>52-488</sub>) AttH *lipocalin-like* consiste em dois barris-beta achatados, compostos por 13 ( $\beta$ 1- $\beta$ 13) e 10 ( $\beta$ 14- $\beta$ 23) fitas-betas antiparalelas, respectivamente (Chiu et al., 2010). As informações sobre esse domínio derivam de dados cristalográficos (PDB:2ICH) de um possível ortólogo codificado pelo gene NE1406 (AttH) de *Nitrosomonas europaea*, uma bactéria Gram-negativa (Chiu et al., 2010). A mesma análise sugeriu a presença de um domínio C-terminal (*LmxM.34.3070*<sub>665-809</sub>) semelhante à Proteína Celular de Ligação a Retinol (CRBP) de *Danio rerio* (PDB 1KQW) (Calderone et al., 2002) (Figura 4). As CRBPs são membros da superfamília das calicinas, pequenas proteínas secretadas que exibem alta afinidade e seletividade por moléculas hidrofóbicas (Flower, 1996). Adicionalmente, as CRBPs compartilham semelhanças estruturais com as FABPs que apresentam dez fitas-beta antiparalelas formando um barril-beta fechado por duas alfa-hélices inter-fita em uma das extremidades (Flower, 1996). Diante disso, **o domínio C-terminal FABP-like codificado pelo gene *LmxM.34.3070* foi denominado “LeiFABP”.**

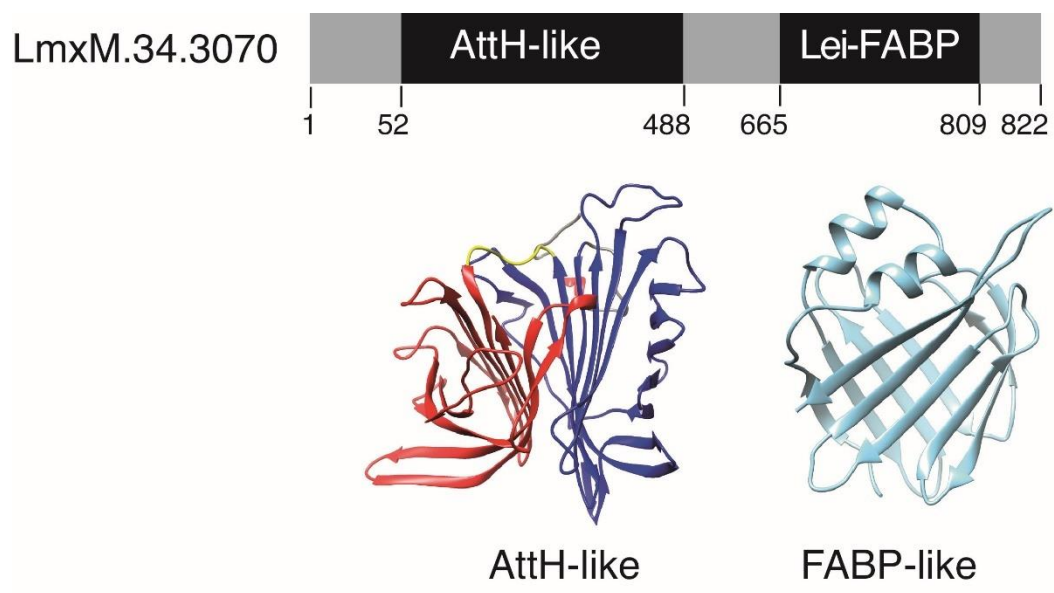
A partir desta análise preliminar, hipotetizamos que o gene *LmxM.34.3070* poderia estar associado ao metabolismo de ácidos graxos e lipídios. Nossas buscas por BLAST no TriTrypDB revelaram a conservação desse gene no gênero, o que reforça a hipótese de que o gene *LmxM.34.3070* seja importante fisiologicamente para os parasitos. Na tentativa de caracterizar uma FABP em *Leishmania*, produzimos a forma recombinante do domínio C-terminal FABP-like de *LmxM.34.3070* de *Leishmania amazonensis* para ensaios de biologia estrutural. Para isso, desenhamos diferentes

construções de biologia molecular e as avaliamos quanto a expressão, solubilidade e rendimento, sendo elas:

- **LmxM.34.3070<sub>652-822</sub>(pET24a)**: DNA codificador do trecho compreendido entre os aminoácidos 652 e 822 de LmxM.34.3070 com sítios de restrição de *Bam*HI (5') e *Hind*III (3') para ser clonado em pET24a.
- **LmxM.34.3070<sub>652-822</sub>(pET28a)**: DNA codificador do trecho compreendido entre os aminoácidos 652 e 822 de LmxM.34.3070 com sítios de restrição de *Nde*I (5') e *Bam*HI (3') para ser clonado em pET28a.
- **LmxM.34.3070<sub>638-822</sub>(pET28a)**: DNA codificador do trecho compreendido entre os aminoácidos 638 e 822 de LmxM.34.3070 com regiões de sobreposição para clonagem em pET28a pela Metodologia de Gibson.
- **LmxM.34.3070<sub>669-822</sub>(pET28a)**: DNA codificador do trecho compreendido entre os aminoácidos 669 e 822 de LmxM.34.3070 com sítios de restrição de *Nde*I (5') e *Hind*III (3') para ser clonado em pET28a.

A construção mais promissora foi analisada por técnicas de biofísica quanto ao seu conteúdo de estrutura secundária, estabilidade térmica e estado de agregação. Esperamos que a caracterização estrutural de LeiFABP contribua para um melhor entendimento do papel de domínios FABP e do envolvimento de LmxM.34.3070 no metabolismo de ácidos graxos e lipídios em espécies do gênero *Leishmania*.





**Figura 4:** Predição de domínios da proteína LmxM.34.3070 por PHYRE2. O programa sugeriu a presença de um domínio semelhante a proteína 2ICH *AttH-like* na região N-terminal (resíduos 52-488), e um domínio FABP-like na região C-terminal (resíduos 665-809). Confiança: 100 e 98% respectivamente.

---

## **2. Objetivos**

---

**Objetivo do trabalho:** Clonar, expressar e purificar diferentes fragmentos da região C-terminal contendo o domínio FABP-like (LeiFABP) de LmxM.34.3070 de *L. amazonensis* e determinar a construção mais adequada para estudos biologia estrutural.

**Objetivos específicos:**

1. Desenhar construções correspondendo a diferentes trechos da região C-terminal de LmxM.34.3070 de *L. amazonensis*.
2. Otimizar as condições de expressão em *E. coli*, purificação e solubilidade das construções obtidas.
3. Avaliar o conteúdo de estrutura secundária do domínio LeiFABP por espectroscopia de dicroísmo circular.
4. Avaliar o estado de agregação do domínio LeiFABP recombinante por espalhamento de luz.
5. Obter espectros de RMN unidimensionais da construção LeiFABP recombinante.

---

## **3. Materiais e métodos**

---

### 3.1. Predição de estrutura secundária e terciária

O desenho dos clones e a escolha de fragmentos codificadores da região C-terminal FABP-like de LmxM.34.3070 foi guiado através de diferentes predições de estrutura utilizando PSIPRED, Raptor-X e ColabFold. Inicialmente, a sequência de aminoácidos de LmxM.34.3070 foi utilizada como *input* para [PSIPRED](#) (Jones, 1999). PSIPRED é um método proposto por David T Jones em 1999 que combina a abordagem das Matrizes Posicionais com um algoritmo de redes neurais para prever o conteúdo de estruturas secundárias em sequências de proteínas (Jones, 1999). Em seguida, a sequência de aminoácidos de LmxM.34.3070 foi submetida a dois programas para predição da estrutura tridimensional: Raptor-X e ColabFold. [Raptor-X](#) é um algoritmo de redes neurais treinadas para prever características de sequências de aminoácidos, como o potencial para prever contatos inter-residuais (*residue pairwise*), conteúdo de estrutura secundária, e o grau de acessibilidade ao solvente (Xu & Wang, 2019)(Wang et al., 2017). Raptor-X calcula uma matriz de co-evolução de pares de resíduos a partir da análise de alinhamentos múltiplos de sequências, e utiliza essa informação para prever contatos inter-residuais (Xu & Wang, 2019) (Wang et al., 2017). De forma similar, o [ColabFold](#) utiliza técnicas de aprendizado de máquina para prever a estrutura de uma proteína a partir da sequência de aminoácidos. A acurácia da estrutura final em função da sequência de aminoácidos é estimada através do *score pLDDT*, que considera violações físicas, geométricas e estereoquímicas de cada resíduo do modelo (Jumper et al., 2021)

### 3.2. Células:

- Promastigotas de *Leishmania amazonensis* cepa PH8;
- *Escherichia coli* cepas de clonagem: DH5alpha e SURE;
- *Escherichia coli* cepas de expressão: BL21(DE3), pLysS, C43, RIL, RIP-L.

### 3.3. Meios de cultura:

- **M199 (Gibco):** pH 7, contendo L-glutamina e suplementado com 10% de soro fetal bovino (FBS) inativado, 0,25% de hemina, 0,1 mM de adenina, 100 U/mL de penicilina, 100 µg/mL de estreptomicina, 40 mM de HEPES e 40 mM de NaHCO<sub>3</sub>;
- **Luria Berthani (LB Broth, Powder (Lennox)):** 20 g em 1 litro de água destilada;
- **Luria Berthani (LB Ágar, (Lennox)):** 35 g em 1 litro de água destilada.

### 3.4. Extração de DNA genômico de *L. amazonensis*

A extração de DNA genômico (Longmire et al., 1987) de *L. amazonensis* foi uma etapa essencial para a obtenção do DNA molde para amplificação dos alvos de interesse por reação em cadeia da polimerase (PCR). Inicialmente, 50 mL de cultura de promastigotas de *L. amazonensis* em fase estacionária, aproximadamente  $5 \times 10^7$  promastigotas, foram coletadas a 3000 x g por 15 minutos. As células foram lavadas uma vez com PBS (NaCl 137 mM, KCl 2,7 mM, Na<sub>2</sub>HPO<sub>4</sub> 10 mM, K<sub>2</sub>HPO<sub>4</sub> 2 mM, pH 7,4) e então ressuspensas em 6 mL de tampão de lise (10 mM Tris, 2 mM EDTA, 400 mM NaCl, SDS 1%). Em seguida foi realizada a adição de Proteinase K (20 mg/mL) e incubação das células a 42°C. Após 2 horas adicionou-se 2 mL de NaCl 6 M, seguido de agitação vigorosa em *vortex* por 15 segundos e centrifugação a 3000 x g por 15 minutos. Para a precipitação do material genético foi adicionado etanol absoluto gelado, seguido de centrifugação a 3000 x g por 10 minutos. O material genético foi coletado, lavado com etanol 70%, e o *pellet* contendo o DNA genômico de *L. amazonensis* foi solubilizado em tampão de eluição TE (10 mM Tris, 1mM EDTA). A amostra foi quantificada por espectrofotometria no equipamento Nanodrop ( $\lambda = 260$  nm). O rendimento de DNA genômico nesta extração foi de 440 ng/µL.

### 3.5. Reação em cadeia da polimerase (PCR)

Os oligonucleotídeos iniciadores foram desenhados para a amplificar diferentes trechos codificantes da região C-terminal de *LmxM.34.3070*. O desenho dos iniciadores baseou-se no

gene homólogo a *LmxM.34.3070* em *L. amazonensis*, que foi recuperado do genoma sequenciado depositado na plataforma *The Leishmania (L.) amazonensis Genome DB* ([LGE - Leishmania amazonensis](#)) (Real et al., 2013). Os iniciadores além de amplificar diferentes trechos codificantes da região C-terminal, também foram desenhados para inserir sítios de restrição específicos nas extremidades dos fragmentos amplificados. O volume final de reação foi de 25 µL, sendo 1 µL de DNA genômico de *L. amazonensis* a 440 ng/µL como molde, 0,25 µL de cada iniciador na concentração de 10 µM, 11 µL de água livre de nucleases, e 12,5 µL de DreamTaq Green PCR Master Mix (Thermo Fisher). Para amplificar o fragmento codificante utilizado na montagem de Gibson foi realizada uma PCR com a Phusion High-Fidelity DNA-Polimerase a partir do DNA genômico de *L. amazonensis*, em volume final de 40 µL (8 µL de tampão HF, 2 µL de cada iniciador na concentração de 10 µM, 4 µL de dNTPs (a 2 mM), 4 µL de DNA genômico de *L. amazonensis*, 0,4 µL de Phusion High-Fidelity DNA-Polimerase e água livre de nucleases). Todos os oligonucleotídeos foram avaliados *in silico* pela ferramenta *OligoAnalyser Tool* ([OligoAnalyzer Tool - Primer analysis and Tm Calculator | IDT](#)) quanto à temperatura de *melting*, conteúdo GC, formação de *hairpins*, formação de heterodímeros e homodímeros (Tabela 2).

Foram realizados 35 ciclos de amplificação no termociclador SureCycler 8800 (Agilent Technologies): desnaturação a 94°C por 30 segundos, anelamento por 30 segundos em temperaturas específicas determinadas após experimentos de PCR com gradiente de temperatura, extensão a 72°C por 45 segundos. Finalmente realizou-se uma extensão final a 72°C por 10 minutos. Os produtos de PCR foram visualizados por eletroforese em gel de agarose (1%) contendo o intercalante Syber Safe (Invitrogen). A corrida eletroforética foi iniciada com a aplicação de uma diferença de potencial de 85V por 1 hora em tampão TAE (Tris 40 mM, acetato 20 mM, EDTA 1 mM).

Sequência	Apelido	Tamanho do produto (pb)	Sequência codificada	$T_m$ °C	%G C	Hairpin (kcal/mol)	Dímeros (máximo de bases complementares)	Homodímeros (máximo de bases complementares)	Método de clonagem	Tipo de inserção
F:5'TATGGATCCATGCTGCAACCGCCCTCCATG'3	Primer 1	552	LmxM34.3070652-822	66.8	56.7	-1.8	4	10	Restrição	Sítio BamHI
R:5'TAAAAGCTTGCTGCCGCGCGGCACCAGTTTCTTGTCGGATGTTGGAGACCC'3	Primer 2	552	LmxM34.3070652-822	73.0	56.9	-4.0	4	12	Restrição	Sítio HindIII
F:5'ATACATATGCTGCAACCGCCCTCCATG'3	Primer 3	534	LmxM34.3070652-822	59.4	66.7	0.39	5	6	Restrição	Sítio NdeI
R:5'ATAGGATCCCTATTTCTTGTCGGATGTTGGAGACCC'3	Primer 4	534	LmxM34.3070652-822	63.8	48.1	-1.5	5	6	Restrição	Sítio BamHI
F:5'CGGGCTTTGCTATTTCTTGTCGGATGTTGGAGACCC'3	Primer 5	574	LmxM34.3070638-822	65.0	51	-1.24	4	3	Gibson	Overlap
R:5'GCCTGGTGATCGACAGCTGCGTCT'3	Primer 6	574	LmxM34.3070638-822	65.0	63	-0.48	4	6	Gibson	Overlap
F:5'ATACATATGGGCACCTGGATCATGGACCC'3	Primer 7	477	LmxM34.3070669-822	63.0	51.7	-1.53	4	6	Restrição	Sítio NdeI
R:5'ATAAAGCTTCTATTTCTTGTCGGATGTTGG'3	Primer 8	477	LmxM34.3070669-822	57.3	36.7	0.06	4	6	Restrição	Sítio HindIII
F:5'TGTCGATCACCAGGCCGCTGCT'3	Primer 9	5253	pET28a linear	65.0	64	-0.97	4	4	Gibson	Overlap
R:5'AAGAAATAGCAAAGCCCGAAAGGAAGCTGAG'3	Primer 10	5253	pET28a linear	62.0	45	-1.95	4	4	Gibson	Overlap

**Tabela 2:** Características físico-químicas do oligonucleotídeos analisadas no [OligoAnalyzer Tool](#), tamanho dos produtos de PCR, sequência codificada por cada par de primers, estratégia de clonagem utilizada na obtenção de cada clone e tipo de inserção realizada na amplificação por cada par de primers.



### 3.6. Tecnologia de DNA recombinante

Para a expressão heteróloga dos diferentes fragmentos da região C-terminal de LmxM.34.3070, recorreremos à tecnologia de DNA recombinante para inserir a sequência de nucleotídeos codificante nos plasmídeos de interesse. Os plasmídeos são vetores de DNA circulares e epissomais, desenhados para conter sequências regulatórias, marcas para seleção e uma origem de replicação reconhecida pelo hospedeiro (Stryjewska et al., 2013). Utilizamos os seguintes plasmídeos de subclonagem e de expressão para produzir os clones de interesse em *Escherichia coli*:

- *pCR2.1-TOPO (3900 pb)*: plasmídeo conjugado a Topoisomerase I para subclonagem de produtos de PCR, contendo *overhangs* de desoxitimina, origem de replicação, gene repórter *lacZ*, região multiclonal, marca de resistência à canamicina, marca de resistência à ampicilina, promotor da RNA Polimerase do fago T7, regiões M13.
- *pGEM®-T Easy Vector (3015 pb)*: plasmídeo para subclonagem de produtos de PCR, contendo *overhangs* de desoxitimina, origem de replicação, gene repórter *lacZ*, região multiclonal, marca de resistência à ampicilina, promotor da RNA Polimerase do fago T7, regiões M13.
- *pET24a (+) (5310 pb)*: plasmídeo de expressão contendo origem de replicação, região multiclonal, marca de resistência à canamicina, promotor da RNA Polimerase do fago T7, operador *lac* O1, sítio RBS (Shine Dalgarno) de ligação ao ribossomo e *tag* de histidinas na região C terminal.
- *pET28a (+) (5369 pb)*: plasmídeo de expressão contendo origem de replicação, região multiclonal, marca de resistência à canamicina, promotor da RNA Polimerase do fago T7, operador *lac* O1, sítio RBS (Shine Dalgarno) de ligação ao ribossomo, *tag* de

histidinas nas regiões N e C terminal, e sítio de clivagem por trombina na região N-terminal.

### 3.7. Subclonagem de fragmentos codificantes obtidos por PCR

As subclonagens foram realizadas com diferentes protocolos no intuito de inserir os produtos de PCR em plasmídeos contendo o gene *lacZ* da Beta-galactosidase para identificação de colônias recombinantes pelo método *Blue/White colony screening*. Além disso, a dupla digestão dos subclones com as endonucleases de interesse geram um perfil eletroforético claro das bandas digeridas e confirmam a liberação dos insertos com as extremidades coesivas. Os fragmentos de PCR **LmxM34.3070<sub>652-822</sub>** para clonagem em pET24a e **LmxM34.3070<sub>652-822</sub>** para clonagem em pET28a foram subclonados em vetor pCR2.1 com o kit TOPO-TA® e o fragmento de PCR **LmxM34.3070<sub>669-822</sub>** para clonagem em pET28a foi subclonado em vetor pGEM®.

#### 3.7.1. Subclonagem em vetor pCR2.1 com o kit TOPO-TA®:

Os produtos das reações por PCR realizadas para obtenção dos DNA codificante da região **LmxM34.3070<sub>652-822</sub>** para clonagem nos vetores de expressão pET24a e pET28, respectivamente, passaram antes por uma subclonagem em pCR2.1 (com o kit TOPO-TA®) (Green & Sambrook, 2021). O protocolo da reação de subclonagem seguiu as recomendações do fabricante. O volume final da reação de ligação foi 6 µL, correspondentes a 4 µL de produto de PCR, 1 µL de pCR 2.1 TOPO-TA® e 1 µL da solução estoque de sais do *kit*. As soluções de reação foram gentilmente pipetadas e incubadas à temperatura ambiente por 30 minutos.

#### 3.7.2. Subclonagem em pGEM®-T Easy Vector Systems:

O DNA codificante da região **LmxM34.3070<sub>669-822</sub>** para clonagem em vetor de expressão pET28a foi obtido por PCR, purificado com o *Kit* de Purificação GeneJET PCR (Thermo

Scientific) seguindo as recomendações do fabricante e quantificado por NanoDrop ( $\lambda = 260$  nm). O produto de PCR purificado e quantificado foi subclonado em pGEM: 5  $\mu$ L de 2X *Rapid Ligation Buffer* de T4 DNA Ligase, 1  $\mu$ L T4 DNA Ligase (3 U/ $\mu$ L), 1  $\mu$ L de pGEM®-T Easy Vector (50 ng/  $\mu$ L), 24 ng de produto de PCR purificado. A massa de inserto utilizada seguiu uma razão molar de 3:1 (inserto:vetor) de acordo com a Equação 1:

$$m_{\text{inserto}} = \frac{3 \text{ pares de base inserto}}{1 \text{ pares de base do vetor}} m_{\text{vetor}} \quad (\text{Equação 1})$$

O restante do volume foi completado com água livre de nucleases para 10  $\mu$ L. A reação foi incubada *overnight* à 4°C.

### 3.8. Transformação de bactérias competentes

As cepas de *E. coli* utilizadas para as subclonagens e clonagens realizadas nesse trabalho foram DH5alpha e SURE, cujos fenótipos fazem delas hospedeiras interessantes para a produção de DNAs recombinantes. O protocolo de transformação foi realizado por eletroporação ou choque térmico. No caso da eletroporação, células de *E. coli* SURE eletrocompetentes foram transformadas com os subclones pCR2.1(TOPO)- **LmxM34.3070<sub>652-822</sub>**. Foram adicionados 2  $\mu$ L de produto de reação pCR.2.1(TOPO)- **LmxM.34.3070<sub>652-822</sub>** ao tubo contendo as células, que foram eletroporadas em cubetas de 2 mm e em seguida, os tubos foram acrescidos de 900  $\mu$ L de LB. Os parâmetros do eletrochoque no eletroporador foram: resistência 100 ohms, capacitância 25 mF,

No caso do choque térmico, o seguinte protocolo foi usado para o subclone pGEM- **LmxM.34.3070<sub>669-822</sub>** e para todos os clones de expressão (**LmxM.34.3070<sub>652-822</sub>(pET24a)**, **LmxM.34.3070<sub>652-822</sub>(pET28a)**, **LmxM.34.3070<sub>638-822</sub>(pET28a)** e **LmxM.34.3070<sub>669-822</sub>(pET28a)**): aos tubos contendo células DH5alpha competentes foram adicionados 5  $\mu$ L de DNA recombinante. Em seguida, as células foram gentilmente suspendidas com a pipeta, incubadas em gelo e após 45 minutos, foram submetidas ao choque térmico de 42°C por 1 minuto em

banho-maria. Após a incubação em gelo por 10 minutos, as células foram acrescidas de 500 µL de meio LB líquido e cultivadas por agitação moderada à 200 rpm na incubadora por 1 hora a 37 °C.

### 3.9. Blue-White colony screening

Posteriormente, as culturas de células de *E.coli SURE* contendo os subclones pCR2.1-**LmxM.34.3070<sub>652-822</sub>** e de DH5alpha contendo o subclone pGEM-**LmxM.34.3070<sub>669-822</sub>** foram plaqueadas em meio LB-Ágar contendo 50 µg/ml de ampicilina, X-Gal e IPTG para a seleção de colônias pelo método *Blue/White colony screening* e incubadas na estufa *overnight* a 37 °C. As colônias brancas escolhidas foram coletadas e inoculadas em LB contendo 50 µg/ml de ampicilina e cultivadas *overnight* a 37°C sob agitação moderada à 200rpm na incubadora. Essas culturas foram direcionadas para purificação dos subclones recombinante através de extração de DNA plasmidial por mini-preparação (Miniprep) e em seguida, ensaios de dupla digestão.

### 3.10. Mini-preparação de plasmídeo bacteriano (MINIPREP)

A miniprep é uma técnica bioquímica utilizada para isolar um plasmídeo de interesse de células bacterianas (Birnboim & Doly, 1979). Inicialmente, 3 mL de cultura (em meio LB contendo ampicilina) de células *E. coli* DH5alpha ou SURE transformadas com os plasmídeos de interesse, foram transferidos para tubos eppendorf e centrifugados a 10000 x g à temperatura ambiente por 1 minuto. As células foram colhidas e suspensas com o tampão P1 (Tris 50 mM, EDTA 50mM e RNase A a 100 µg/mL). Após a ressuspensão, adicionou-se o tampão de lise P2 (NaOH 0.2 N e SDS 1%) e o tubo foi invertido 5 vezes. Neste momento, a amostra adquiriu aspecto viscoso devido à ação do detergente e do álcali levando ao rompimento da membrana e lise celular das bactérias. Após 5 minutos, a amostra foi acrescida de tampão de precipitação P3 (acetato de potássio 5M e ácido acético glacial 12%). A amostra

foi então centrifugada a 10000 x g e o sobrenadante coletado. Após a coleta do sobrenadante, adicionou-se 600µl de isopropanol absoluto para precipitar os plasmídeos. A amostra foi novamente centrifugada, lavada com etanol 70% e eluída em tampão de eluição TE (Tris-HCl 10mM, EDTA 1 mM, pH=8) ou em água livre de nucleases. O rendimento da preparação foi quantificado por espectrofotometria a 260 nm no Nanodrop (Thermo Fisher). Todos os produtos de subclonagem e clonagem em vetor de expressão foram purificados seguindo este protocolo.

### **3.11. Clonagem por restrição em vetor de expressão pET24a e pET28a**

#### **3.11.1. Digestões dupla**

Após a purificação dos subclones, estes foram duplamente digeridos com as endonucleases correspondentes aos sítios de restrição inseridos na PCR para os ensaios de digestão e liberação do fragmento codificante. Todas as digestões foram realizadas de acordo com as recomendações dos fabricantes quanto ao tempo de digestão, atividade *star*, temperatura, escolha do tampão de digestão e concentração final de glicerol. Foram digeridos 70 µg de DNA de cada subclone com 20 U da respectiva enzima de restrição (New England BioLabs); o tampão de digestão (New England BioLabs) utilizado foi o recomendado pelo fabricante de acordo com o *set* de enzimas de restrição de cada ensaio de dupla digestão; o volume final de cada reação foi calculado para que a concentração final de glicerol não excedesse 5%. Os vetores de expressão pET24a e pET28a passaram pelas digestões duplas com os mesmos *sets* de enzimas de seus respectivos insertos para a formação de extremidadesceptoras no sítio multiclonal. Para a montagem de cada construção, foram digeridos 50 µg de vetor de expressão.

#### **3.11.2. Purificação de fragmentos de DNA de gel de agarose**

Os fragmentos de DNA de interesse foram visualizados por eletroforese em gel de agarose (1%) e as bandas referentes aos insertos liberados das duplas digestões, bem como os

vetores linearizados foram excisadas com bisturi e purificadas com o *kit* QIAquick Gel Extraction (Qiagen) de acordo com o protocolo disponibilizado pelo fabricante. Os produtos de digestão foram analisados quanto a concentração e pureza por espectrofotometria no equipamento Nanodrop (Thermo Fisher).

### 3.11.3. Reação de ligação com a T4 Ligase

As reações de ligação para clonagem do fragmento no vetor de expressão foram realizadas partindo-se de 50 ng do vetor de expressão escolhido, sendo que a quantidade de inserto respeitou a razão molar 3:1, 4  $\mu\text{L}$  de Reaction Buffer Ligase (5X), 1  $\mu\text{L}$  de T4 Ligase e o volume final completado com água *nuclease free* para 20  $\mu\text{L}$ . Um volume de 5  $\mu\text{L}$  de cada reação de ligação foi utilizado para transformar células competentes de *E.coli* DH5alpha por choque térmico utilizando 50  $\mu\text{g/mL}$  de Canamicina para seleção. As colônias que vingaram em cada placa foram repicadas em novas placas para extração dos clones por miniprep, e submetidas à reação de PCR de colônias para confirmação da clonagem.

### 3.12. Clonagem direta por Gibson em vetor de expressão pET28a

A construção **LmxM.34.3070<sub>638-822</sub>** foi clonada em pET28a através da metodologia de Gibson (Gibson, 2011). Os fragmentos de inserto e vetor foram amplificados por PCR usando os primers 5 e 6 para o inserto, e 9 e 10 para o vetor (Tabela 2). Os produtos de PCR foram incubados com *DpnI* por 1 hora a 37°C para degradação do DNA parental, submetidos a eletroforese em gel de agarose 1% e as bandas excisadas e purificadas. Os produtos de purificação das bandas em gel foram quantificados para os cálculos de razão molar inserto:vetor. A reação de Gibson foi realizada em um volume final de 15  $\mu\text{L}$ : 100 ng de vetor de expressão, 33 ng de inserto (a quantidade de inserto seguiu a razão molar 3:1 e foi calculada de acordo com a Equação 1), 7,5  $\mu\text{L}$  de *Gibson Assemble Master Mix* contendo exonuclease de reconhecimento de extremidade 5', DNA polimerase e DNA ligase (New England BioLabs) e água livre de nucleases para um volume final de 15  $\mu\text{L}$ . A montagem de Gibson ocorreu

isotermicamente por 1 hora a 50°C no termociclador SureCycler 8800 (Agilent Technologies). O produto dessa reação (15µL) foi utilizado para transformar células *E.coli* DH5alpha por choque térmico. As células foram plaqueadas em meio LB-Ágar contendo canamicina para seleção e incubadas na estufa a 37°C *overnight*. No dia seguinte, as colônias que vingaram foram submetidas à PCR de colônia e repicadas em novas culturas para extração do clone por miniprep.

### 3.13. PCR de colônia

As colônias de DH5alpha recombinantes contendo os clones de expressão obtidos em cada clonagem realizada foram repicadas e depois inoculadas em 10 µL de água livre de nucleases. A suspensão contendo as células recombinantes foi incubada a 98°C no termociclador por 10 minutos para lisar as células por choque térmico. A amostra foi, então, centrifugada a 16000 x g, o sobrenadante foi coletado e utilizado como *template* em experimentos de PCR para detecção do gene de interesse nas colônias recombinantes. O volume final da reação de PCR foi de 25 µL, sendo 10 µL de sobrenadante proveniente do lisado de colônias recombinantes, 0,25 µL de cada iniciador (os mesmos utilizados na etapa de amplificação de cada trecho de interesse a partir do DNA genômico de *L.amazonensis*) na concentração de 10 µM, 12,5 µL de DreamTaq Green PCR Master Mix (Thermo Fisher) e água livre de nucleases para um volume final de 25 µL. Foram realizados 35 ciclos de amplificação no termociclador SureCycler 8800 (Agilent Technologies): desnaturação a 94°C por 30 segundos, associação dos iniciadores por 30 segundos com temperatura de *annealing* ideal para cada *set de primers*, extensão a 72°C por 45 segundos. Foi realizada uma extensão final a 72°C por 10 minutos. A detecção dos insertos ocorreu por visualização de fragmentos por eletroforese em gel de agarose (1%) contendo o intercalante Syber Safe (Invitrogen). A corrida eletroforética foi realizada com uma diferença de potencial de 85V por 1 hora em tampão TAE.

### **3.14. Sequenciamento de clones (confirmados por PCR de colônia) pelo método de Sanger**

Os clones confirmados por PCR de colônias foram submetidos a sequenciamento pela metodologia de Sanger utilizando o kit BigDye Terminator v3.1 Cycle Sequencing (Thermo Fisher). Os iniciadores utilizados contêm os trechos das regiões do promotor T7 e terminador T7, cujas sequências são:

**T7 Promoter: 5'- TAATACGACTCACTATAGGG-3' Forward,**

**T7 Terminator: 5'- GCTAGTTATTGCTCAGCGG-3' Reverse.**

A reação foi preparada da seguinte forma: 7 µL de tampão de sequenciamento (200 mM Tris-HCl pH=9, e 5mM de MgCl<sub>2</sub>), 1 µL de primer T7 promotor ou T7 terminator na concentração de 10 µM, 5 µL de DNA proveniente de Miniprep (500 ~ 1000 ng µL), 2 µL de Big Dye e 5 µL de água nuclease *free*. Utilizando o termociclador SureCycler 8800 (Agilent Technologies), procedeu-se à reação com os seguintes passos: *hot start* de 96°C por 2 minutos e 40 ciclos de desnaturação a 96°C por 45 segundos, associação do iniciador por 30 segundos a 50°C, extensão a 60°C a 4 minutos. Ao final da reação, foram adicionados 80 µL de isopropanol a 75% aos produtos de PCR, que foram guardados a 4°C durante à noite. No dia seguinte, as amostras foram centrifugadas a 1200 x g por 1 hora e o sobrenadante descartado. O resíduo de isopropanol foi eliminado com *spins* de 1000 x g com as placas invertidas e os tubos abertos. O sequenciamento foi realizado pela Central Analítica do IQ-USP no sequenciador automático 3500 XL Genetics Analyzers. Os resultados do sequenciamento foram analisados no *software* SnapGene e avaliados quanto à sequência esperada.

### **3.15. Teste de expressão heteróloga dos diferentes fragmentos recombinantes**

A expressão heteróloga dos diferentes clones foi realizada em células de *E. coli*. BL21(DE3) de diferentes cepas (ver Tabela 2) (Casali, 2003). As células BL21 contêm em seu



genoma um prófago T7 com genes de T7 RNA Polimerase, que reconhecem o promotor T7 presente nos plasmídeos de expressão que foram utilizados na etapa de clonagens dos diferentes fragmentos codificadores do domínio C-terminal contendo o barril FABP-like de LmxM.34.3070. As células foram transformadas com o clone de expressão de interesse devidamente sequenciado, e cultivadas em placas LB-Ágar contendo 50 µg/mL de canamicina e/ou 50 µg/mL de cloranfenicol a 37 °C durante à noite. No dia seguinte, uma colônia foi coletada e inoculada em 5ml de meio LB contendo os mesmos antibióticos e incubada a 37°C e 270 rpm *overnight*. No dia seguinte, foi realizada uma diluição de 1:100 do pré-inóculo em 50 mL de LB contendo canamicina e/ou cloranfenicol nas mesmas concentrações. A expressão inicial da proteína foi induzida pela adição de 600 µM de IPTG quando a densidade óptica a 600 nm (OD<sub>600</sub>) atingiu 0,6-0,8 (*mid-log*). A OD<sub>600</sub> foi medida no equipamento AJAX-3000PC (MiCRONAL). A indução da expressão heteróloga dos diferentes fragmentos recombinantes foi realizada a 37° por 4 horas. Amostras de 1ml de cultura não induzida, bem como amostras de culturas depois de 1, 2, 3 e 4 horas da indução foram coletadas para acompanhamento da evolução da expressão. Essas amostras foram centrifugadas, o sobrenadante descartado e o *pellet* suspenso em tampão de Laemmli para aplicação em gel de acrilamida SDS-PAGE.

Após a confirmação da expressão heteróloga, os clones de interesse foram utilizados para a expressão em larga escala dos fragmentos recombinantes. As células de *E. coli* RIP-L foram transformadas com o clone de interesse por choque térmico, crescidas a 37°C e cultivadas em placas LB-Ágar contendo 50 µg/mL de canamicina e 50 µg/mL cloranfenicol. No dia seguinte, um pré-inóculo de 50 ml de LB com os mesmos antibióticos foi preparado e depositado na incubadora a 37°C *overnight* e sob agitação moderada (200 rpm). Na manhã seguinte, foi realizada uma diluição de 1:100 de pré-inóculo em 0,5 litro de LB (50 µg/mL de canamicina e 50 µg/mL cloranfenicol) e a adição de IPTG ocorreu em OD<sub>600</sub> entre 0.6 e 0.8.

Foram testadas diferentes condições de expressão como concentração de IPTG e temperatura de indução e diferentes métodos e soluções para lise. A purificação por cromatografia de afinidade também passou por testes, em que diferentes composições de tampões A (de ligação) e B (de eluição) foram avaliados. Estas condições se encontram sumarizadas nas tabelas 4.1, 4.2 e 4.3.

### **3.16. Eletroforese unidimensional em gel de poliacrilamida (SDS-PAGE)**

O fracionamento de proteínas em malhas de poliacrilamida foi realizado de acordo com o protocolo descrito por Laemmli em 1970 (LAEMMLI, 1970). Quando submetidas a um campo elétrico, as proteínas carregadas apresentam uma velocidade de migração na malha que é proporcional ao seu tamanho. A adição de carga negativa às proteínas, neste experimento, é conferida pela adição de detergente aniônico dodecil sulfato de sódio (SDS) (LAEMMLI, 1970). Dois géis foram preparados: o gel de empilhamento, numa malha de 5% de poliacrilamida, e o gel de fracionamento numa malha de 12% de poliacrilamida.

Amostras de 1 mL culturas de bactérias induzidas para expressão heteróloga dos diferentes fragmentos da região C-terminal de LmxM.34.3070 foram centrifugadas e o *pellet* suspenso em tampão de Laemmli (Beta-mercaptoetanol 0,7 M, Glicerol 50%, (SDS) 10%, Tris-HCl 0,25 M, pH= 8, azul de bromofenol 0,25 M, concentração final 5X) que contém agentes desnaturantes. As amostras foram desnaturadas a 90 °C por 10 minutos e posteriormente aplicadas nas canaletas do gel. O fracionamento das proteínas se deu pela aplicação de uma corrente constante de 30mA por gel em tampão Tris-Glicina (Tris-base 0,025 M, pH=8,6, Glicina 200 mM, SDS 1%).

Após a corrida, os géis contendo as proteínas fracionadas foram corados por azul de Coomassie, e em seguida descorados por solução descorante (Ácido Acético Glacial 5%, Etanol 10%) para identificação das bandas correspondentes aos diferentes fragmentos clonados.

Foi utilizado o fotodocumentador ChemiDoc Imaging Systems (BIO-RAD) para registrar os resultados desses experimentos.

Cepa	Características e efeitos no fenótipo
BL21(DE3)	Transcrição do clone de interesse pela RNA polimerase de T7. A transcrição está sob controle de <i>lacI</i> , ela é reprimida até a indução por IPTG, gene da T7 RNA polimerase <i>upstream</i> a um operador <i>lac</i> .
BL21(DE3) pLYSs	O plasmídeo pLysS codifica a lisozima do fago T7, um inibidor da polimerase T7, o que reduz e quase elimina a expressão basal a partir do promotor T7 na ausência do indutor.
<i>OverExpression</i> BL21 (DE3) C43	As cepas <i>OverExpress</i> contêm mutações genéticas para conferir tolerância a proteínas tóxicas.
BL21(DE3) RIL	Expressão de proteínas contendo códons raros de Arginina, Isoleucina e Leucina
BL21(DE3) RIP-L	Expressão de proteínas contendo códons raros de Arginina, Isoleucina, Prolina e Leucina

**Tabela 3:** Fenótipos das cepas utilizadas nos testes de expressão dos diferentes clones obtidos neste projeto. As células foram principalmente modificadas para não terem atividade endonuclease, RNase e protease o que poderia comprometer a produção de proteínas recombinantes por expressão heteróloga. Outra propriedade interessante é a capacidade das cepas códon plus expressarem tRNAs de códons raros presentes na composição gênica de outros organismos, o que aumenta a probabilidade de sucesso na produção de proteínas recombinantes por expressão heteróloga em bactéria (Casali, 2003).

### 3.17. Western blot Anti-His

Amostras de culturas induzidas foram centrifugadas e lisadas em tampão de Laemmli. As proteínas foram fracionadas em géis por SDS-PAGE e posteriormente transferidas para membrana de nitrocelulose ativadas pelo tampão de *blotting* (Tris-base 20 mM pH = 8, Glicina 150 mM, SDS 0,03%, Metanol 20%) através do sistema de transferência semi-seco Amersham (45 mA por cassete de transferência/1,5 h).

O objetivo dos experimentos de western blot foi validar a expressão dos diferentes fragmentos da região C-terminal de LmxM.34.3070 recombinantes que foram clonados através do reconhecimento da cauda de histidinas por interação imuno específica com anticorpo Anti-His (Booth et al., 2018). Foi aplicada uma solução corante de Ponceau para confirmar a transferência das proteínas fracionadas do gel para a membrana de nitrocelulose. Após a confirmação da transferência, a membrana foi lavada com TBS (Tris 20 mM, NaCl 150 mM) e posteriormente bloqueada com leite (Molico desnatado) a 5% em TBS *overnight* a 4°C sob agitação lenta.

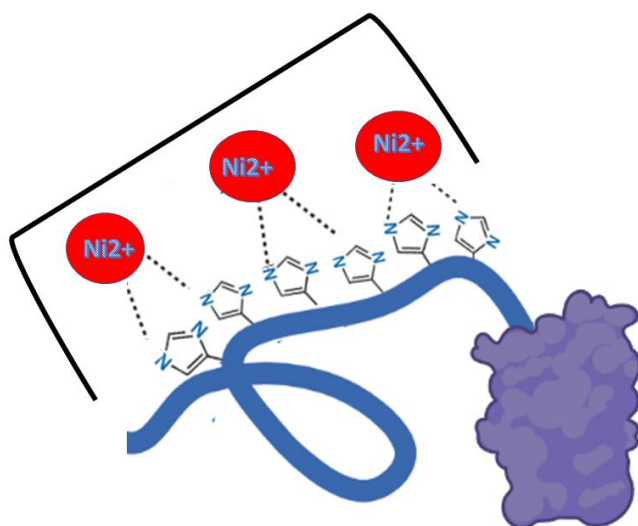
Após o bloqueio, a membrana passou por vários ciclos de lavagens com TBS e foi incubada com o anticorpo monoclonal Anti-His *tag* conjugado à fosfatase alcalina (Invitrogen) numa diluição de 1:4000 em PBS contendo 1% de albumina sérica bovina (BSA). A incubação foi realizada por 1 hora, à temperatura ambiente no agitador lento. Após a incubação, a membrana passou por lavagens com PBS e a revelação das bandas de expressão foi feita pela adição de BCIP-NBT a 37°C. O registro desses experimentos foi realizado através do fotodocumentador ChemiDoc Imaging Systems (BIO-RAD).

### 3.18. Cromatografia de Afinidade

O princípio da técnica de separação de proteínas com cauda de histidinas na coluna de afinidade baseia-se na interação do grupo imidazol dos resíduos de histidina com os cátions

de níquel immobilizados na resina, formando complexos de coordenação que retém o alvo (Block et al., 2009) (Waugh, 2005) (Figura 5). A eluição das proteínas retidas pode ser alcançada aumentando-se a concentração de imidazol na fase móvel da cromatografia de afinidade, através de um tampão de eluição B (foram usadas as concentrações de 250 e 500mM de imidazol na eluição), onde o imidazol em alta concentração irá competir e deslocar as histidinas dos sítios de coordenação com níquel (Block et al., 2009) (Waugh, 2005).

Os experimentos de FPLC (*Fast Protein Liquid Chromatography*) foram realizados no equipamento ÄKTAFPLC UPC-900 (Amersham Biosciences) para purificar os diferentes fragmentos da região C-terminal de LmxM.34.3070 recombinante. As células induzidas foram coletadas e lisadas por diferentes métodos de lise, em diferentes soluções de lise (ver Tabelas [4.1](#), [4.2](#) e [4.3](#)). O lisado foi clarificado por centrifugação a 35000 x g à 4°C por 1.5 hora. O precipitado foi descartado e o sobrenadante foi aplicado em coluna de níquel-HiTrap (GE Healthcare) previamente carregada com uma solução 0.1M de sulfato de níquel e equilibrada com Tampão de Ligação A. A eluição dos diferentes fragmentos da região C-terminal de LmxM.34.3070 recombinantes obtidos nesse trabalho foi realizada através de um gradiente linear de imidazol, por meio da alteração das concentrações entre o Tampão de Ligação A e o Tampão de Eluição B na fase móvel da cromatografia de afinidade, com fluxo de 2 ml/minuto em 12 volumes de coluna (0 a 100% de B em 60ml). Diferentes composições dos tampões A e B também foram testadas e encontram-se resumidas na tabela 4.



**Figura 5: Mecanismo de captura em HiTrap:** Uma sequência linear de 6-12 histidinas pode ancorar a proteína recombinante em uma superfície immobilizada com cátions de níquel (bolas vermelhas), que são hexavalentes e possuem dois sítios de coordenação com átomos de nitrogênio do grupo imidazol das histidinas.

Experimento	Construção testada	[IPTG]	Temperatura e tempo de indução	Lise	Tampão de Lise	Tampão de Afinidade A	Tampão de Afinidade B	Solubilidade
Teste 1	<b>LmxM.34.3070</b> <b>652-822(pET28a)</b>	600 μM	37°C / 3hrs	Sonicação	Tris 50 mM pH= 8 NaCl 200 mM PMSF 1mM	Imidazol 10mM NaPi 50mM pH 7.2 NaCl 300 mM B-mercaptoetanol 1mM	Imidazol 0.5M NaPi 50mM pH 7.2, NaCl 300 mM, B-mercaptoetanol 1mM	Insolúvel
Teste 2	<b>LmxM.34.3070</b> <b>652-822(pET28a)</b>	300 μM	18°C / overnight	French Press	Imidazol 20mM NaPi 50mM pH 7.2 NaCl 300 mM B-mercaptoetanol 1mM PMSF 1mM	Imidazol 20mM NaPi 50mM pH 7.2 NaCl 300 mM B-mercaptoetanol 1mM	Imidazol 0.5M, NaPi 50mM pH 7.2, NaCl 300 mM, B-mercaptoetanol 1mM	Parcialmente Solúvel
Teste 3	<b>LmxM.34.3070</b> <b>652-822(pET28a)</b>	300 μM	18°C / overnight	French Press	Tris 50 mM pH=8 Imidazol 30 mM Sacarose 10% NaCl 100 mM B-mercaptoetanol 1 mM 0.01% TWEEN 20 PMSF 1mM 0.5 mg/mL lisozima	Imidazol 30mM NaPi 50mM pH 7.2 NaCl 400 mM B-mercaptoetanol 1mM	Imidazol 250mM NaPi 50mM pH 7.2 NaCl 400 mM B-mercaptoetanol 1mM	Parcialmente Solúvel
Teste 4	<b>LmxM.34.3070</b> <b>652-822(pET28a)</b>	300 μM	16°C / 20 hrs	French Press	Tris 50 mM pH= 8 Imidazol 30 mM NaCl 100 mM Sacarose 10% B-mercaptoetanol 1mM L- Arginina 50 mM L- Glutamina 50mM 0.03% TWEEN 20 0.03% Triton-X PMSF 1mM 0.5 mg/mL lisozima	Imidazol 30mM NaPi 50mM pH 7.2 NaCl 400 mM B-mercaptoetanol 1 mM	Imidazol 0.25 M NaPi 50mM pH 7.2 NaCl 400 mM B-mercaptoetanol 1mM	Parcialmente Solúvel

**Tabela 4.1:** Testes de condições de expressão (concentração de IPTG e Temperatura), de métodos e soluções de lise, e de tampões A e B de cromatografia de afinidade para os clones selecionados para expressão em larga escala.

Experimento	Construção testada	[IPTG]	Temperatura e tempo de indução	Lise	Tampão de Lise	Tampão de Afinidade A	Tampão de Afinidade B	Solubilidade
Teste 5	<b>LmxM.34.3070</b> 638-822(pET28a)	300 $\mu$ M	16°C / 20 hrs	French Press	Tris 50 mM pH= 8 Imidazol 30 mM NaCl 100 mM Sacarose 10% B-mercaptoetanol 1mM L- Arginina 50 mM L-Glutamina 50mM 0.03% TWEEN 20 0.03% Triton-X PMSF 1mM 0.5 mg/mL lisozima	Imidazol 30mM  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1mM	Imidazol 0.25M  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1mM	Muito solúvel
Teste 6	<b>LmxM.34.3070</b> 669-822(pET28a)	300 $\mu$ M	16°C / 20 hrs	French Press	Tris 50 mM pH= 8 Imidazol 30 mM NaCl 100 mM Sacarose 10% B-mercaptoetanol 1mM L- Arginina 50 mM L-Glutamina 50 mM 0.03% TWEEN 20 0.03% Triton-X PMSF 1mM 0.5 mg/mL lisozima	Imidazol 30mM  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1 mM	Imidazol 0.25M  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1mM	Sem pico no cromatograma de afinidade

**Tabela 4.2:** Testes de condições de expressão (concentração de IPTG e Temperatura), de métodos e soluções de lise, e de tampões A e B de cromatografia de afinidade para os clones selecionados para expressão em larga escala.



**Tabela 4.3:** Testes de condições de expressão (concentração de IPTG e Temperatura), de métodos e soluções de lise, e de tampões A e B de cromatografia de afinidade para os clones selecionados para expressão em larga escala.

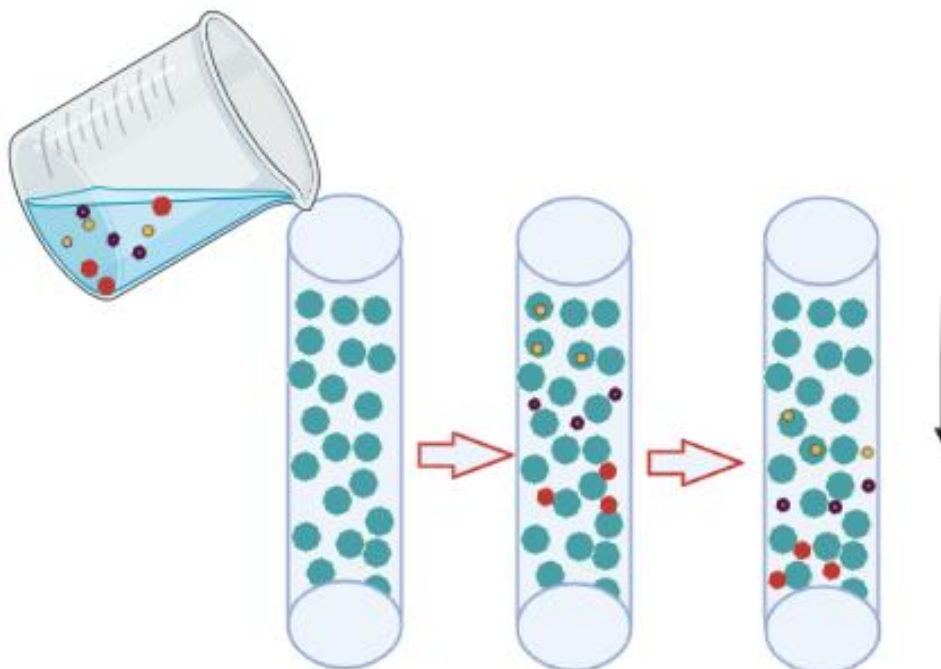
Experimento	Construção testada	[IPTG]	Temperatura e tempo de indução	Lise	Tampão de Lise	Tampão de Afinidade A	Tampão de Afinidade B	Solubilidade
Teste 7	<b>LmxM.34.3070</b> <b>669-822(pET28a)</b>	300 $\mu$ M	16°C / 20 hrs	French Press	Tris 50 mM pH= 8 Imidazol 30 mM NaCl 100 mM Sacarose 10% B-mercaptoetanol 1mM L- Arginina 50 mM L-Glutamina 50 mM 0.03% TWEEN 20 0.03% Triton-X PMSF 1mM 0.5 mg/mL lisozima	Imidazol 30mM  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1 mM	Imidazol 0.5M  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1mM	Sem pico no cromatograma de afinidade
Teste 8	<b>LmxM.34.3070</b> <b>669-822(pET28a)</b>	300 $\mu$ M	16°C / 20 hrs	French Press	Tris 50 mM pH= 8 Imidazol 30 mM NaCl 100 mM Sacarose 10% B-mercaptoetanol 1mM L- Arginina 50 mM L-Glutamina 50 mM 0.03% TWEEN 20 0.03% Triton-X PMSF 1mM 0.5 mg/mL lisozima	Imidazol 30mM  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1 mM	Imidazol 0.25M  NaPi 50mM pH 7.2  NaCl 400 mM  B-mercaptoetanol 1mM	Sem pico no cromatograma de afinidade

### 3.19. Cromatografia por exclusão molecular

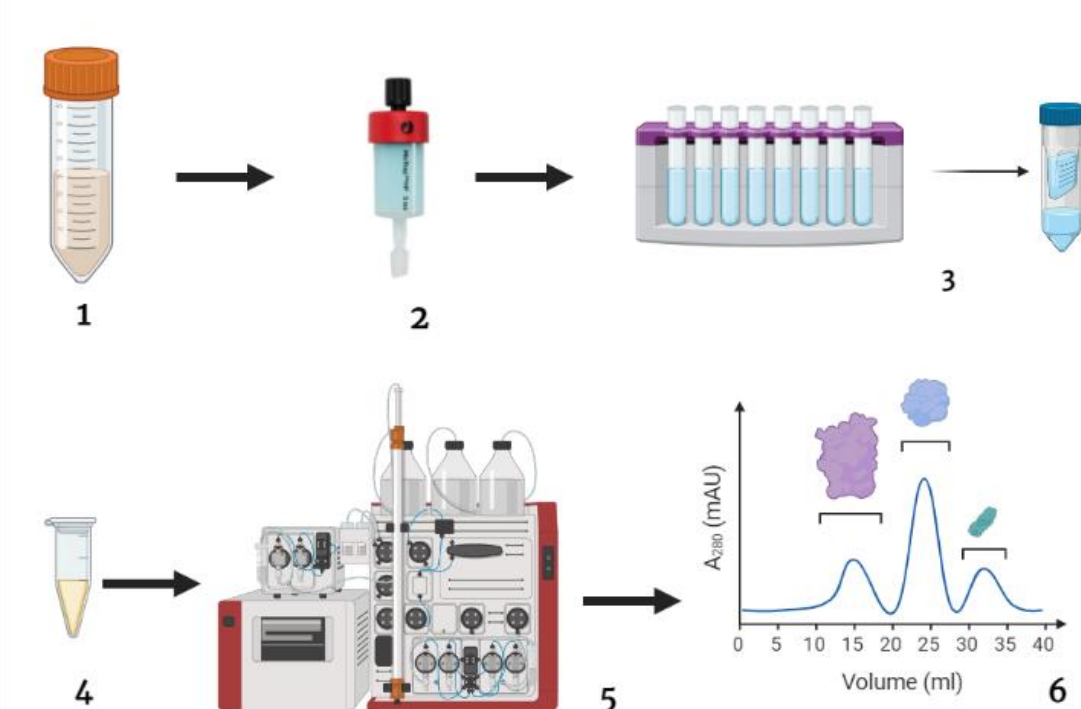
O grau de pureza das amostras purificadas por cromatografia de afinidade foi constantemente avaliado por SDS-PAGE, onde verificamos a purificação parcial dos diferentes fragmentos recombinantes da região C-terminal de LmxM.34.3070. Para os estudos de biologia estrutural, amostras com maior nível de polimento são mais adequadas e para isto, recorreremos à cromatografia por exclusão molecular.

A cromatografia de exclusão molecular, também conhecida como filtração em gel, é uma técnica de separação de proteínas baseada em seu tamanho (Hagel, 2011). A fase estacionária na coluna é composta por polímeros unidos por ligações cruzadas que formam poros de diferentes tamanhos e espessuras. Proteínas maiores são incapazes de penetrá-los, restando-lhes um menor volume para ser percorrido e, portanto, são as primeiras a serem eluídas (Hagel, 2011). Já as proteínas menores atravessam esses poros, percorrendo um maior volume e levando maior tempo até a sua completa eluição (Hagel, 2011) (Figura 6).

A etapa de purificação por exclusão molecular foi otimizada com base no melhor tampão e fluxo de eluição (ver Tabela 5) em coluna preparativa Superdex S75 16/600. As frações de interesse que eluíram nas cromatografias de afinidade realizadas em condições otimizadas foram concentradas em 500µl com membranas Amicon® Ultra 0.5 (*cutoff* 10 kDa) e aplicadas na coluna (Figura 7). As corridas cromatográficas foram realizadas isocraticamente em um volume de coluna (120ml) no equipamento ÄKTAFPLC UPC-900.



**Figura 6: Mecanismo de separação por exclusão molecular:** Proteínas maiores (vermelhas) migram mais rapidamente. Proteínas menores (amarelas) demoram mais para serem eluídas por penetrarem nos poros formados pelos polímeros interligados por pontes cruzadas. A seta preta indica a direção do fluxo da fase móvel.



**Figura 7:** Esquema geral do protocolo de purificação dos diferentes fragmentos recombinantes da região C-terminal de LmxM.34.3070 em 6 etapas: **1:** clarificação do lisado e coleta de sobrenadante. **2:** injeção do sobrenadante em coluna de afinidade Ni-Hitrap. **3-** eluição do fragmento por gradiente linear de imidazol no equipamento AKTA e concentração das frações de interesse em membranas Amicon para um VF= 500 $\mu$ l. **4-** Injeção de amostra em na coluna Superdex S75 16/600. **5-** Eluição isocrática por FPLC no equipamento AKTA. **6-** Análise do perfil cromatográfico e avaliação de rendimento.

**Tabela 5:** Testes de condições de purificação de diferentes construções por exclusão molecular. Os resultados foram avaliados a qualidade do cromatograma, rendimento e formação de oligômeros.

Experimentos	Origem da amostra	Construção	Fluxo	Composição do tampão de exclusão molecular
GF1	Frações de interesse da cromatografia de afinidade no teste 3	<b>LmxM.34.3070</b> <b>652-822(pET28a)</b> (expresso em RIP-L)	1ml/min	Tris 50mM pH=8, NaCl 200mM, Glicerol 1% , B-mercaptoetanol 1mM, EDTA 1mM,
GF2	Frações de interesse da cromatografia de afinidade no teste 4	<b>LmxM.34.3070</b> <b>652-822(pET28a)</b> (expresso em RIP-L)	1ml/min	Tris 50mM pH=8, NaCl 200mM, Glicerol 1% , B-mercaptoetanol 1mM, EDTA 1mM,
GF3	Frações de interesse da cromatografia de afinidade no teste 5	<b>LmxM.34.3070</b> <b>638-822(pET28a)</b> (expresso em RIP-L)	1ml/min	Tris 50mM pH=8, NaCl 200mM, Glicerol 1% , B-mercaptoetanol 1mM, EDTA 1mM,
GF4	Purificado por Cromatografia de afinidade (pH=7.2) <b>Tampão A:</b> NaPi 50 mM, NaCl 400 mM, Imidazol 30mM e B-mercaptoetanol 1mM. <b>Tampão B:</b> NaPi 50mM, NaCl 400mM, Imidazol 0.25M e B-mercaptoetanol 1mM.	<b>LmxM.34.3070</b> <b>638-822(pET28a)</b> (expresso em RIP-L)	1.5 ml/min	Tris 50 mM pH=8 NaCl 400 mM, Glicerol 1%, B-mercaptoetanol 10mM EDTA 5mM,

### 3.20. Método do coeficiente de extinção molar

A estimativa das concentrações dos fragmentos recombinantes puros obtidos após cada etapa de purificação por exclusão molecular foi realizada por espectrofotometria usando a Lei de Lambert-Beer, dada pela equação:

$$\text{Log} \left( \frac{I_0}{I} \right) = \epsilon \cdot b \cdot c, \quad \text{Equação 2}$$

onde  $\epsilon$  é o coeficiente de extinção molar, uma propriedade característica da espécie química submetida à absorção da luz incidente,  $I_0$  intensidade de luz incidente,  $I$  a intensidade de luz irradiada, “b” a espessura de um meio transparente e “c” a concentração da amostra.

Cadeias laterais de resíduos de aminoácidos contendo aromáticos como fenilalanina, tirosina e triptofano, atribuem às proteínas a capacidade de absorver luz no comprimento de onda de 280 nm (Wilkins et al., 1999). O coeficiente de extinção molar teórico dos fragmentos obtidos foi calculado com base em suas sequências de aminoácidos pelo servidor ProtParam ([ExPASy - ProtParam tool](#)) (Wilkins et al., 1999) e foi estimado em 37.47 mM<sup>-1</sup>cm<sup>-1</sup> para todos os fragmentos de LmxM.34.3070 testados neste trabalho.

Os cálculos das concentrações pelo método de coeficiente de extinção molar em cada fase de testes foram essenciais para avaliar a eficácia dos métodos empregados desde a lise até a exclusão molecular, uma vez que esses valores de concentrações refletiam o rendimento das preparações de amostras candidatas a experimentos de biologia estrutural, que exigem altas quantidades (miligramas) do alvo a ser estudado.

### 3.21. Espalhamento de luz a múltiplos ângulos acoplado à cromatografia por exclusão molecular (SEC-MALS)

A verificação da massa bem como do estado de oligomerização do fragmento estudado foi realizada por Espalhamento de Luz a múltiplos ângulos acoplado a uma coluna de separação por Exclusão Molecular (SEC-MALS) no equipamento MALS Wyatt Technology.

Foram aplicados 100µL das frações de interesse em uma coluna analítica de gel filtração Superdex S200 10/300 com um fluxo de 1ml/min para separação de oligômeros por exclusão molecular. Os valores de absorbância durante a eluição foram medidos por um detector de UV na faixa de 280 nm.

Paralelamente à detecção de espécies absorventes de UV, sinais de dispersão de luz em múltiplos ângulos obtidos através da aplicação de um Raio Laser incidente foram registrados por um detector de MALS (*Multi-Angle Light Scattering*) (Some et al., 2019) para a quantificação analítica das massas das espécies presentes nessas amostras. A relação entre a massa das espécies que eluíram da coluna e a sua dispersão multiangular de raio laser foi processada através da equação:

$$M = \frac{R}{Kc \left( \frac{dn}{dc} \right)^2}, \quad \text{Equação 3}$$

onde  $M$  é a massa do analito,  $R$  a quantidade de raio laser espalhada medida no detector de MALS,  $c$  a concentração de proteína calculada pela absorbância a 280 nm,  $K$  a constante ótica do sistema, e  $dn/dc$  a normalização da medida obtida pela razão entre as diferenças de difração do branco e da amostra (Some et al., 2019).

### 3.22. Espectroscopia de dicroísmo circular (CD)

Proteínas são macromoléculas quirais formadas por resíduos de aminoácidos e por isso apresentam atividade óptica, podendo ser examinadas por dicroísmo circular (Kelly et al., 2005). O dicroísmo circular refere-se à absorção diferencial dos vetores L e R que compõem o plano da luz circularmente polarizada incidente sobre um cromóforo (Greenfield, 2007). A absorção diferencial dos vetores L e R resulta em na rotação do plano da luz polarizada, reportada neste experimento através dos valores do ângulo  $\theta$ , onde:

$$\theta = \arctan\left(\frac{b}{a}\right), \quad \text{Equação 4}$$

em que a e b referem-se às dimensões dos eixos do plano da elipse que resulta da absorção diferencial de L e R.

Amostras de frações contendo apenas monômeros do fragmento recombinante mais solúvel obtidas por exclusão molecular foram preparadas e submetidas a experimentos de dicroísmo circular. A concentração de proteína nesta análise foi de 12  $\mu\text{M}$ . As medidas foram obtidas na faixa do UV distante, e a elipticidade resultante de 6 acumulações, caminho óptico de 1mm em cubeta de quartzo, velocidade 20 nm/min e temperatura 25 °C no polarímetro J815 (Jasco) da Central Analítica do Instituto de Química da USP. Os dados obtidos em *miligrau* foram convertidos em elipticidade molar pela equação:

$$\theta m = \frac{\theta \cdot MW \cdot 100}{c \cdot l \cdot n}, \quad \text{Equação 5}$$

onde  $\theta m$  é a elipticidade molar,  $\theta$  elipticidade em miligraus,  $MW$  a/ massa molecular da proteína em kDa,  $c$  a concentração das amostras em mg/ml,  $l$  o caminho ótico em centímetros e  $n$  o número de resíduos da proteína (Kelly et al., 2005) (Greenfield, 2007). O conteúdo de estrutura secundária do domínio FABP-like recombinante foi examinado por deconvolução espectral usando DichroWeb (Miles et al., 2022), que reconstruiu os dados

experimentais através do algoritmo Continn e utilizando o modelo SMP180t (Abdul-Gader et al., 2011) de pequenas proteínas globulares cujas estruturas experimentais são conhecidas.

### 3.23. Desnaturação térmica

Para os experimentos de desnaturação foram realizados em diferentes condições:

- 1) LeiFABP purificado em **tampão Tris-HCl** 50 mM pH= 8, 200 mM de NaCl, 2 mM B-Mercaptoetanol, 1% de Glicerol, 2 mM de EDTA;
- 2) LeiFABP purificado em **tampão acetato**: Acetato de sódio 45 mM, Ácido acético glacial 5 mM pH= 5.5, 200 mM de NaCl, 2 mM B-Mercaptoetanol, 1% de Glicerol, 2 mM de EDTA
- 3) LeiFABP purificado em **tampão PBS**: fosfato de sódio monobásico 36 mM, fosfato de sódio dibásico 14 mM pH 7,2, pH= 7.2, 200 mM de NaCl, 2 mM B-Mercaptoetanol, 1% de Glicerol, 2 mM de EDTA.

As amostras foram preparadas exatamente como descrito na seção 3.21, entretanto, a curva de desnaturação foi obtida através da média dos valores de  $\theta$  de 3 acumulações registrados em 213 nm ou 220 nm, entre 25 e 90 °C, onde os registros de  $\theta$  foram efetuados em passos de 1 °C. Como o domínio C-terminal de LmxM.34.3070 foi predito para possuir uma topologia em barril-beta FABP-like, o comprimento de onda foi fixado em 213 e 220 nm (Greenfield, 2007).

O ajuste dos dados experimentais de desnaturação térmica foi realizado através de um *script* interno elaborado pela Dra. Angy Liseth Davalos que se baseia no ajuste dos dados experimentais à equação abaixo que assume a lei de distribuição de Boltzman para um sistema de dois estados (nativo e desnovelado), cuja diferença de energia é equivalente ao  $\Delta G$  associado à desnaturação térmica:



$$S_{\lambda}(T) = S_N \left( \frac{1}{1 + \exp\left\{\frac{-\Delta H}{RT} \left(1 - \frac{T}{T_m}\right)\right\}} \right) + S_D \left( \frac{\exp\left\{\frac{-\Delta H}{RT} \left(1 - \frac{T}{T_m}\right)\right\}}{1 + \exp\left\{\frac{-\Delta H}{RT} \left(1 - \frac{T}{T_m}\right)\right\}} \right) \quad (\text{Equação 6})$$

onde a  $S(T)$  é o sinal total medido num comprimento de onda fixo ( $\lambda$ ) em função da temperatura, que é equivalente à soma da contribuição do sinal da proteína na forma enovelada  $S(N)$  com o sinal da forma desnovelada e  $S(D)$  multiplicados pelas respectivas probabilidades.  $T$  a temperatura em Kelvin,  $R$  é a constante dos gases. A equação acima descreve a curva de desnaturação obtida experimentalmente através de um ajuste não linear dos dados assumindo a temperatura de *melting* ( $T_m$ ), a variação de entalpia ( $\Delta H$ ),  $S_N$  e  $S_D$  como parâmetros livres.

### 3.24. Espectroscopia de Ressonância Magnética Nuclear

Amostras do fragmento C-terminal de LmxM.34.3070 mais solúvel foram preparadas em tampão de RMN: fosfato de sódio monobásico 36 mM, fosfato de sódio dibásico 14 mM pH 7,2, 400 mM de NaCl, 10 mM de beta-mercaptoetanol, 5 mM de EDTA e transferidas para tubos de RMN de 5mm (Sigma). Os dados de RMN foram coletados no espectrômetro Bruker da Central Analítica do Instituto de Química da USP, operando em 800 MHz e processados usando o TopSpin (Bruker-Biospin). A concentração da amostra foi 202  $\mu$ M de proteína. A temperatura de aquisição foi 25 °C. O espectro foi adquirido com supressão do sinal da água por *Watergate* usando a sequência p3919, acumulação de 64 transientes, janela espectral de 16025 Hz, e aquisição de 3202 pontos. O espectro foi multiplicado por uma função  $\cos^2$  e preenchido com zeros até 65536 pontos antes da correção de fases.

---

## **4. Resultados e Discussão**

---

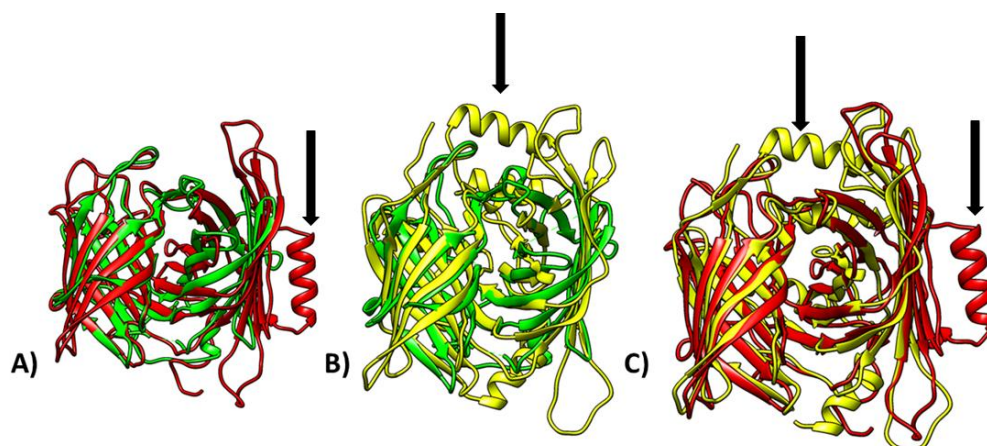
#### 4.1. Os modelos preditos por Raptor-X e ColabFold são semelhantes aos modelos gerados por PHYRE2

Os modelos dos domínios N e C-terminal de LmxM.34.3070 preditos com Raptor-X apresentaram semelhanças com os modelos de homologia sugeridos pelo PHYRE2 para as regiões N-terminal (52-488) e C-terminal (665-809) de LmxM.34.3070. Para a região N-terminal, Raptor-X gerou modelos apresentando dois barris-beta formados por fitas-beta antiparalelas. Todavia,  $\alpha$ -hélices foram modeladas por Raptor-X (seta preta na figura 8A), as quais não estão presentes na estrutura da proteína *AttH-lipocalin like* de *Nitrosomonas europaea* (PDB 2ICH), escolhida como molde para esta região por PHYRE2. Considerando a região de estrutura beta, observou-se uma semelhança com a proteína *AttH lipocalin-like* que é composta por dois motivos *lipocalin-like* duplicados, cada qual consistindo em 10 e 12 fitas-beta antiparalelas. O RMSD obtido do melhor alinhamento entre a estrutura predita para a região N-terminal de LmxM.34.3070 por Raptor-X e a estrutura cristalográfica da proteína *AttH-lipocalin like* de *N. europaea* (PDB 2ICH) é de 1.233 Angstroms (102 pares de resíduos) (Figura 8A). Para a região C-terminal compreendida pelos aminoácidos 665-809, o Raptor-X modelou um barril-beta composto por 10 fitas beta antiparalelas e o motivo hélice-alça-hélice na alça entre as fitas  $\beta$ A e  $\beta$ B, assim como sugerido por PHYRE2 (Figura 9). O RMSD entre as estruturas preditas por Raptor-X e PHYRE2 para o domínio C-terminal (aminoácidos 665-809) de LmxM.34.3070 foi de 1.035 Angstroms sobre 70 pares de resíduos (Figura 9A). É importante mencionar que a modelagem de um barril-beta na região C-terminal de LmxM.34.3070 por Raptor-X só foi possível quando a predição foi realizada utilizando apenas o trecho de aminoácidos 665-809 de LmxM.34.3070. Quando a predição foi realizada utilizando-se como *input* a sequência de amino ácidos completa de LmxM.34.3070, a modelagem de um domínio C-terminal em forma de barril não foi obtida, uma vez que o algoritmo apenas identificou uma folha-beta antiparalela (Figura 10).

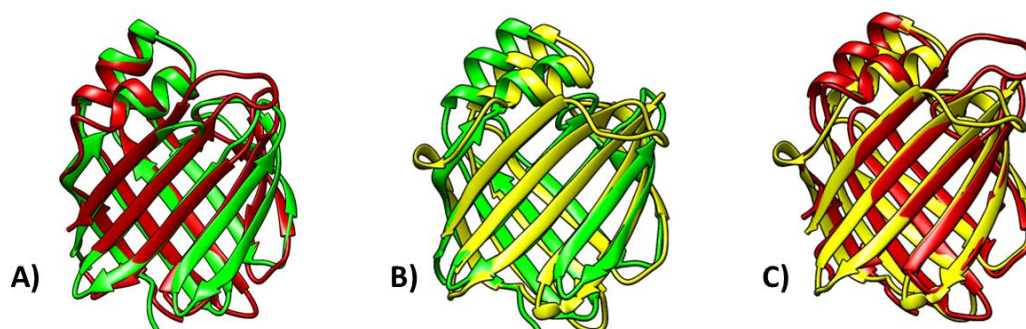
Provavelmente, isso se deve à baixa qualidade do alinhamento na região C-terminal quando a sequência de aminoácidos completa de *LmxM.34.3070* é utilizada para a busca.

A plataforma ColabFold (Jumper et al., 2021) também foi utilizada para prever a estrutura da proteína codificada pelo gene *LmxM.34.3070*. Similarmente à análise realizada em Raptor-X, para a região N-terminal, o programa sugeriu um domínio composto por dois barris-beta achatados formando uma estrutura similar à da proteína *AttH lipocalin-like* de *N. europaea* (PDB 2ICH). Assim como em Raptor-X, Colab Fold também modelou hélices (seta preta na figura 8B) não encontradas na proteína *AttH-lipocalin like*, que foi sugerida como o melhor modelo de homologia para esta região por PHYRE2 (Figura 8). O RMSD entre os modelos preditos para esta região por PHYRE2 e ColabFold foi de 1.067 Angstroms quando 126 pares de resíduos foram comparados (Figura 8A). Para a região C-terminal, o modelo gerado por ColabFold consistiu em um barril-beta composto de 10 fitas antiparalelas e duas alfa-hélices na alça inter-fita AB, e é consistente com as previsões por PHYRE2 uma vez que o RMSD entre ambas as estruturas foi de 1.031 Angstroms (alinhamento entre 70 pares de resíduos) (Figura 9B).

De modo geral, para a região C-terminal, as modelagens em PHYRE2, Raptor-X e Colab Fold convergiram para a mesma solução. Para a região N-terminal, as modelagens foram parcialmente congruentes, uma vez que hélices foram apontadas nas previsões realizadas com Raptor-X e Colab Fold mas que não estão presentes na proteína *AttH-lipocalin like* usada como *template* por PHYRE2. No entanto, essas previsões por ColabFold e RaptorX não coincidem uma com a outra (Figura 8C). A estrutura cristalográfica da proteína *AttH-lipocalin like* (PDB 2ICH) e os modelos de ColabFold e RaptorX são semelhantes **nos trechos** que compõem o barril-beta de fitas antipararelas (Figura 8).



**Figura 8:** Sobreposição par a par das estruturas geradas para a região 52-488 correspondente ao domínio N-terminal de LmxM.34.3070. Modelo PHYRE: em verde. Modelo Raptor-X: vermelho. Modelo ColabFold: Amarelo. Em todas as predições, os algoritmos inferiram que a região N-terminal forma um domínio formado por dois barris-beta de 10 e 12 fitas-beta antiparalelas. A) Sobreposição entre os modelos de PHYRE e Raptor-X. RMSD: 1.233 Angstroms (calculado sobre 102 pares de resíduos). B) Sobreposição entre os modelos de PHYRE e Colab Fold. RMSD: 1.067 Angstroms (calculado entre 126 pares de resíduos). C) Sobreposição entre os modelos de Raptor-X e ColabFold. RMSD: 1.240 Angstroms (calculado entre 190 pares de resíduos). O modelo gerado por PHYRE2 para a região 52-488 de Lmx.M.34.3070 contém os dois motivos *lipocalin-like*. No entanto, ColabFold e Raptor-X previram hélices em alças inter-fita que não foram preditas por PHYRE2 (setas pretas).

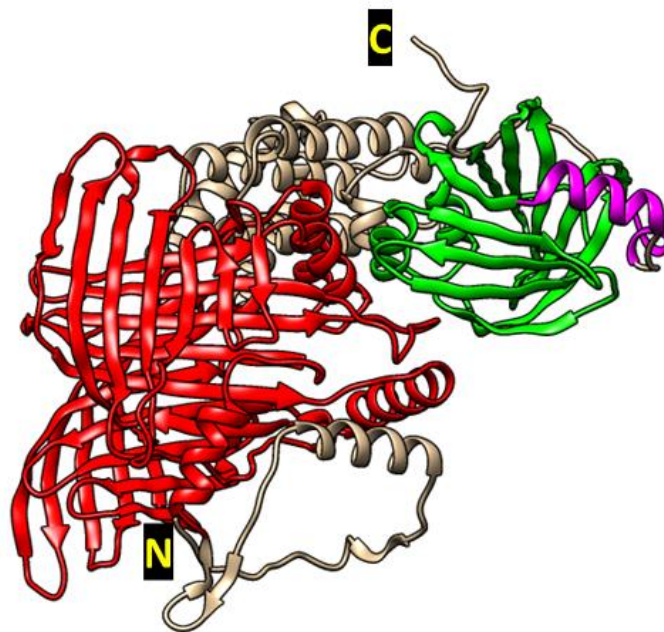


**Figura 9:** Sobreposição par a par dos modelos gerados para a região C-terminal (aminoácidos 665-809) de LmxM.34.3070. Modelo PHYRE: em verde. Modelo Raptor-X: vermelho. Modelo ColabFold: Amarelo. Todas as predições convergiram para a formação de um barril-beta de 10 fitas-beta antiparalelas e um motivo hélice-alça-hélice entre as fitas AB. A) Sobreposição entre os modelos de PHYRE e Raptor-X. RMSD: 1.035 Angstroms (calculado entre 70 pares de resíduos). B) Sobreposição entre os modelos de PHYRE e ColabFold. RMSD: 1.031 Angstroms (calculado entre 58 pares de resíduos). C) Sobreposição entre os modelos de Raptor-X e Colab Fold. RMSD: 1.266 Angstroms (calculado entre 76 pares de resíduos).



**Figura 10:** Predição de estrutura da proteína LmxM.34.3070 com Raptor-X. Observa-se a presença de um domínio N-terminal composto de dois barris-beta *AttH-like* (vermelho). Na região C-terminal não houve predição de um domínio FABP-like, apenas uma folha beta antiparalela foi predita (verde).

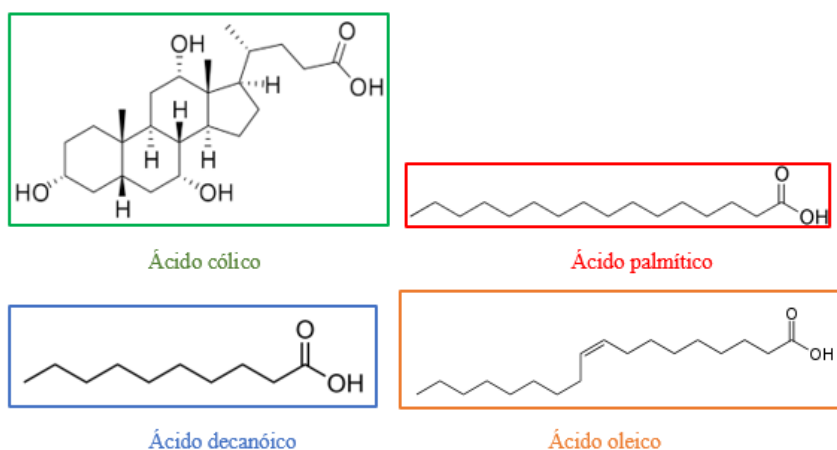
As FABPs são proteínas ligadoras de lipídeos e ácidos graxos formadas por um barril-beta de 10 fitas-beta antiparalelas com um motivo hélice-alça-hélice na alça entre as fitas  $\beta A$  e  $\beta B$ . A modelagem comparativa realizada por PHYRE2, bem como as predições usando *machine learning* em Raptor-X e ColabFold, convergiram na predição estrutural da região C-terminal de LmxM.34.3070, uma vez que em todas as análises verificou-se uma topologia FABP-like. Todavia, predições em ColabFold utilizando como *input* toda a sequência de LmxM.34.3070, encontraram  $\alpha$ -hélices adicionais na região N-terminal (Figura 11, em magenta) do domínio FABP-like de LmxM.34.3070, que não foram observadas para a grandes maiorias das FABPs descritas.



**Figura 11:** Predição de estrutura da proteína LmxM.34.3070 em ColabFold demonstrando a presença de um domínio N-terminal composto de dois barris-beta AttH-like (vermelho). Na região C-terminal, ColabFold previu domínio FABP-like, formado por 10 fitas beta antiparalelas e um motivo hélice-alça-hélice entre as fitas AB (verde). Adicionalmente, ColabFold previu duas novas  $\alpha$ -hélices na região N-terminal (magenta) deste barril FABP-like.

O modelo estrutural do domínio C-terminal FABP-like de LmxM.34.3070 obtido com ColabFold foi comparado, por alinhamento estrutural, a outras proteínas depositadas no PDB usando DALI (Holm, 2020). Dentre as nove estruturas mais bem classificadas, foram analisadas aquelas que possuíam ligantes de natureza hidrofóbica (Figura 12). Nesta busca foram encontradas proteínas com estruturas semelhantes ao domínio C-terminal FABP-like de LmxM.34.3070 e que estão associadas à ligação e ao transporte de moléculas hidrofóbicas em outros organismos. Mais especificamente: a **FABP** de *Manduca sexta* complexada ao ácido palmítico (PDB 1MDC); a proteína de ligação a ácido biliar de *Zebrafish* complexada ao ácido cólico (PDB 3ELZ); a luciferina 2-monoxigenase de ligação ao ácido decanóico (PDB 5IBO Holo, 7SNS Apo) e a **nemFABP** (As-p18) de ligação ao ácido oleico do helminto parasito *Ascaris suum* (PDB 6I9F) (Figura 12). Esse resultado fornece informações sobre possíveis ligantes do domínio FABP-like de LmxM.34.3070.

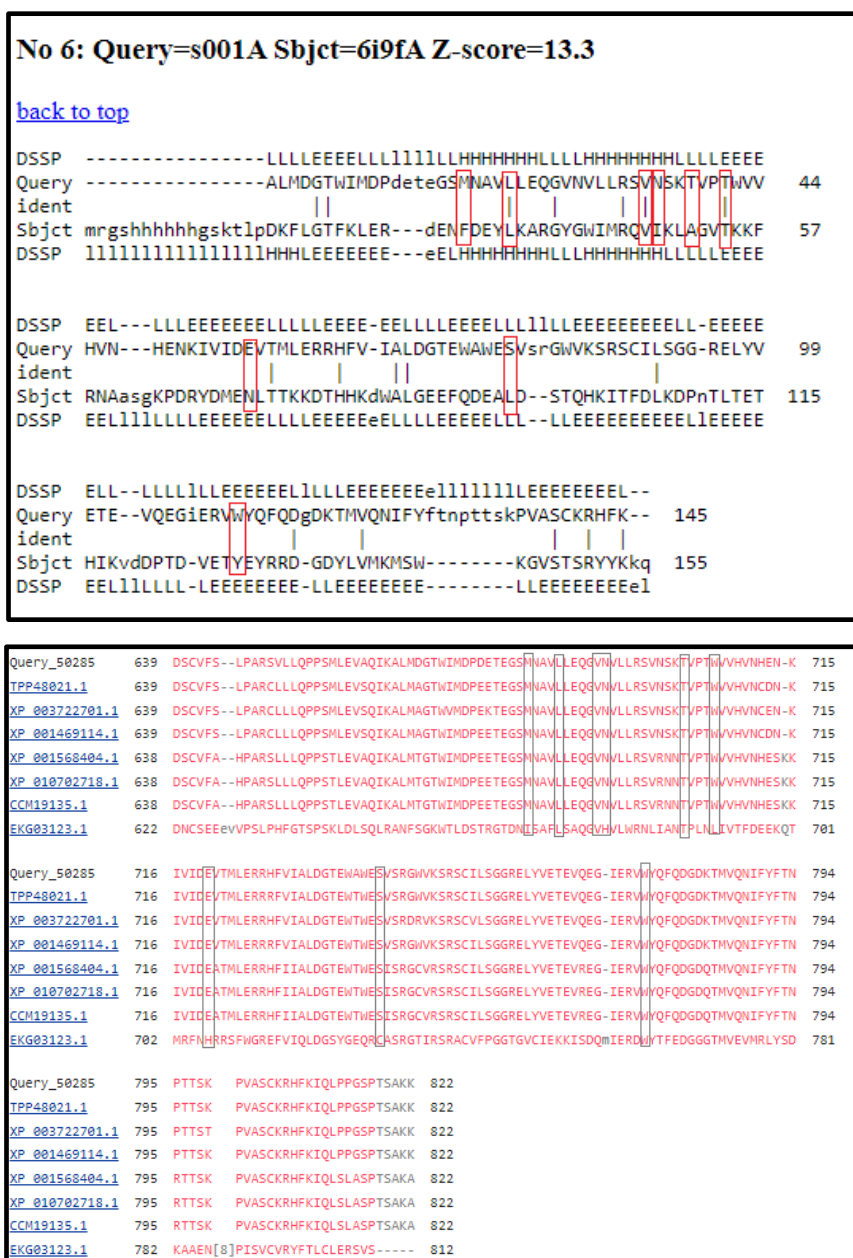
No:	Chain	Z	rmsd	lali	nres	%id	PDB	Description	
<input type="checkbox"/>	1:	1mdc-A	14.9	2.1	131	131	13	<a href="#">PDB</a>	MOLECULE: INSECT FATTY ACID BINDING PROTEIN;
<input type="checkbox"/>	2:	5z4g-A	13.2	2.4	133	145	11	<a href="#">PDB</a>	MOLECULE: SAHS4;
<input type="checkbox"/>	3:	7sns-C	12.8	2.3	132	172	14	<a href="#">PDB</a>	MOLECULE: OPLOPHORUS-LUCIFERIN 2-MONOOXYGENASE CATALYTIC SU
<input type="checkbox"/>	4:	6i9f-A	12.8	2.8	142	155	17	<a href="#">PDB</a>	MOLECULE: FATTY ACID-BINDING PROTEIN HOMOLOG;
<input type="checkbox"/>	5:	3elz-B	12.5	2.8	126	135	13	<a href="#">PDB</a>	MOLECULE: ILEAL BILE ACID-BINDING PROTEIN;
<input type="checkbox"/>	6:	7oxw-A	11.5	2.7	126	143	14	<a href="#">PDB</a>	MOLECULE: CELLULAR RETINOIC ACID-BINDING PROTEIN 2;
<input type="checkbox"/>	7:	1sa8-A	9.8	2.3	103	106	18	<a href="#">PDB</a>	MOLECULE: FATTY ACID-BINDING PROTEIN, INTESTINAL;
<input type="checkbox"/>	8:	7sfn-B	9.5	2.9	123	153	11	<a href="#">PDB</a>	MOLECULE: OLMO - OLIGOMYCIN SPIROCYCLASE;
<input type="checkbox"/>	9:	4tq2-A	9.4	3.6	128	177	8	<a href="#">PDB</a>	MOLECULE: PUTATIVE PHYCOERYTHRIN LYASE;



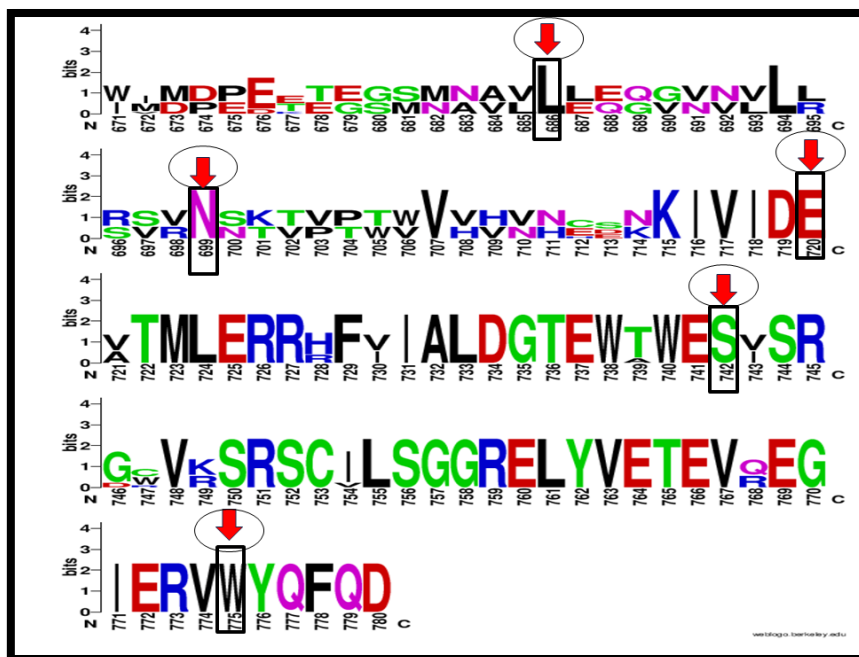
**Figura 12:** Resultados da busca comparativa no servidor DALI entre a estrutura predita do domínio FABP-like de *LmxM.34.3070* por ColabFold e todas as estruturas depositadas no PDB. 1MDC: FABP de *Manduca sexta* complexada a ácido palmítico; 3ELZ: proteína de ligação a ácido biliar de *Zebrafish* complexada à ácido cólico; 7SNS: luciferina 2-monoxigenase apo, mas com afinidade por ácido decanoico (PDB 5IBO); 6I9F: nem-FABP As-p18 de *Ascaris suum* complexada ao ácido oleico. As demais proteínas apontadas nesta busca compartilham semelhanças estruturais com as FABPs.

O alinhamento das sequências de aminoácidos de nemFABP de *Ascaris suum* (As-p18) com a sequência do domínio LeiFABP de *LmxM.34.3070* evidencia a conservação de resíduos envolvidos na interação desse domínio com a molécula de ácido oleico em *Leishmania* spp. e *T. cruzi* (Figura 13 e 14). Assim, pode-se sugerir que o domínio LeiFABP possui como potenciais ligantes ácidos graxos saturados e insaturados, encontrados abundantemente no lipídoma de *Leishmania*.





**Figura 13: Acima:** Alinhamento entre a sequência de amino ácidos do domínio LeiFABP e de As-p18 (6I9F). Os resíduos envolvidos na interação com o ácido oleico estão destacados em caixas vermelhas. **Abaixo:** Alinhamento da sequência de aminoácidos de LmxM.34.3070 e potenciais ortólogos em outros tripanosomatídeos. Os resíduos envolvidos na interação com o ácido oleico estão destacados em caixas cinzas, demonstrando a conservação desses resíduos em diferentes tripanosomatídeos: *L. donovani* (TPP48021.1), *L. major* (XP\_003722701.1), *L. infantum* (XP\_001469114.1), *L. braziliensis* (XP\_001568404.1), *L. panamensis* (XP\_010702718.1), *L. guyanensis* (CCM19135.1), *T. cruzi* (EKG03123.1).

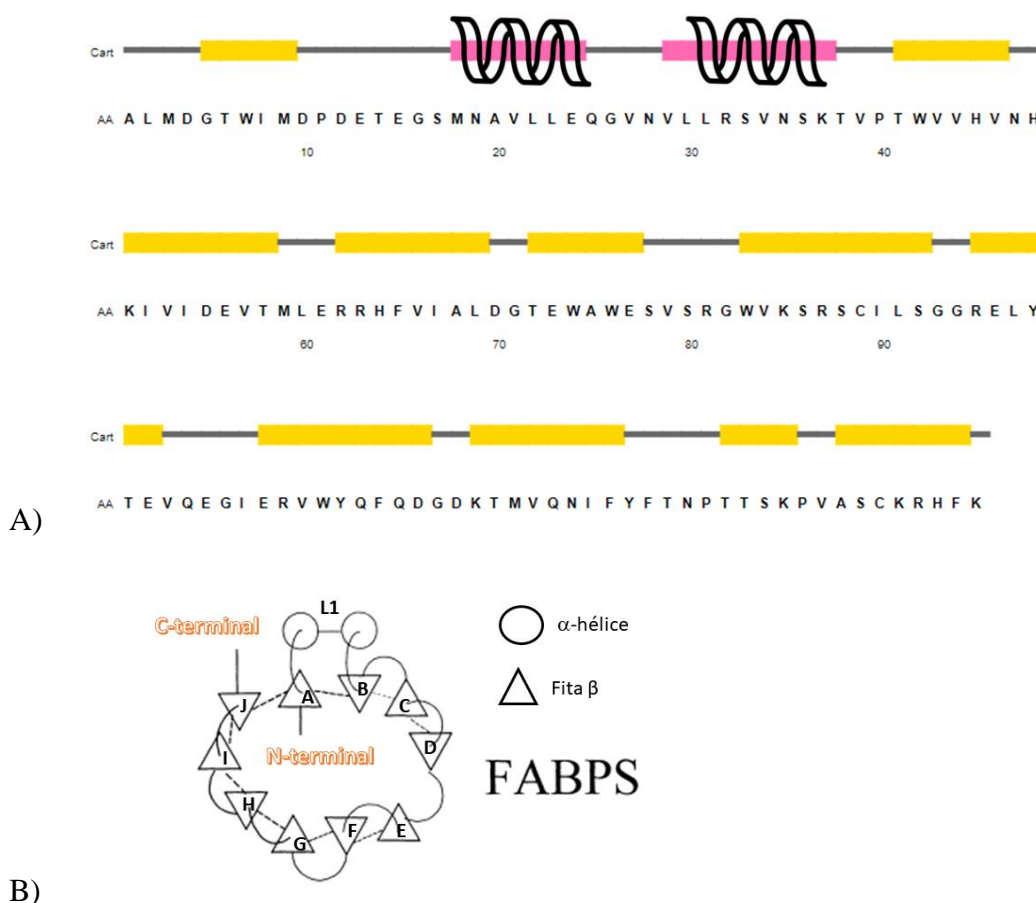


**Figura 14:** Representação gráfica em WebLogo (Crooks et al., 2004) do alinhamento múltiplo obtido entre a região C-terminal de LmxM.34.3070 em diferentes espécies de *Leishmania*. As setas vermelhas indicam a conservação de resíduos importantes para a ligação do ácido oleico como visto na proteína Asp-18 (estrutura 6I9F) no gênero *Leishmania*.

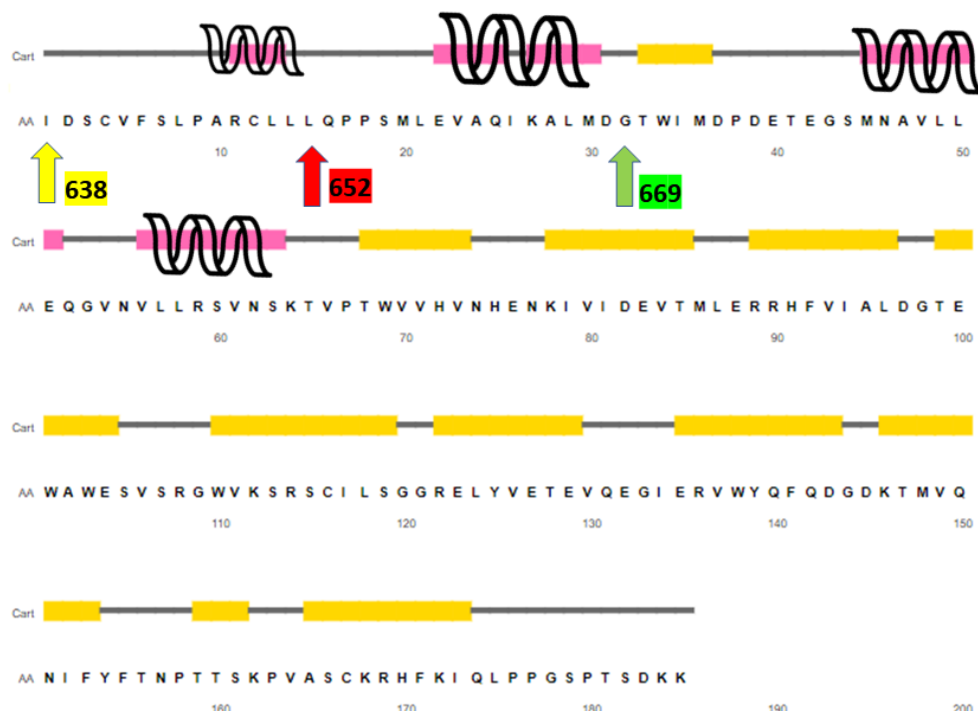
## 4.2. Conteúdo de estrutura secundária do domínio FABP-like foi estimado por PSIPRED

A predição por PSIPRED para a região 665-809 de LmxM.34.3070 fornece um conteúdo de estrutura secundária que é consistente com a topologia clássica das FABPs. (Figura 15A). A topologia das FABPs consiste de 10 fitas- $\beta$  antiparalelas formando um barril-beta, e um motivo hélice-alça-hélice conectando as fitas  $\beta$ A e  $\beta$ B (alça AB) (Figura 15B). Entretanto, hélices adicionais (resíduos em rosa entre as posições 638 e 669 na figura 16) foram preditas na extremidade N-terminal do domínio FABP-like de LmxM.34.3070 por PSIPRED quando realizamos a análise com toda a sequência de aminoácidos de LmxM.34.3070. Essas hélices também foram preditas nos modelos de ColabFold. Portanto, hipotetizamos que elas poderiam ser necessárias para o enovelamento correto desse domínio. Diante disso, decidimos testar três diferentes fragmentos da região C-terminal de LmxM.34.3070 de acordo com as predições realizadas com PSIPRED e ColabFold, e avaliá-

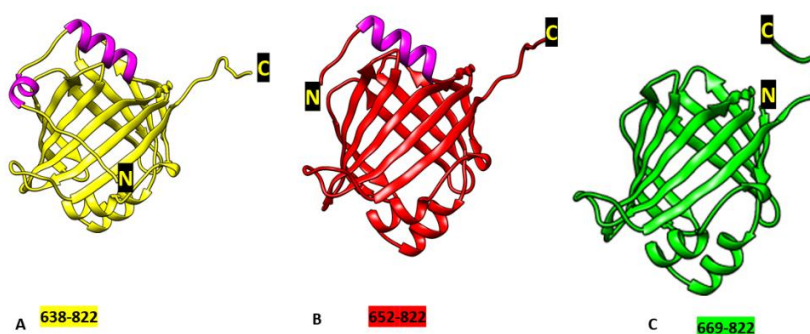
los quanto ao nível de expressão e solubilidade: **LmxM.34.3070<sub>638-822</sub>** contendo o barril FABP-like e todas as hélices N-terminais (Figura 17A), **LmxM.34.3070<sub>652-822</sub>** contendo o barril FABP-like e uma hélice N-terminal (Figura 17B), e **LmxM.34.3070<sub>669-822</sub>** apenas o barril FABP-like (Figura 17C). As setas em amarelo, vermelho e verde indicam o conteúdo de estrutura secundária de **LmxM.34.3070<sub>638-822</sub>**, **LmxM.34.3070<sub>652-822</sub>** e **LmxM.34.3070<sub>669-822</sub>** respectivamente na Figura 16. A construção **LmxM.34.3070<sub>652-822</sub>** foi clonada nos vetores de expressão pET24a e pET28a para a expressão em fusão com uma cauda de seis histidinas no C-terminal ou no N-terminal, respectivamente. As demais construções foram clonadas em pET28a para expressão em fusão com a cauda de histidinas no N-terminal (Tabela 6).



**Figura 15:** **A)** Conteúdo de estrutura secundária predita para a região C-terminal (aminoácidos 665-809) de LmxM.34.3070 por PSIPRED. Nesta análise, PSIPRED inferiu que esta região é formada por 11 fitas betas (caixas amarelas), e duas hélices interfita (entre as duas primeiras fitas A e B). **B)** A predição em PSIPRED é congruente com a topologia clássica das FABPs, onde são encontradas 10 fitas beta antiparalelas (A-J) e duas hélices interfita (L1) (Extraído e adaptado de Flower et.al, 2000).



**Figura 16:** Conteúdo de estrutura secundária predita para o fragmento 638-822 de LmxM.34.3070 por PSIPRED. Nesta análise, PSIPRED inferiu um conteúdo de estrutura secundária FABP-like, onde encontramos 11 fitas-beta (caixas amarelas) e duas hélices inter-fita. Todavia, PSIPRED também modelou hélices adicionais na região N-terminal (resíduos em rosa entre os aminoácidos 638 e 669). As setas indicam o início de cada um dos fragmentos que foram avaliados quanto a expressão, solubilidade e rendimento.



**Figura 17:** Modelos ColabFold para os fragmentos C-terminais de LmxM.34.3070 testados. **A)** Fragmento **LmxM.34.3070 638-822** contendo o barril FABP-like e todas as hélices N-terminais adicionais (em magenta). **B)** Fragmento **LmxM.34.3070 652-822** contendo o barril FABP-like e uma hélice N-terminal adicional (em magenta). **C)** Fragmento **LmxM.34.3070 669-822** contendo apenas o barril FABP-like.

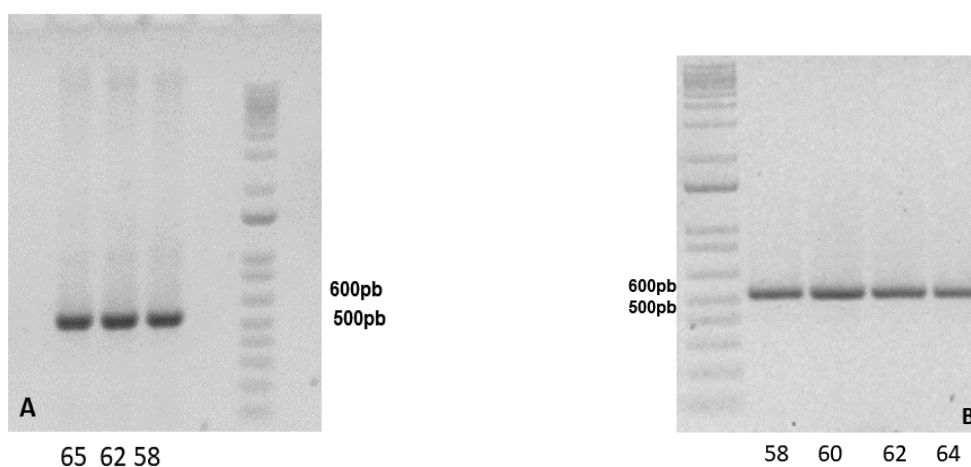
Construção	Sequências de aminoácidos	Número de hélices adicionais na região N-terminal	His-tag	MW	Tamanho (AA)	Método De clonagem
LmxM.34.3070 <sub>652-822</sub> + pET24a	MASMTGGQQMGRGSM LQPPSML <b>EVAQIKALMD</b> <u>GTWIMDPDETE</u> <u>GSMNAVLLEQGVNVLLRSVNSK</u> <u>TVPTWVVHVNHENKIVIDEVTMLERRHFVIALDGTEWAWESVSRG</u> <u>WVKSR</u> <u>SCILSGGRELYVETEVQEGIERVWYQFDGDKTMVQNIFYFTNPTT</u> <u>SKPVAS</u> <u>CKRHFQIQLPPGSPTSDDK</u> <u>LVPRGSKLAAALE</u> <b>HHHHHH</b>	1	C-terminal	23kDa	205	Restrição Com <i>Bam</i> HI e <i>Hind</i> III (com subclonagem em pCR2.1- TOPO)
LmxM.34.3070 <sub>652-822</sub> + pET28a	MGSS <b>HHHHHH</b> <u>SSGLVPRGSHMLQPPSML</u> <b>EVAQIKALMD</b> <u>GTWIMD</u> <u>PDETEGSMNAVLLEQGVNVLLRSVNSKTVPTW</u> <u>VVHVNHENKIVIDEVTMLERRHFVIALDGTEWAWESVSRGWVKSR</u> <u>SCILSGGRELY</u> <u>VETEVQEGIERVWYQFDGDKTMVQNIFYFTNPTT</u> <u>SKPVASCKRH</u> <u>FKIQLPPGSPT</u> <u>DDK</u>	1	N-terminal	21kDa	192	Restrição Com <i>Bam</i> HI e <i>Nde</i> I (com subclonagem em pCR2.1- TOPO)
LmxM.34.3070 <sub>669-822</sub> + pET28a	MGSS <b>HHHHHH</b> <u>SSGLVPRGSHM</u> <u>GTWIMDPDETEGSMNAVLLEQGV</u> <u>NVLLRSVNSKTVPTWVVHVNHENKIVIDE</u> <u>VTMLERRHFVIALDGTEWAWESVSRGWVKSRSCILSGGRELYVET</u> <u>EVQEGIER</u> <u>VWYQFDGDKTMVQNIFYFTNPTT</u> <u>SKPVASCKRHFQIQLPPGSPTS</u> <u>DDK</u>	-	N-terminal	19kDa	175	Restrição Com <i>Nde</i> I e <i>Hind</i> III (com subclonagem em pGEM)
LmxM.34.3070 <sub>638-822</sub> + pET28a	MGSS <b>HHHHHH</b> <u>SSGLVIDSCVFLPAR</u> <b>RCLL</b> <u>LQPPSML</u> <b>EVAQIKALMD</b> <u>GTWIMDPDETEGSM</u> <u>NAVLLEQGVNVLLRSVNSKTVPTWVVHVNHENKIVIDEVTMLERR</u> <u>HFVIALDGTEWAWESVSRGWVKSRSCILSGGRELYVETEVQEGIER</u> <u>VWYQFDGDKTMVQNIFYFTNPTT</u> <u>SKPVASCKRHFQIQLPPGSPTS</u> <u>DDK</u>	2	N-terminal	22kDa	200	Gibson (não houve subclonagem)

**Tabela 6:** Sequência, número de hélices adicionais em N-terminal, posição da *his-tag*, vetor de expressão, massa molecular, quantidade de resíduos de aminoácidos e método de clonagem utilizado para cada uma das quatro construções obtidas neste trabalho. Em amarelo destaca-se a posição da cauda de histidinas, em magenta os resíduos preditos para formar as hélices N-terminal adicionais modeladas por PSIPRED e ColabFold no domínio C-terminal FABP-like (LeiFABP), e em ciano os resíduos de aminoácidos preditos para compor o barril FABP-like em LmxM.34.3070. O primeiro aminoácido de LmxM.34.3070 está identificado em negrito. O sítio de trombina está sublinhado. Em nenhuma das sequências dos fragmentos recombinantes foi predito sinal de tráfego para membrana.

### 4.3. Amplificação por PCR dos fragmentos de interesse

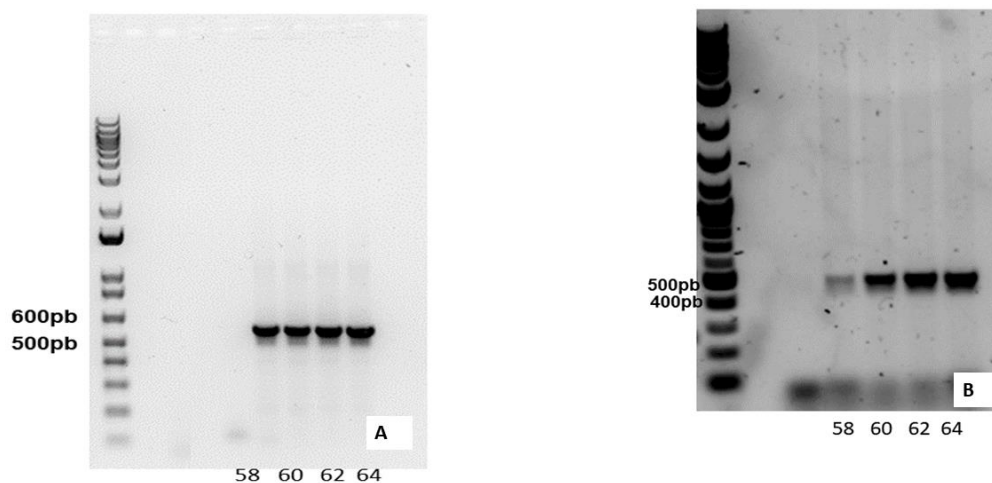
A primeira etapa executada para a amplificação dos fragmentos de interesse por PCR foi a realização de testes com os iniciadores desenhados. Previamente aos testes, utilizamos ferramentas *in silico* para prever temperatura de *melting*, conteúdo GC e formação de dímeros. Entretanto, para validar a eficiência e a especificidade dos oligonucleotídeos, também realizamos experimentos de PCR com gradiente de temperatura, onde testamos diferentes temperaturas de *annealing* para a verificação de bandas inespecíficas que poderiam se formar devido a *off targets* ou dimerização de *primers*.

Os oligonucleotídeos, em geral, tiveram uma boa performance, uma vez que o resultado da amplificação demonstrou apenas bandas consistentes com o tamanho, em pares de base, dos diferentes trechos escolhidos para codificar os diferentes fragmentos recombinantes de cada construção (Figuras 18 e 19). Não foram observadas bandas inespecíficas. Além disso, também confirmamos que as concentrações de reagentes e volumes estavam adequadas uma vez que não houve a observação de produtos inespecíficos. A formação de produtos inespecíficos poderia indicar excesso de DNA molde ou de *primers* (Roux, 2009).



**Figura 18:** **A)** Gradiente de temperatura (58, 62 e 65°C) para amplificação do fragmentos de DNA codificante de **LmxM.34.3070<sub>652-822</sub>** com os sítios de restrição de *Bam*HI e *Hind*III para clonagem em pET24a. *Set de primers*: 1 e 2 (Tabela 2). Produto esperado: 552 pares de base. **B)** Gradiente de temperatura (58, 60, 62 e 64°C) para amplificação

do fragmento de DNA codificante **LmxM.34.3070<sub>652-822</sub>** com os sítios de restrição de *NdeI* e *BamHI* para clonagem em em pET28a. Produto esperado: 534 pares de base. *Set de primers*: 3 e 4 (Tabela 2).

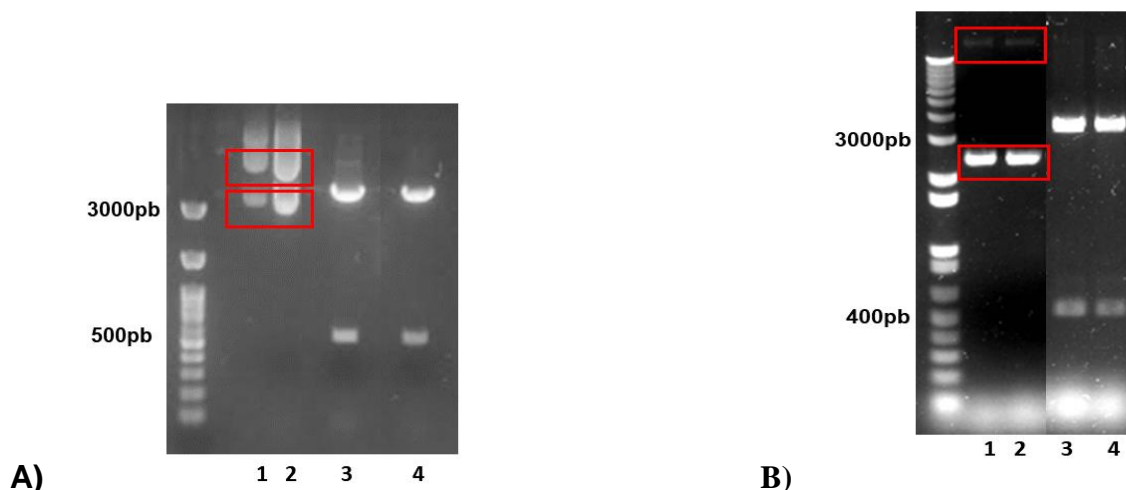


**Figura 19:** **A)** Gradiente de temperatura (58, 60, 62 e 64°C) para amplificação do fragmento de DNA codificante de **LmxM.34.3070<sub>638-822</sub>** com regiões de sobreposição para clonagem em pET28a pela metodologia de Gibson. Produto esperado: 575 pares de base. *Set de primers*: 5 e 6 (Tabela 2). **B)** Gradiente de temperatura (58, 60, 62 e 64°C) para amplificação do fragmento de DNA codificante **LmxM.34.3070<sub>669-822</sub>** com os sítios de restrição de *NdeI* e *HindIII* para clonagem em pET28a. Produto esperado: 483 pares de base. *Set de primers*: 7 e 8 (Tabela 2).

#### 4.4. Confirmação da subclonagem por ensaios de dupla de digestão: **pCR2.1-TOPO-LmxM.34.3070<sub>652-822</sub>** e **pGEM-LmxM.34.3070<sub>669-822</sub>**

A subclonagem dos fragmentos de DNA codificantes de **LmxM.34.3070<sub>652-822</sub>**, obtidos por PCR para clonagem em pET24A e pET28A, foi realizada em pCR2.1 com o kit TOPO-TA®. As vantagens do uso desse *kit* são a rapidez do procedimento, uma vez que o vetor pCR2.1 se encontra já linearizado e com extremidades de timidina prontas para ligação direta de produtos de PCR amplificados por Taq Polimerases, sem a necessidade de adição de DNA Ligase (Green & Sambrook, 2021). O princípio da técnica de subclonagem com o kit TOPO-TA® reside na conjugação do plasmídeo pCR2.1 à Topoisomerase I de *Vaccinia virus* que reconhece e cliva sítios 5'-(C/T) CCTT-3 em regiões *supercoiled* do DNA replicado, gerando as saliências de timidina aceptoras de adeninas adicionadas ao *amplicon* durante uma PCR com enzima Taq Polimerase (Green & Sambrook, 2021). A re-

circularização do vetor pCR2.1(TOPO)-inserto também é mediada pela Topoisomerase I (Green & Sambrook, 2021). Após a realização de miniprep, cada subclone purificado foi submetido a uma dupla digestão com os seus *sets* de enzimas de restrição específicas para confirmar a subclonagem. Os vetores pGEM são plasmídeos de alto número de cópias, e assim como o vetor pCR2.1 do kit TOPO-TA® possuem extremidades aceptoras de produtos de PCR amplificados com Taq Polimerases. O vetor pGEM foi utilizado para subclonar o fragmento de DNA codificante **LmxM.34.3070<sub>669-822</sub>**. O produto de miniprep deste subclone também foi submetido a ensaios de dupla digestão com o seu *set* de enzimas de restrição para confirmação da subclonagem.



**Figura 20:** A) Eletroforese em gel dos produtos de digestão de **LmxM.34.3070<sub>652-822</sub>** subclonado em pCR2.1-TOPO. Em 1 e 2 foram aplicados subclones não digeridos, onde observa-se a ocorrência de bandas em vários níveis da canaleta. Produto de digestão esperado para pCR2.1-TOPO linearizado sem o inserto: 3900 pb. **3:** Tamanho do inserto codificante de **LmxM.34.3070<sub>652-822</sub>** para clonagem em pET24a: 546 pb. **4:** Tamanho do inserto codificante de **LmxM.34.3070<sub>652-822</sub>** para clonagem em pET28a: 528 pb. B) Eletroforese em gel dos produtos de digestão de **LmxM.34.3070<sub>669-822</sub>** subclonado em pGEM. Em 1 e 2 foram aplicados subclones não digeridos, onde observa-se a ocorrência de bandas em vários níveis da canaleta. Produto de digestão esperado para pGEM linearizado sem o inserto: 3015 pb. 3 e 4: Tamanho do inserto codificante de **LmxM.34.3070<sub>669-822</sub>** para clonagem em pET28a: 476 pb.

Pelo perfil eletroforético (Figura 20 A e B) dos ensaios de dupla digestão observamos que houve a liberação de insertos dos subclones. Esse resultado mostra que as amplificações realizadas para obtenção dos fragmentos de DNA codificantes de **LmxM.34.3070<sub>652-822</sub>** e de **LmxM.34.3070<sub>669-822</sub>** com os sítios de restrição nas extremidades 5' e 3', foram bem sucedidas.

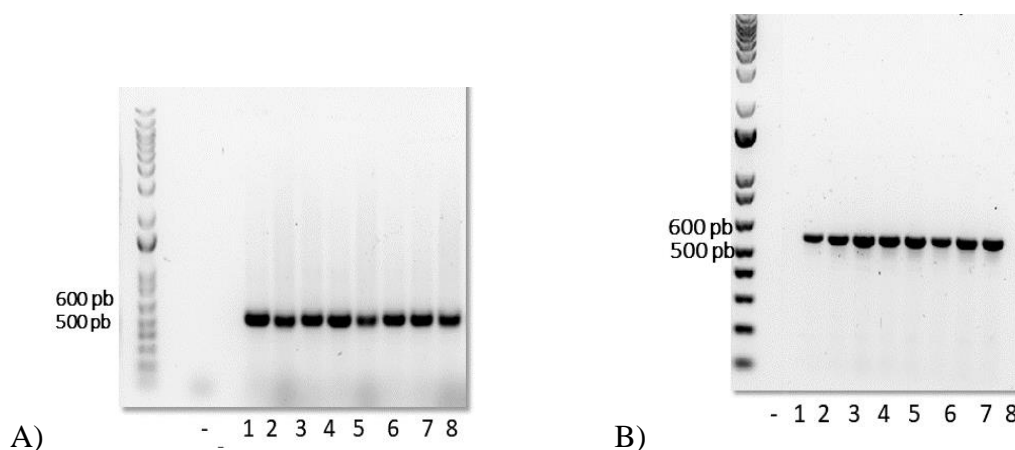


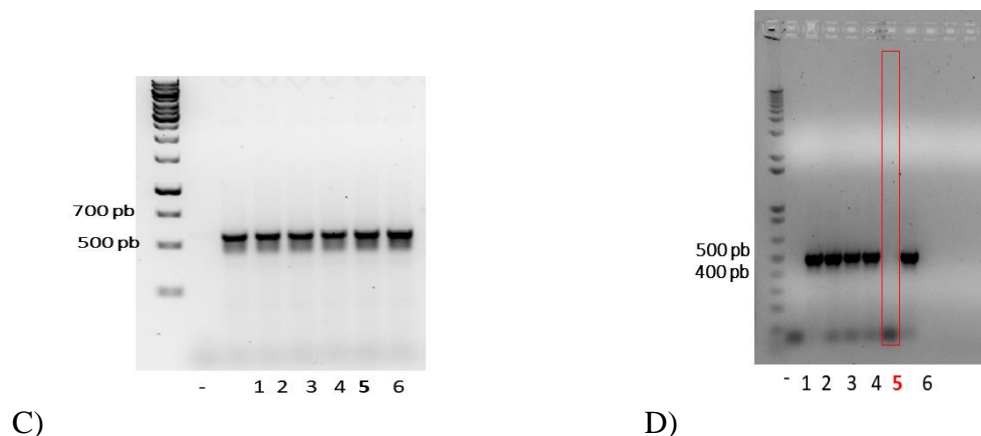
Portanto, as clonagens em vetor de expressão poderiam prosseguir, uma vez que esses insertos foram liberados com as extremidades coesivas.

Também é interessante notar que houve a circularização dos vetores pCR2.1-TOPO e pGEM a partir da reação com *amplicons* contendo *overhangs* de adenina. A migração de DNAs circulares não obedece ao princípio de proporcionalidade entre tamanho e velocidade podendo migrar diferencialmente na malha de agarose (E. G. Gibson et al., 2020). Isso é demonstrado pelo perfil eletroforético de amostras não digeridas, onde observa-se a presença de bandas em diferentes pontos da canaleta (caixas vermelhas nas figuras 20 A e B) que não são consistentes com o tamanho, em pares de base, dos subclones recombinantes pCR2.1-TOPO e pGEM contendo o inserto que resultaram da etapa de subclonagem.

#### 4.5. Confirmação da clonagem por PCR de colônia

A confirmação da presença de inserto foi realizada por um *screening* de colônias, que foram submetidas a uma PCR (Figuras 21 A, B, C e D). As colônias contendo os insertos de interesse foram cultivadas em meio líquido contendo canamicina e, posteriormente, essas culturas foram submetidas a miniprep e sequenciamento de DNA pela metodologia de Sanger (Sanger et al., 1977).





**Figura 21: PCRs de colônias dos clones construídos para expressão de diferentes construções de LeiFABP.** Cada número refere-se a uma colônia recombinante replicada na placa e submetida a PCR. "-" controle negativo de PCR. **A)** Eletroforese em gel das PCRs de colônias de **LmxM.34.3070<sub>652-822</sub>** clonado em pET24a para expressar em fusão com uma *tag* de histidinas em C-terminal. Esta PCR foi realizada com os mesmos iniciadores utilizados para amplificar o trecho codificante a partir do DNA genômico de *L.amazonensis*. Observa-se a detecção do inserto em todas as colônias: Produto de PCR esperado: 552 pb. **B)** Eletroforese em gel das PCRs de colônias de **LmxM.34.3070<sub>652-822</sub>** clonado em pET28a para expressar em fusão com uma *tag* de histidinas em N-terminal. Esta PCR foi realizada com os mesmos iniciadores utilizados para amplificar o trecho codificante **LmxM.34.3070<sub>652-822</sub>** a partir do DNA genômico de *L. amazonensis*. Observa-se a detecção do inserto em todas as colônias: Produto de PCR esperado: 534 pb. **C)** Eletroforese em gel das PCRs de colônias de **LmxM.34.3070<sub>638-822</sub>** clonado em pET28a para expressar a proteína de interesse em fusão com uma *tag* de histidinas em N-terminal. Esta PCR foi realizada com os mesmos iniciadores utilizados para amplificar o trecho codificante **LmxM.34.3070<sub>638-822</sub>** a partir do DNA genômico de *L. amazonensis*. Observa-se a detecção do inserto em todas as colônias: Produto de PCR esperado: 575 pb. **D)** Eletroforese em gel das PCRs de colônias de **LmxM.34.3070<sub>669-822</sub>** clonado em pET28a para expressar em fusão com uma *tag* de histidinas em N-terminal. Esta PCR foi realizada com os mesmos iniciadores utilizados para amplificar o trecho codificante de **LmxM.34.3070<sub>669-822</sub>** a partir do DNA genômico de *L.amazonensis*. Produto de PCR esperado: 483 pb. Observa-se que a **colônia 5** coletada da placa não apresentava o inserto **LmxM.34.3070<sub>669-822</sub>**.

#### 4.6. Sequenciamento dos clones de expressão

As sequências experimentais extraídas dos cromatogramas de sequenciamento foram alinhadas por BLAST com as respectivas sequências esperadas para os clones desenhados neste trabalho (Tabela 6). O algoritmo do BLAST (ferramenta BLASTN-NCBI) foi novamente utilizado para alinhar as duas sequências. BLAST fornece parâmetros importantes para avaliação de um alinhamento, como percentual de identidade e a cobertura (Madden et al., 2019). Avaliamos as sequências experimentais quanto a mutações que pudessem prejudicar a fase de leitura, tais como a incorporação incorreta de nucleotídeos que ocasionassem a formação de um stop códon prematuro, e *indels*. Para o clone **LmxM34.3070<sub>652-822</sub>(pET24a)** não foi obtido um percentual de identidade de 100% (Figura

22). Foram observadas duas substituições de DNA na fita antisenso: CTC → TTC (Figura 22, caixa amarela) que resultou na mudança de uma glutamato para uma lisina na posição 8 deste fragmento, e CCC → CCA (Figura 22, caixa preta) que não alterou a tradução de uma glicina na posição 147 deste fragmento.

As regiões correspondentes a estes trechos contendo substituições foram analisadas no cromatograma de sequenciamento. Para a substituição observada na trinca CTC → TTC que resulta na tradução de uma lisina no lugar de um glutamato na posição 8 do fragmento **LmxM34.3070<sub>652-822</sub>** clonado em pET24a, o cromatograma apresenta uma boa resolução, sem sobreposição de picos neste trecho, indicando que provavelmente durante a clonagem, ocorreu uma mutação (caixa amarela na Figura 23). Entretanto, para a região do cromatograma que corresponde ao trecho onde observou-se substituição na trinca CCC → CCA (caixa preta na Figura 24), que resultou numa mutação sinônima de glicina na posição 147, foi observada uma baixa resolução com sobreposição de picos, indicando que provavelmente, a diferença observada entre sequência experimental e a sequência esperada deveu-se a um artefato do sequenciamento. Além disso, o cromatograma mostra a existência de um pico de menor amplitude correspondente a uma citosina onde ocorre esta substituição, que é a base esperada para ocorrer nesta posição (seta preta na Figura 24). A polimerase utilizada para amplificar os fragmentos de DNA dessas construções mutantes foi a DreamTaq Polimerase, que apresenta maior probabilidade de incorporação incorreta quando comparada a outras polimerases com maior fidelidade, como a Phusion High Fidelity. Isso poderia explicar o *mismatch* observado no alinhamento entre sequências experimental e sequência esperada do clone **LmxM34.3070<sub>652-822</sub>(pET24a)** (Figura 22). No caso das construções **LmxM34.3070<sub>652-822</sub>(pET28a)** (Figura 25), **LmxM34.3070<sub>638-822</sub>(pET28a)** (Figura 26A) e **LmxM34.3070<sub>669-822</sub>(pET28a)** (Figura 26B), não se observou nenhuma mutação

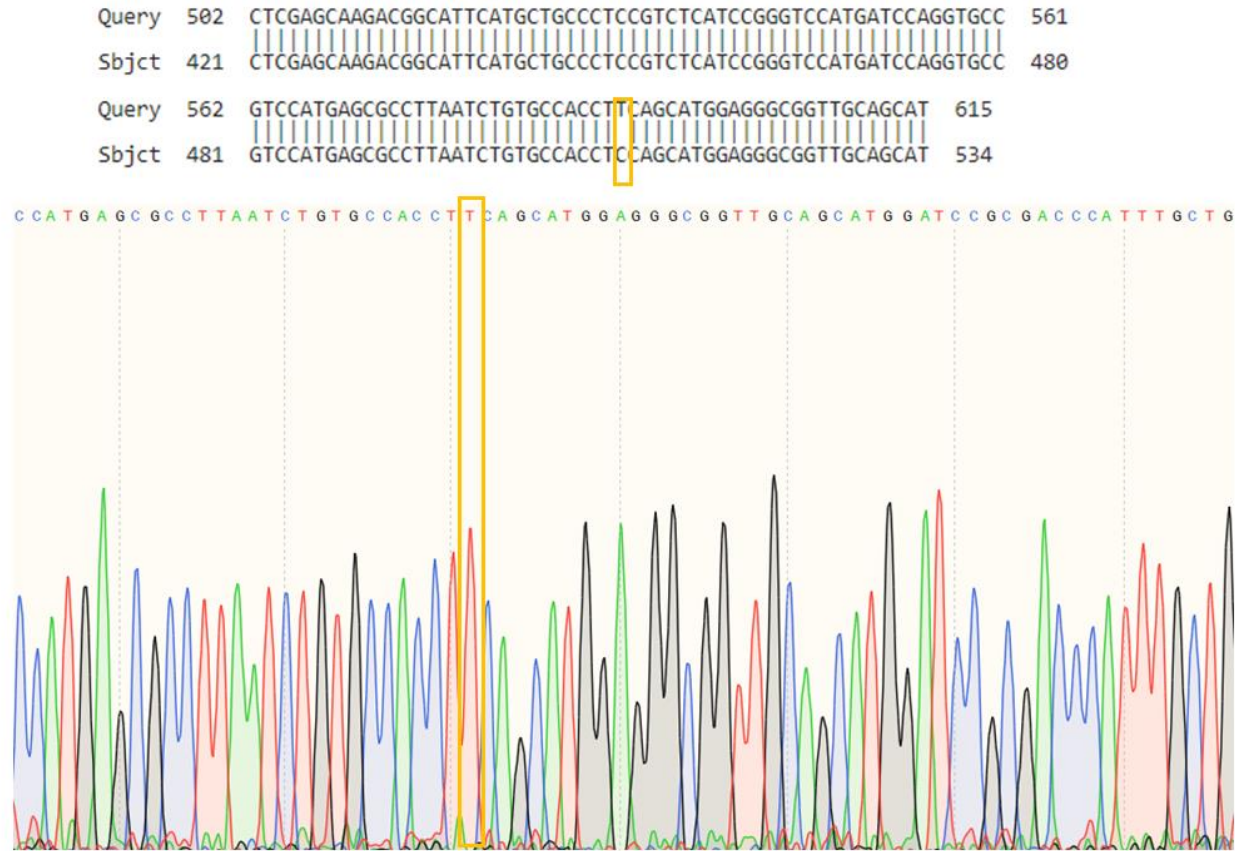
introduzida durante a clonagem. Todos os alinhamentos tiveram uma cobertura de 100% entre as sequências experimentais (Query) e sequências esperadas (Subject).

Query	82	GCTGCCGCGCGCACCAAGTTTCTTGCCGATGTTGGAGACCAGGGGGGAGCTGATCTT	141
Sbjct	1	GCTGCCGCGCGCACCAAGTTTCTTGCCGATGTTGGAGACCAGGGGGGAGCTGATCTT	60
Query	142	AAAGTGCCGCTTGACGAAGCCACAGGCTTCGACGTCGTCGGGTTTCGTGAAGTAGAATAT	201
Sbjct	61	AAAGTGCCGCTTGACGAAGCCACAGGCTTCGACGTCGTCGGGTTTCGTGAAGTAGAATAT	120
Query	202	GTTCTGCACCATTTGCTTGTACCCGCTGGAACCTGGTACCAGACGCGCTCGATTCCCTC	261
Sbjct	121	GTTCTGCACCATTTGCTTGTACCCGCTGGAACCTGGTACCAGACGCGCTCGATTCCCTC	180
Query	262	TTGCACCTCCGCTCCACATACAGCTCGCGCCCTCTGACAAGATGCACGAGCGGACTT	321
Sbjct	181	TTGCACCTCCGCTCCACATACAGCTCGCGCCCTCTGACAAGATGCACGAGCGGACTT	240
Query	322	CACCCAGCCGCGGAGACACTCTCCATGCCCACTCCGTCGGTTCGAGTGCATGACAAA	381
Sbjct	241	CACCCAGCCGCGGAGACACTCTCCATGCCCACTCCGTCGGTTCGAGTGCATGACAAA	300
Query	382	GTGCCCGCTCGAGCATGGTCACCTCGTCAATCACGATCTTGTCTCGTGGTTCACGTG	441
Sbjct	301	GTGCCCGCTCGAGCATGGTCACCTCGTCAATCACGATCTTGTCTCGTGGTTCACGTG	360
Query	442	GACAACCCAGGTGCGCACCGCTTGTCTGTTTACGCTGCGCAGCAAGACGTTGACGCCCTG	501
Sbjct	361	GACAACCCAGGTGCGCACCGCTTGTCTGTTTACGCTGCGCAGCAAGACGTTGACGCCCTG	420
Query	502	CTCGAGCAAGACGGCATTATGCTGCCCTCCGCTCATCCGGTCCATGATCCAGGTGCC	561
Sbjct	421	CTCGAGCAAGACGGCATTATGCTGCCCTCCGCTCATCCGGTCCATGATCCAGGTGCC	480
Query	562	GTCCATGAGCGCCTTAATCTGTGCCACTTCAGCATGGAGGGCGGTTGCAGCAT	615
Sbjct	481	GTCCATGAGCGCCTTAATCTGTGCCACTTCAGCATGGAGGGCGGTTGCAGCAT	534

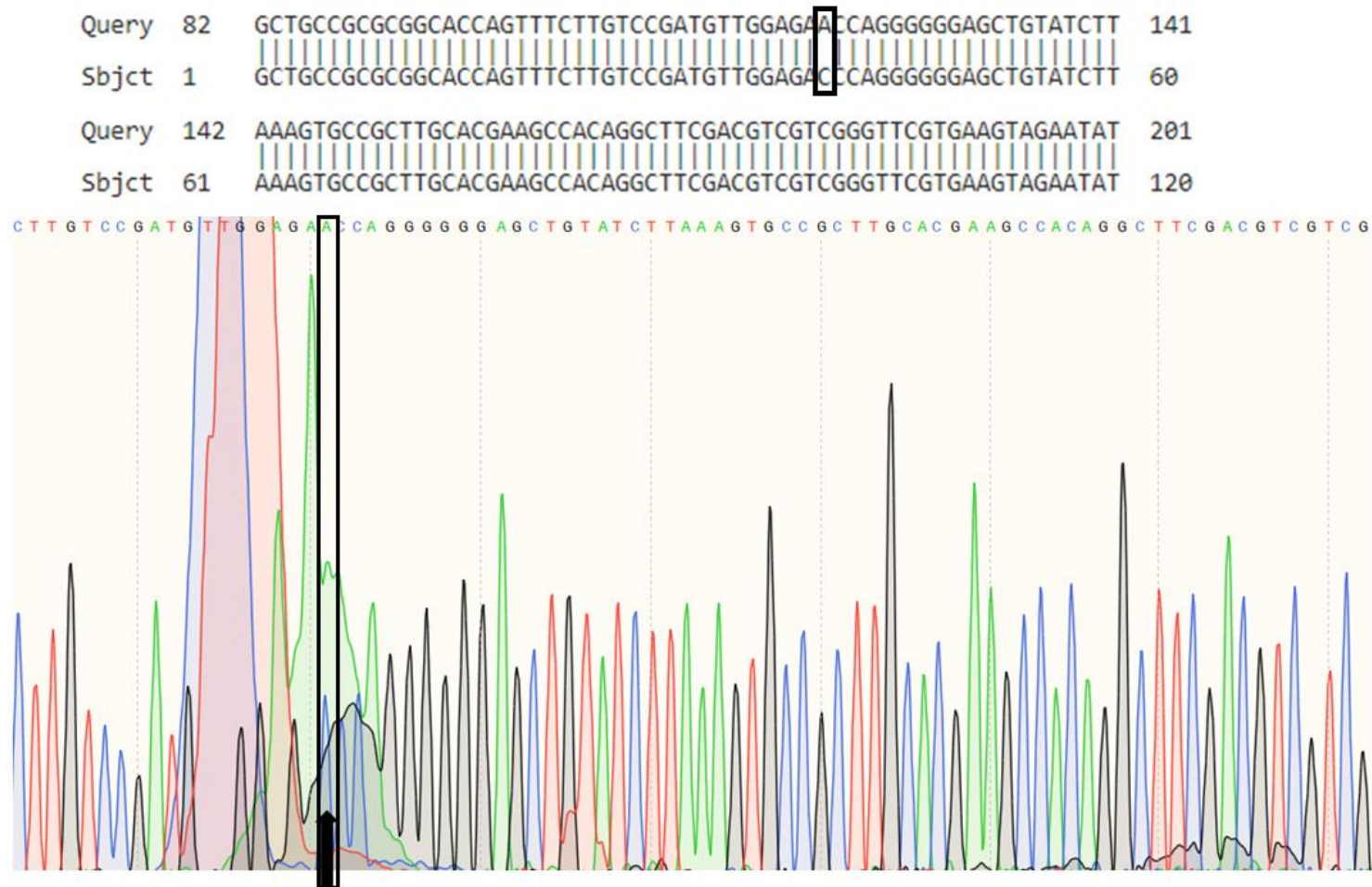
  

Query	1	LQPPSMLKVAQIKALMDGTWIMDPDETEGSMNAVLLLEQGVNLLRSVNSKTVPTWVHVN	60
Sbjct	1	LQPPSMLKVAQIKALMDGTWIMDPDETEGSMNAVLLLEQGVNLLRSVNSKTVPTWVHVN	60
Query	61	HENKIVIDEVTMLERRHFVIALDGTAWESVSRGWKSRSCILSGGRELYVETEVQEGI	120
Sbjct	61	HENKIVIDEVTMLERRHFVIALDGTAWESVSRGWKSRSCILSGGRELYVETEVQEGI	120
Query	121	ERVWYQFDGDKTMVQNIIFYFTNPPTTSKPVASCKRHFQIQLPPGSPTSDDK	171
Sbjct	121	ERVWYQFDGDKTMVQNIIFYFTNPPTTSKPVASCKRHFQIQLPPGSPTSDDK	171

**Figura 22:** Alinhamento por BLAST entre as sequências experimentais e as sequências de aminoácidos do fragmento **LmxM34.3070<sub>652-822</sub>(pET24a)**. Query: Sequência experimental do clone **LmxM34.3070<sub>652-822</sub>(pET24a)** extraída do cromatograma de sequenciamento. Subject: Sequência esperada do clone **LmxM34.3070<sub>652-822</sub>(pET24a)**. Foram observadas duas mutações (caixas pretas). As alterações na sequência de aminoácidos são mostradas abaixo (caixas vermelhas). Percentual de identidade: 99.61%. Não foram detectados *stop codons* prematuros.



**Figura 23:** A substituição observada na trinca CTC -> TTC que resulta na tradução de uma lisina no lugar de um glutamato na posição 8 do fragmento [LmxM34.3070<sub>652-822</sub>](#) clonada em pET24a está destacada pela caixa amarela. No cromatograma observa-se que este trecho apresenta uma boa resolução, indicando que provavelmente, esta mutação realmente ocorreu na etapa de clonagem.



**Figura 24:** A substituição observada na trinca CCC → CCA que resulta na tradução sinônima de uma glicina na posição 147 do fragmento *LmxM34.3070*<sub>652-822</sub> clonada em pET24a está destacada pela caixa preta. No cromatograma observa-se que este trecho não apresenta uma boa resolução, indicando que provavelmente, esta mutação não ocorreu na etapa de clonagem. A seta preta indica um pico de menor amplitude correspondente a uma citosina, que é a base esperada para esta posição da sequência.

Query	1	CTGCAACCGCCCTCCATGCTGGAGGTGGCACAGATTAAGGCGCTCATGGACGGCACCTGG	60
Sbjct	1	CTGCAACCGCCCTCCATGCTGGAGGTGGCACAGATTAAGGCGCTCATGGACGGCACCTGG	60
Query	61	ATCATGGACCCGGATGAGACGGAGGGCAGCATGAATGCCGTCTTGCTCGAGCAGGGCGTC	120
Sbjct	61	ATCATGGACCCGGATGAGACGGAGGGCAGCATGAATGCCGTCTTGCTCGAGCAGGGCGTC	120
Query	121	AACGCTTTGCTGCGCAGCGTGAACAGCAAGACGGTGCCGACCTGGGTTGTCCACGTGAAC	180
Sbjct	121	AACGCTTTGCTGCGCAGCGTGAACAGCAAGACGGTGCCGACCTGGGTTGTCCACGTGAAC	180
Query	181	CACGAGAACAAGATCGTGATTGACGAGGTGACCATGCTCGAGCGGGCGGCACCTTTGTCATC	240
Sbjct	181	CACGAGAACAAGATCGTGATTGACGAGGTGACCATGCTCGAGCGGGCGGCACCTTTGTCATC	240
Query	241	GCACTCGACGGGACGGAGTGGGCATGGGAGAGTGTCTCGCGCGCTGGGTGAAGTCGCGC	300
Sbjct	241	GCACTCGACGGGACGGAGTGGGCATGGGAGAGTGTCTCGCGCGCTGGGTGAAGTCGCGC	300
Query	301	TCGTGCATCTTGTGAGGAGGGCGCAGCTGTATGTGGAGACGGAGGTGCAAGAGGGGAATC	360
Sbjct	301	TCGTGCATCTTGTGAGGAGGGCGCAGCTGTATGTGGAGACGGAGGTGCAAGAGGGGAATC	360
Query	361	GAGCGCGTCTGGTACCAGTTCAGGACGGTGACAAGACAATGGTGAGAACATATTCTAC	420
Sbjct	361	GAGCGCGTCTGGTACCAGTTCAGGACGGTGACAAGACAATGGTGAGAACATATTCTAC	420
Query	421	TTCACGAACCCGACGACGTCGAAGCCTGTGGCTTCGTGCAAGCGGCACCTTAAGATACAG	480
Sbjct	421	TTCACGAACCCGACGACGTCGAAGCCTGTGGCTTCGTGCAAGCGGCACCTTAAGATACAG	480
Query	481	CTCCCCCTGGGTCTCCAACATCGGACAAGAAATAG	516
Sbjct	481	CTCCCCCTGGGTCTCCAACATCGGACAAGAAATAG	516

Query	1	LQPPSMLEVAQIKALMDGTWIMDPDETEGSMNAVLLLEQGVNLLRSVNSKTVPWVHVN	60
Sbjct	1	LQPPSMLEVAQIKALMDGTWIMDPDETEGSMNAVLLLEQGVNLLRSVNSKTVPWVHVN	60
Query	61	HENKIVIDEVTMLERRHFVIALDGTAWAWESVSRGWKSRSCILSGGRELYVETEVQEGI	120
Sbjct	61	HENKIVIDEVTMLERRHFVIALDGTAWAWESVSRGWKSRSCILSGGRELYVETEVQEGI	120
Query	121	ERVWYQFDGDKTMVQNIIFYFTNPTTSKPVASCKRHFQIQLPPGSPTSDDK	171
Sbjct	121	ERVWYQFDGDKTMVQNIIFYFTNPTTSKPVASCKRHFQIQLPPGSPTSDDK	171

**Figura 25:** Alinhamento por BLAST entre as sequências experimentais e as sequências de aminoácidos do fragmento **LmxM34.3070<sub>652-822</sub>(pET28a)**. Query: Sequência experimental do clone **LmxM34.3070<sub>652-822</sub>(pET28a)** extraída do cromatograma de sequenciamento. Subject: Sequência esperada do clone **LmxM34.3070<sub>652-822</sub>(pET28a)**. Percentual de identidade: 100%. Não foram detectados *stop codons* prematuros.



```

Query 1  ATCGACAGTGCCTTTCTCGTGCCTGGCGCGCTGTCTACTACTGCAACCGCCCTCCATG 60
Sbjct 558 ATCGACAGTGCCTTTCTCGTGCCTGGCGCGCTGTCTACTACTGCAACCGCCCTCCATG 499

Query 61  CTGGAGGTGGCACAGATTAAGGCGCTCATGGACGGCACCTGGATCATGGACCCGGATGAG 120
Sbjct 498  CTGGAGGTGGCACAGATTAAGGCGCTCATGGACGGCACCTGGATCATGGACCCGGATGAG 439

Query 121  ACGGAGGGCAGCATGAATGCCGCTTTGCTCGAGCAGGGCGTCAACGCTTTGCTGCGCAGC 180
Sbjct 438  ACGGAGGGCAGCATGAATGCCGCTTTGCTCGAGCAGGGCGTCAACGCTTTGCTGCGCAGC 379

Query 181  GTGAACAGCAAGACGGTGCCGACCTGGGTTGTCCACGTGAACCACGAGAACAAGATCGTG 240
Sbjct 378  GTGAACAGCAAGACGGTGCCGACCTGGGTTGTCCACGTGAACCACGAGAACAAGATCGTG 319

Query 241  ATTGACGAGGTGACCATGCTCGAGCGGCGGCACTTTGTCATCGCACTCGACGGGACGGAG 300
Sbjct 318  ATTGACGAGGTGACCATGCTCGAGCGGCGGCACTTTGTCATCGCACTCGACGGGACGGAG 259

Query 301  TGGGCATGGGAGAGTGTCTCGCGCGGCTGGGTGAAGTCGCGCTCGTGATCTTTGTCAGGA 360
Sbjct 258  TGGGCATGGGAGAGTGTCTCGCGCGGCTGGGTGAAGTCGCGCTCGTGATCTTTGTCAGGA 199

Query 361  GGGCGCAGCTGTATGTGGAGACGGAGGTGCAAGAGGGAATCGAGCGCTCTGGTACCAG 420
Sbjct 198  GGGCGCAGCTGTATGTGGAGACGGAGGTGCAAGAGGGAATCGAGCGCTCTGGTACCAG 139

Query 421  TTCCAGGACGGTGACAAGACAATGGTGCAGAACATATTTACTTTCACGAACCCGACGACG 480
Sbjct 138  TTCCAGGACGGTGACAAGACAATGGTGCAGAACATATTTACTTTCACGAACCCGACGACG 79

Query 481  TCGAAGCCTGTGGCTTCGTGCAAGCGGCACTTTAAGATACAGCTCCCCCTGGGTCTCCA 540
Sbjct 78  TCGAAGCCTGTGGCTTCGTGCAAGCGGCACTTTAAGATACAGCTCCCCCTGGGTCTCCA 19

Query 541  ACATCGGACAAGAAATAG 558
Sbjct 18  ACATCGGACAAGAAATAG 1

```

```

Query 1  GGCACCTGGATCATGGACCCGGATGAGACGGAGGGCAGCATGAATGCCGCTTTGCTCGAG 60
Sbjct 465  GGCACCTGGATCATGGACCCGGATGAGACGGAGGGCAGCATGAATGCCGCTTTGCTCGAG 406

Query 61  CAGGGCGTCAACGCTTTGCTGCGCAGCGTGAACAGCAAGACGGTGCCGACCTGGGTTGTC 120
Sbjct 405  CAGGGCGTCAACGCTTTGCTGCGCAGCGTGAACAGCAAGACGGTGCCGACCTGGGTTGTC 346

Query 121  CACGTGAACCACGAGAACAAGATCGTGATTGACGAGGTGACCATGCTCGAGCGGCGGCAC 180
Sbjct 345  CACGTGAACCACGAGAACAAGATCGTGATTGACGAGGTGACCATGCTCGAGCGGCGGCAC 286

Query 181  TTTGTCATCGCACTCGACGGGACGGAGTGGGCATGGGAGAGTGTCTCGCGCGGCTGGGTG 240
Sbjct 285  TTTGTCATCGCACTCGACGGGACGGAGTGGGCATGGGAGAGTGTCTCGCGCGGCTGGGTG 226

Query 241  AAGTCGCGCTCGTGATCTTTGTCAGGAGGGCGCGAGCTGTATGTGGAGACGGAGGTGCAA 300
Sbjct 225  AAGTCGCGCTCGTGATCTTTGTCAGGAGGGCGCGAGCTGTATGTGGAGACGGAGGTGCAA 166

Query 301  GAGGGAATCGAGCGCTCTGGTACCAGTTCAGGACGGTGACAAGACAATGGTGCAGAAC 360
Sbjct 165  GAGGGAATCGAGCGCTCTGGTACCAGTTCAGGACGGTGACAAGACAATGGTGCAGAAC 106

Query 361  ATATTCTACTTCACGAACCCGACGACGTCGAAGCCTGTGGCTTCGTGCAAGCGGCACTTT 420
Sbjct 105  ATATTCTACTTCACGAACCCGACGACGTCGAAGCCTGTGGCTTCGTGCAAGCGGCACTTT 46

Query 421  AAGATACAGCTCCCCCTGGGTCTCCAACATCGGACAAGAAATAG 465
Sbjct 45  AAGATACAGCTCCCCCTGGGTCTCCAACATCGGACAAGAAATAG 1

```

```

Query 1  IDSCVFLPARCLLLQPPSHLEVAQIKALMDGTWIMDPDETEGSMNAVLLLEQGVNVLRS 60
Sbjct 1  IDSCVFLPARCLLLQPPSHLEVAQIKALMDGTWIMDPDETEGSMNAVLLLEQGVNVLRS 60

Query 61  VNSKTVPTWVHVHNHENKIVIDEVTMLERRHFVIALDGTENAWESVSRGWKSRSCILSG 120
Sbjct 61  VNSKTVPTWVHVHNHENKIVIDEVTMLERRHFVIALDGTENAWESVSRGWKSRSCILSG 120

Query 121  GRELYVETEVEQGIERVWYQFDGDKTMVQNIIFYFTNPTTSKPVASCKRHFQIQLPPGSP 180
Sbjct 121  GRELYVETEVEQGIERVWYQFDGDKTMVQNIIFYFTNPTTSKPVASCKRHFQIQLPPGSP 180

Query 181  TSDKK 185
Sbjct 181  TSDKK 185

```

```

Query 1  GTWIMDPDETEGSMNAVLLLEQGVNVLRSVNSKTVPTWVHVHNHENKIVIDEVTMLERRH 60
Sbjct 1  GTWIMDPDETEGSMNAVLLLEQGVNVLRSVNSKTVPTWVHVHNHENKIVIDEVTMLERRH 60

Query 61  FVIALDGTENAWESVSRGWKSRSCILSGGRELYVETEVEQGIERVWYQFDGDKTMVQNI 120
Sbjct 61  FVIALDGTENAWESVSRGWKSRSCILSGGRELYVETEVEQGIERVWYQFDGDKTMVQNI 120

Query 121  IFYFTNPTTSKPVASCKRHFQIQLPPGSPTSDDK 154
Sbjct 121  IFYFTNPTTSKPVASCKRHFQIQLPPGSPTSDDK 154

```

A)

B)

**Figura 26:** Alinhamento por BLAST entre as sequências de DNA experimental e esperada, e das sequências de aminoácidos experimental e esperada. **A)** Query: Sequência experimental do clone **LmxM34.3070<sub>638-822</sub>** extraída do cromatograma de sequenciamento. Subject: Sequência esperada do clone **LmxM34.3070<sub>638-822</sub>**. Não foram observadas mutações. Percentual de identidade: 100%. Não foram detectados *stop codons* prematuros. **B)** Query: Sequência experimental do clone **LmxM34.3070<sub>669-822(pET28a)</sub>** extraída do cromatograma de sequenciamento; Subject: Sequência esperada do clone **LmxM34.3070<sub>669-822(pET28a)</sub>**. Não foram observadas mutações. Percentual de identidade: 100%. Não foram detectados *stop codons* prematuros.

#### 4.7. **LmxM.34.3070<sub>652-822</sub>** é mais expresso quando fusionado a uma *his-tag* N-terminal

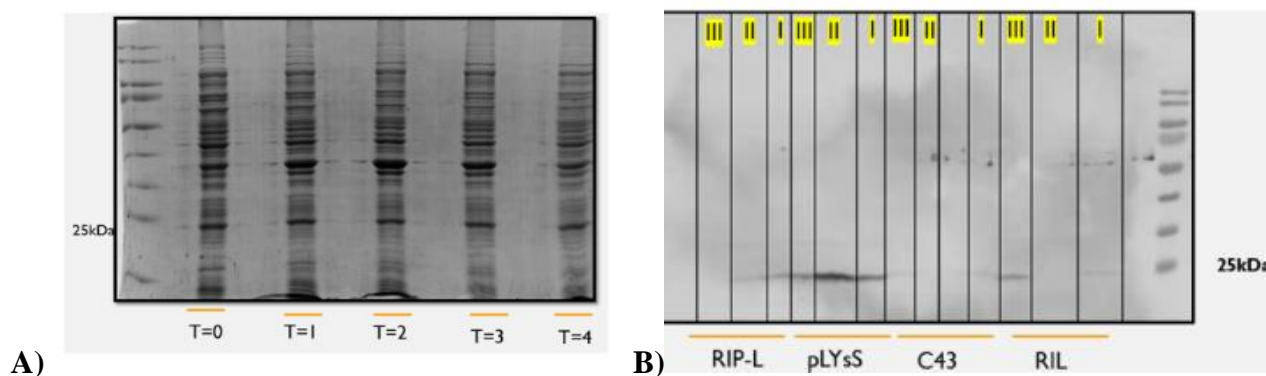
Inicialmente, os testes foram realizados com **LmxM.34.3070<sub>652-822</sub>** clonado em pET24a e pET28a pois gostaríamos de obter resultados sobre o melhor vetor a ser utilizado para a expressão heteróloga dos nossos fragmentos recombinantes, bem como gostaríamos de descobrir a melhor posição para a inserção da *tag* de histidinas. A primeira etapa do teste de expressão foi a realização de um *screening* de cepas. Nosso banco de células de *E.coli* para expressão de proteínas incluía as cepas: BL21 (DE3), BL21 (DE3) pLysS, C43 (DE3), BL21 (DE3) Codon plus (RIL-P e RIL). As cepas de expressão de *E.coli* foram engenheiradas e possuem genótipos específicos para aumentar o sucesso da produção de uma proteína recombinante de acordo com suas características como toxicidade, solubilidade e composição de códons (Tabela 3).

O fragmento **LmxM.34.3070<sub>652-822</sub>** clonado em pET24a e expresso em fusão com um *tag* de histidinas na região C-terminal não apresentou alta taxa de expressão heteróloga em diferentes cepas de BL21. O perfil eletroforético das amostras induzidas não demonstrou claramente as bandas diferencialmente expressas quando comparadas com o perfil eletroforético de amostras não induzidas no gel SDS-PAGE (Figura 27A). Portanto, para verificar se houve ou não a expressão de **LmxM.34.3070<sub>652-822</sub>** clonado em pET24a recorremos a western blot. Os resultados desse experimento demonstraram que células transformadas e induzidas com **LmxM.34.3070<sub>652-822</sub>(pET24a)**, para expressão do domínio FABP-like com *his-tag* em C-terminal apresentaram nenhum ou baixos níveis de expressão, sendo apenas reveladas bandas pouco aparentes que foram reconhecidas pelo anticorpo Anti-His (Figura 27B).

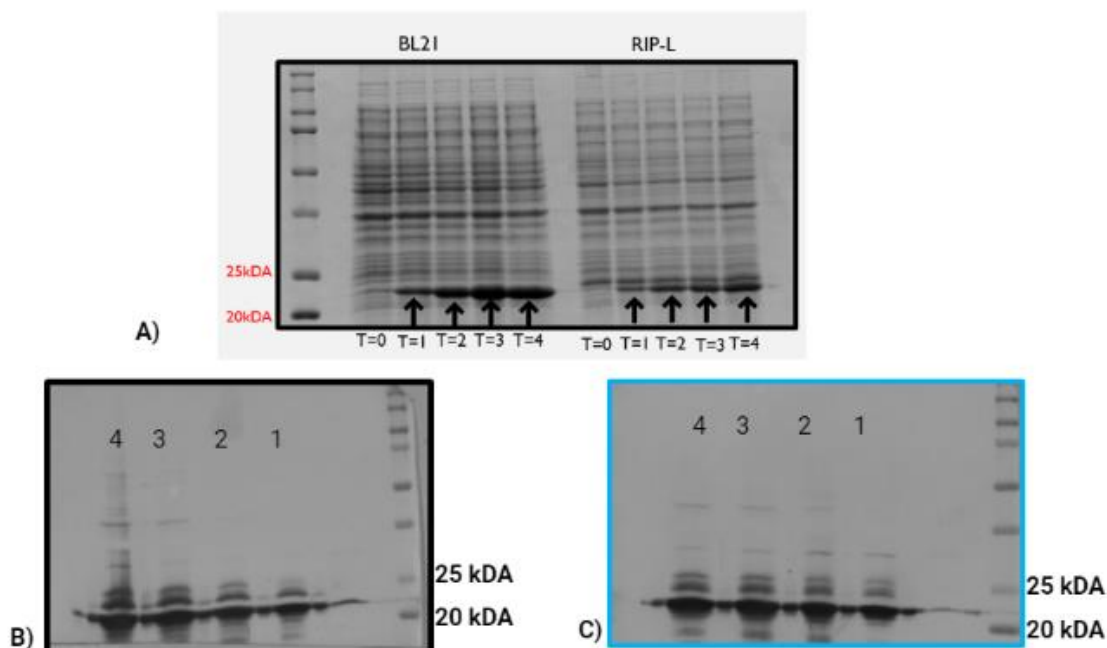
Posteriormente, iniciamos os testes de expressão heteróloga com o clone **LmxM.34.3070<sub>652-822</sub>(pET28a)**, para expressar a *his-tag* na região N-terminal. Quando

comparado com o primeiro clone, obtido em pET24A, **LmxM.34.3070<sub>652-822</sub>(pET28a)**, apresentou ótimo perfil de expressão, com bandas espessas e bem aparentes no perfil eletroforético de amostras induzidas não presentes na amostra não induzida, sendo claramente visualizadas tanto por SDS-PAGE quanto por western blot (Figura 28).

É bem aceito na literatura que a posição da *tag* de histidinas na região N-terminal pode aumentar o nível de expressão de uma proteína recombinante por facilitar o processo de iniciação da tradução pelo ribossomo (Francis & Page, 2010). Nossos testes iniciais mostraram que houve um aumento considerável da expressão de **LmxM.34.3070<sub>652-822</sub>** recombinante quando clonado em pET28a para expressão em fusão com a *tag* de histidinas em N-terminal. Todavia, também não é possível excluir a possibilidade de que a baixa taxa de expressão do clone LmxM.34.3070<sub>652-822</sub> clonado em pET24a seja devido à presença da mutação de um Glu por uma Lys no N-terminal (Figura 23). Por essa razão, abandonamos o **LmxM.34.3070<sub>652-822</sub>(pET24a)** nas etapas seguintes de expressão em larga escala e testes de purificação.

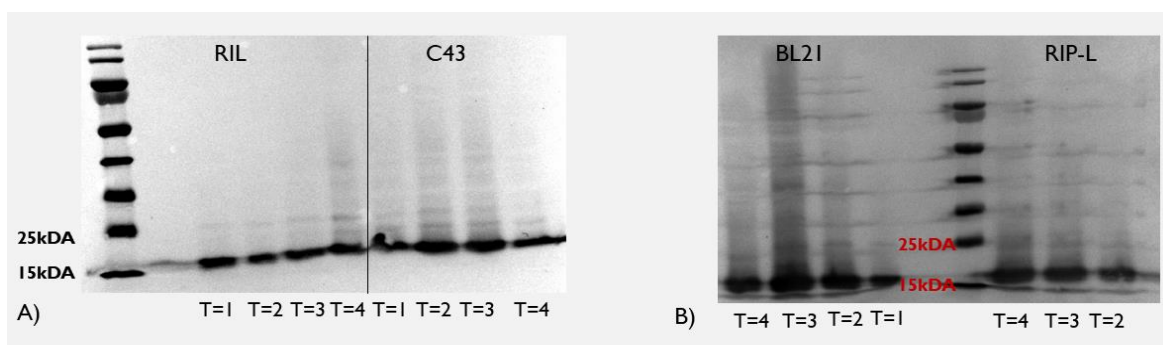


**Figura 27:** A) Perfil eletroforético de expressão do clone **LmxM.34.3070<sub>652-822</sub>(pET24a)**, com uma *tag* de histidinas fusionada na região C-terminal (~23kDa) em BL21 avaliado por SDS-PAGE. T=0 amostra não induzida, T=1 após 1 hora de indução, T=2 após 2 horas de indução, T=3 após 3 horas de indução, T=4 após 4 horas de indução. Não foram observadas bandas diferencialmente expressas em amostras induzidas. B) Detecção da expressão do fragmento **LmxM.34.3070<sub>652-822</sub>** por western blot em culturas induzidas de diferentes cepas de BL21 transformadas com o clone **LmxM.34.3070<sub>652-822</sub>(pET24a)**. I= após 1 hora de indução, II= após duas horas de indução, III= após 3 horas de indução. Apenas em RIL, pLysS e RIP-L foram reveladas bandas de expressão. Entretanto, as bandas demonstram uma baixa expressão do fragmento **LmxM.34.3070<sub>652-822</sub>**.

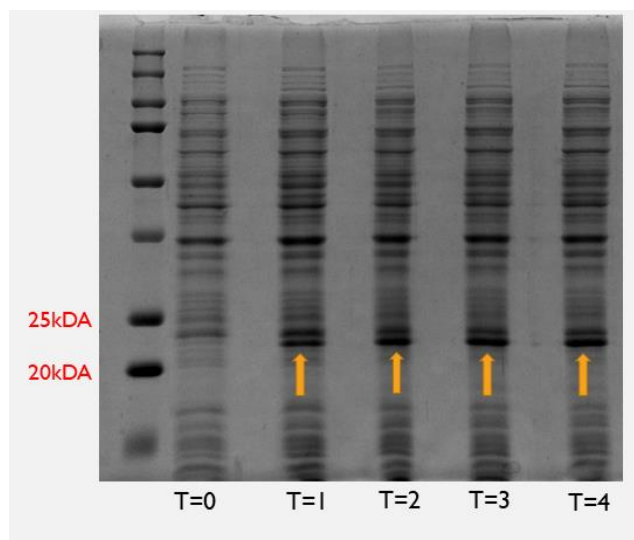


**Figura 28:** A) Perfis eletroforéticos de expressão do clone **LmxM.34.3070<sub>652-822</sub>(pET28a)**, com uma *tag* de histidinas fusionada na região N-terminal (~21kDa) em BL21 e RIP-L avaliados por SDS-PAGE. T=0 amostra não induzida, T=1 após 1 hora de indução, T=2 após 2 horas de indução, T=3 após 3 horas de indução, T=4 após 4 horas de indução. As bandas de expressão de **LmxM.34.3070<sub>652-822</sub>** foram indicadas por setas pretas. B) Detecção da expressão por western blot em amostras de cultura induzida de BL21 transformadas com o clone **LmxM.34.3070<sub>652-822</sub>(pET28a)**, 1= após 1 hora de indução, 2= após duas horas de indução, 3= após 3 horas de indução, 4= após 4 horas de indução. C) Detecção do fragmento **LmxM.34.3070<sub>652-822</sub>** por western blot em amostras de cultura induzida de RIP-L transformadas com o clone **LmxM.34.3070<sub>652-822</sub>(pET28a)**, 1= após 1 hora de indução, 2= após duas horas de indução, 3= após 3 horas de indução, 4= após 4 horas de indução.

Similarmente ao clone **LmxM.34.3070<sub>652-822</sub>(pET28a)**, os clones **LmxM.34.3070<sub>669-822</sub>(pET28a)** (Figura 29-A e B) e **LmxM.34.3070<sub>638-822</sub>(pET28a)** (Figura 30) apresentaram boa taxa de expressão heteróloga em *E.coli*. Portanto, esses três clones seguiram nas etapas posteriores de produção em larga escala e purificação e foram avaliados quanto ao seu rendimento e solubilidade. Essas avaliações foram empenhadas para que pudéssemos escolher o melhor fragmento para estudar o domínio LeiFABP.



**Figura 29:** **A)** Bandas de expressão do clone **LmxM.34.3070<sub>669-822</sub>(pET28a)** com uma *tag* de histidinas na região N-terminal (~19kDa) em RIL e C43 detectadas por western blot. T=1 após 1 hora de indução, T=2 após 2 horas de indução, T=3 após 3 horas de indução, T=4 após 4 horas de indução. **B)** Bandas de expressão do clone **LmxM.34.3070<sub>669-822</sub>(pET28a)** com uma *tag* de histidinas fusionada na região N-terminal (~19kDa) em BL21 e RIP-L detectadas por western blot. T=1 após 1 hora de indução, T=2 após 2 horas de indução, T=3 após 3 horas de indução, T=4 após 4 horas de indução.



**Figura 30:** Perfil eletroforético de expressão do clone **LmxM.34.3070<sub>638-822</sub>(pET28a)** com uma *tag* de histidinas fusionada na região N-terminal (~22kDa) em RIP-L avaliado por SDS-PAGE. T=0 amostra não induzida, T=1 após 1 hora de indução, T=2 após 2 horas de indução, T=3 após 3 horas de indução, T=4 após 4 horas de indução. O perfil eletroforético demonstra claramente as bandas de expressão (setas amarelas) com massas consistentes com **LmxM.34.3070<sub>638-822</sub>**, e por isso, não fizemos a validação da expressão por western blot.

Construção	Cepa	Deteção de bandas de expressão de LeIFABP por SDS PAGE	Deteção de bandas de LeIFABP por Western Blot
<b>LmxM.34.3070<sub>652-822</sub> + pET24a</b>	BL21(DE3) pLysS	-	+
<b>LmxM.34.3070<sub>652-822</sub> + pET24a</b>	BL21(DE3) RIL	-	+
<b>LmxM.34.3070<sub>652-822</sub> + pET24a</b>	BL21(DE3) RIP-L	-	+
<b>LmxM.34.3070<sub>652-822</sub> + pET24a</b>	(DE3) C43	-	+
<b>LmxM.34.3070<sub>652-822</sub> + pET28a</b>	BL21(DE3)	+	+
<b>LmxM.34.3070<sub>652-822</sub> + pET28a</b>	BL21(DE3) RIL-P	+	+
<b>LmxM.34.3070<sub>669-822</sub> + pET28a</b>	BL21(DE3)	NA	+
<b>LmxM.34.3070<sub>669-822</sub> + pET28a</b>	BL21(DE3) RIL	NA	+
<b>LmxM.34.3070<sub>669-822</sub> + pET28a</b>	(DE3) C43	NA	+
<b>LmxM.34.3070<sub>669-822</sub> + pET28a</b>	BL21 (DE3) RIP-L	NA	+
<b>LmxM.34.3070<sub>638-822</sub> + pET28a</b>	BL21(DE3) RIP-L	+	NA

**Tabela 7:** Resultados dos testes de expressão dos diferentes fragmentos da região C-terminal de LmxM.34.3070. A confirmação da expressão foi realizada por SDS-PAGE e western blot. (-) não detectado, (+) detectado, (NA) não se aplica (Nesses casos, a confirmação de expressão foi realizada apenas por SDS-PAGE ou apenas por western blot).

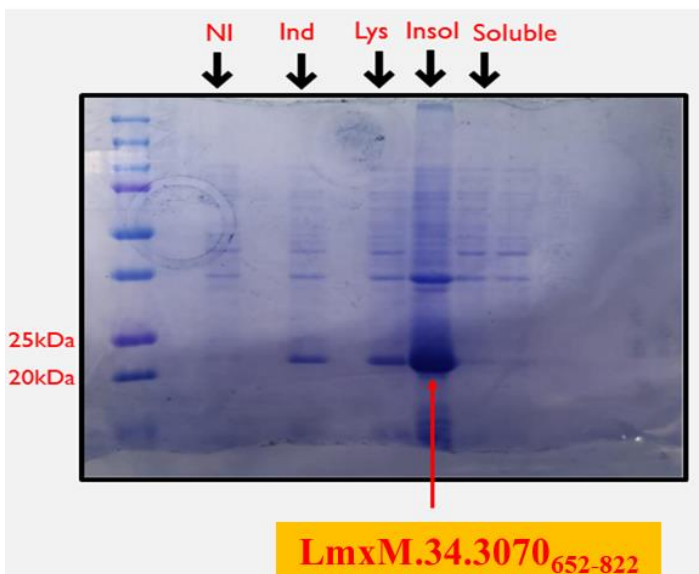
#### 4.8. Expressão em larga escala e purificação de **LmxM.34.3070<sub>652-822</sub>**

Os resultados preliminares dos testes de expressão dos **LmxM.34.3070<sub>652-822(pET24a)</sub>**, e **LmxM.34.3070<sub>652-822(pET28a)</sub>**, demonstraram que a *tag* de histidinas em N-terminal teve contribuição importante para aumentar a taxa de expressão dos diferentes fragmentos recombinantes, ou que a introdução de uma mutação em **LmxM.34.3070<sub>652-822(pET24a)</sub>**, afetou negativamente a expressão desse clone. As próximas etapas que foram realizadas para obtenção de amostras puras para ensaios de biologia estrutural seriam a indução da expressão em larga escala a partir de preparos de 0.5 L de cultura e otimizações de purificação por cromatografia. Todas as tentativas de otimização das etapas de lise e purificação ocorreram a partir de *pellets* obtidos de 0.5 L de cultura.

Como estávamos lidando com fragmentos gênicos de um organismo eucarioto codificante de um domínio solúvel, apostamos que RIP-L poderia ser a melhor hospedeira para a expressão heteróloga dos fragmentos contendo o domínio FABP-like de LmxM.34.3070. O viés de códon é um problema comum em expressão heteróloga de proteínas (Pacheco et al., 2012). A produção de proteínas recombinantes contendo códons raros pode levar à depleção de tRNAs raros ocasionando incorporação incorreta de aminoácidos e assim, prejudicar a tradução da proteína heteróloga (Francis & Page, 2010). Apesar de outras cepas de *E.coli* resultarem em expressão de nossos clones de LeiFABP, RIP-L foi a escolhida por carregar plasmídeos para a expressão de tRNAs raros de arginina, leucina, isoleucina e prolina

#### 4.9. A temperatura de indução e a concentração de IPTG influenciaram na solubilidade de **LmxM.34.3070<sub>652-822</sub>** recombinante

Iniciamos os testes com o clone **LmxM.34.3070<sub>652-822</sub>** para produção em larga escala de do domínio FABP-like recombinante. Esta construção compreende o barril-beta e o motivo hélice-alça-hélice, e uma alfa-hélice adicional na região N-terminal (Tabela 6). O primeiro teste (Teste 1, Tabela 4.1) foi realizado com células RIP-L, 600  $\mu$ M de IPTG e indução a 37°C. Nessas condições, o fragmento **LmxM.34.3070<sub>652-822</sub>** demonstrou-se insolúvel (Figura 31). A insolubilidade de **LmxM.34.3070<sub>652-822</sub>** nesse experimento foi atribuída à alta temperatura de indução bem como à alta concentração de IPTG. A temperaturas menores, proteínas podem ser expressas com maior solubilidade devido à desaceleração das taxas de transcrição e tradução. Nestas condições a probabilidade de incorporação incorreta de aminoácidos na sequência da proteína recombinante diminui, assim como a proteína possui mais tempo para se enovelar corretamente (Francis & Page, 2010). Pelas mesmas razões, concentrações menores de indutor podem reduzir a formação de agregados da proteína



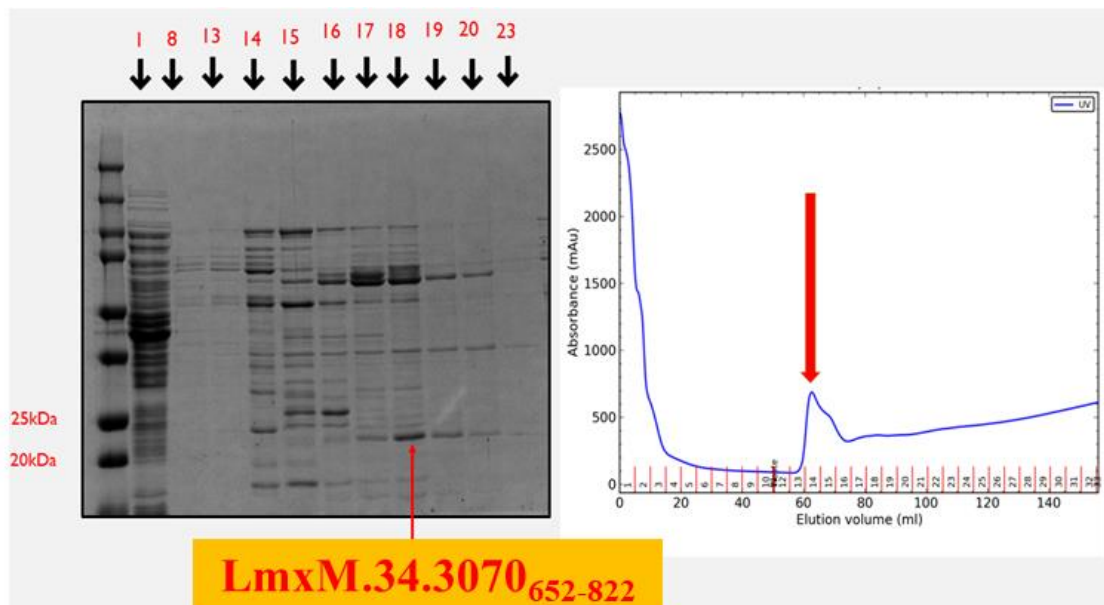
recombinante em *E.coli* (Gomes et al., 2020).

**Figura 31:** Perfil eletroforético do primeiro teste de lise de células RIP-L transformadas com **LmxM.34.3070<sub>652-822</sub>** + pET28a (~21 kDa). As células foram induzidas com 600  $\mu$ M de IPTG à 37°C. O pellet foi ressuscitado em tampão Tris 50mM pH = 8, NaCl 200 mM. As células foram lisadas por sonicação e aplicadas em gel para fracionamento por SDS-PAGE. NI= Não induzido, Ind= induzido, Lys= lisado, Insol= insolúvel. Nessas condições, o fragmento **LmxM.34.3070<sub>652-822</sub>** encontrou-se na porção insolúvel.

Alternativamente, a precipitação do fragmento **LmxM.34.3070<sub>652-822</sub>** também pode ter ocorrido devido ao método de lise celular utilizado, a sonicação. Sonicação é o processo de ruptura da membrana celular por aplicações de ondas ultrassônicas. É descrito na literatura que a lise por sonicação pode aumentar a formação de agregados proteicos devido a alterações na estrutura nativa de proteínas (Stathopoulos et al., 2008). Portanto, substituímos este método de lise celular pela *French Press*. Além disso, pode ser que o tampão de lise utilizado não seja adequado para a melhor solubilidade da proteína. Neste mesmo experimento realizamos a suspensão do *pellet* de células em Tampão A da cromatografia de afinidade como descrito no teste 2 da Tabela 4.1.

No teste 2 (Tabela 4.1) conseguimos expressar o fragmento **LmxM.34.3070<sub>652-822</sub>** de forma parcialmente solúvel com quatro modificações do protocolo: a) diminuindo a concentração de IPTG para 300  $\mu$ M, b) diminuindo a temperatura de indução para 18°C, c) lisando as células com uma *French press* e d) alterando o tampão de lise. O aumento da solubilidade de **LmxM.34.3070<sub>652-822</sub>** foi evidenciado pela presença de um pico no cromatograma da cromatografia de afinidade e pelo aparecimento de uma banda com massa molecular consistente com **LmxM.34.3070<sub>652-822</sub>** no perfil eletroforético das frações componentes deste pico no gel de SDS-PAGE (Figura 32). No entanto, os resultados também demonstraram que nessas condições não estávamos obtendo sucesso na purificação do fragmento **LmxM.34.3070<sub>652-822</sub>** por cromatografia de afinidade devido ao alto grau de contaminação nas frações e o baixo rendimento na porção solúvel (Figura 32).





**Figura 32:** À esquerda: Perfil eletroforético da primeira segunda tentativa (Teste 2 Tabela 4.1) de purificação do fragmento **LmxM.34.3070<sub>652-822</sub>** (21 kDa) onde as frações de interesse (números em vermelho acima do gel) foram avaliadas por SDS-PAGE. O *pellet* de células RIP-L induzidas por 300  $\mu$ M de IPTG à 18°C foi ressuspensionado em tampão de afinidade A (Imidazol 20 mM, NaCl 300 mM, NaPi 50mM, Beta-mercaptoetanol 1 mM, pH=7.2) e a lise ocorreu por *French Press*. À direita: Cromatograma de purificação do fragmento **LmxM.34.3070<sub>652-822</sub>** através de cromatografia de afinidade em coluna HiTrap carregada com Ni<sup>2+</sup> no teste 2. A proteína foi eluída durante um gradiente linear de 20 a 500 mM de imidazol em 140 mL. Nota-se a presença de um pico de absorvância entre as frações 14-17 (seta vermelha). O perfil eletroforético das frações demonstra bandas finas com massas moleculares consistentes com o fragmento **LmxM.34.3070<sub>652-822</sub>** na fração solúvel. Entretanto, o grau de pureza obtido nesta tentativa de purificação estava baixo.

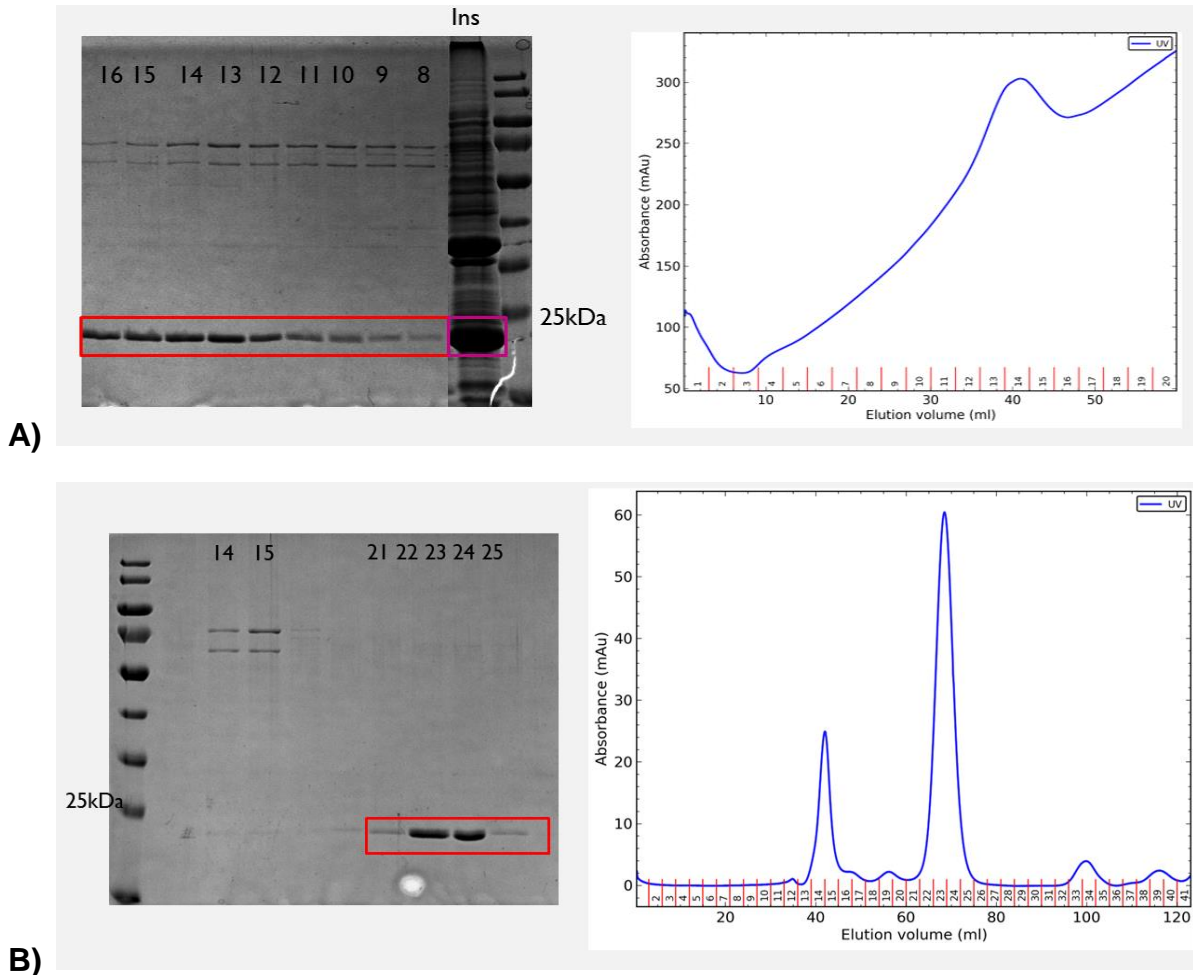
#### 4.10. Teste de aditivos no tampão de lise e otimização da cromatografia de afinidade

Outra estratégia implementada em nossas triagens foi o uso de aditivos no tampão de lise que poderiam aumentar a solubilidade do fragmento **LmxM.34.3070<sub>652-822</sub>**. O fragmento **LmxM.34.3070<sub>652-822</sub>** possui uma alta porcentagem de aminoácidos com cadeias laterais apolares preditas para compor um domínio com um *core* interno de carácter hidrofóbico. Proteínas ricas em resíduos apolares tendem a formar agregados se não se encontram bem enoveladas (Francis & Page, 2010). Detergentes como TWEEN 20 e Triton são comumente utilizados em baixas concentrações para reduzir a formação de corpos de inclusão (S. M. Singh & Panda, 2005) (A. Singh et al., 2015). Foi sugerido que o mecanismo de ação dos detergentes na solubilização de proteínas consiste na formação de

micelas ao redor de proteínas não enoveladas ou parcialmente enoveladas, onde tais micelas mimetizam sistemas de "refolding" presentes nas células, como as chaperonas (S. M. Singh & Panda, 2005) (A. Singh et al., 2015). Portanto, na fase de experimentos do Teste 3 (Tabela 4.1) adicionamos o detergente TWEEN 20 na concentração de 0.01% na solução de lise. À esta solução também foram adicionados sacarose, sabidamente conhecida por sua capacidade reduzir a formação de agregados (A. Singh et al., 2015), e lisozima para aumentar a eficiência da lise celular.

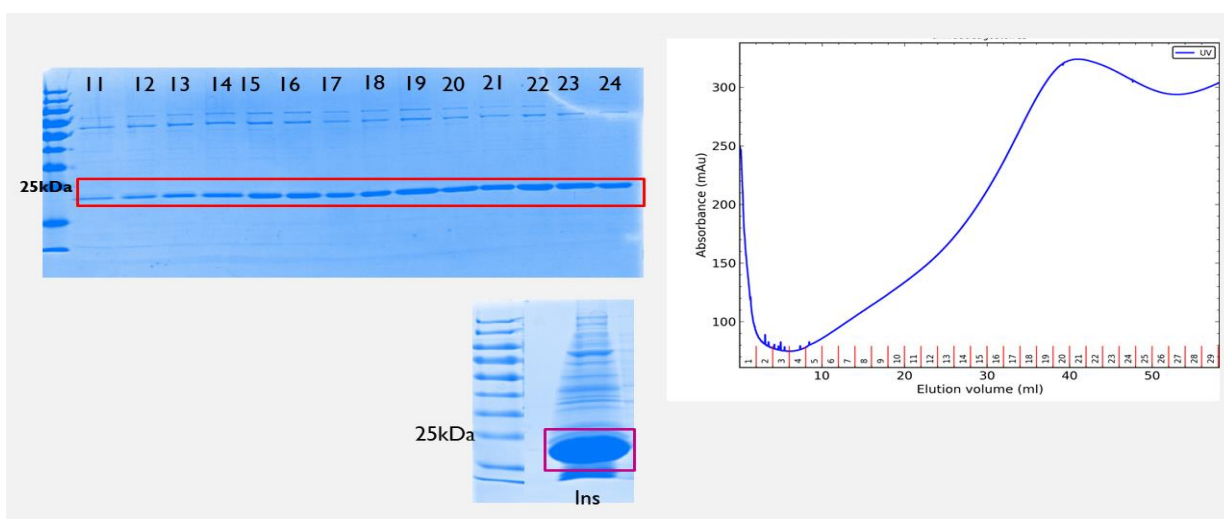
Neste experimento (Teste 3, Tabela 4.1), as concentrações de Imidazol e NaCl também foram alteradas. No tampão de lise e no tampão A da afinidade a concentração final de Imidazol foi aumentada para 30 mM. Nosso intuito ao aumentar a concentração de Imidazol nas soluções e de lise e no tampão A da cromatografia de afinidade foi induzir a competição pelos sítios de interação com o níquel e reduzir a retenção de proteínas inespecíficas na coluna HiTrap. Adicionalmente, a concentração de Imidazol no tampão B foi reduzida para 250 mM para aumentar a resolução da separação e assim diminuir a co-eluição do fragmento **LmxM.34.3070<sub>652-822</sub>** com contaminantes durante o gradiente linear de imidazol. Além das mudanças nas concentrações de imidazol, a concentração de NaCl nos tampões A e B da afinidade foi aumentada para 400 mM para reduzir possíveis interações eletrostáticas de proteínas inespecíficas com o níquel. Nessas condições, progredimos na purificação do fragmento **LmxM.34.3070<sub>652-822</sub>** por cromatografia de afinidade, onde alcançamos bons níveis de pureza nas frações de interesse que foram analisadas por SDS-PAGE (Figura 33A à esquerda). Seguimos com a exclusão molecular (GF1- Tabela 5) com as frações de interesse que eluíram dessa cromatografia de afinidade. No entanto, apesar de ter sido alcançado um melhor nível de pureza e da monodispersidade da amostra, o rendimento da preparação foi baixo conforme pode ser verificado pela absorbância da fração 23, equivalente a 60 mAu (Figura 33B à direita). As frações 21, 22,

23, 24 e 25 (Figura 33B à direita) foram concentradas até 500  $\mu$ L em membranas Amicon e o rendimento final, avaliado por espectrofotometria a 280 nm, foi de 0.4 mg de proteína purificada a partir de 0,5 L de LB. Em resumo, ainda estávamos com um problema de baixa solubilidade do fragmento **LmxM.34.3070<sub>652-822</sub>**.

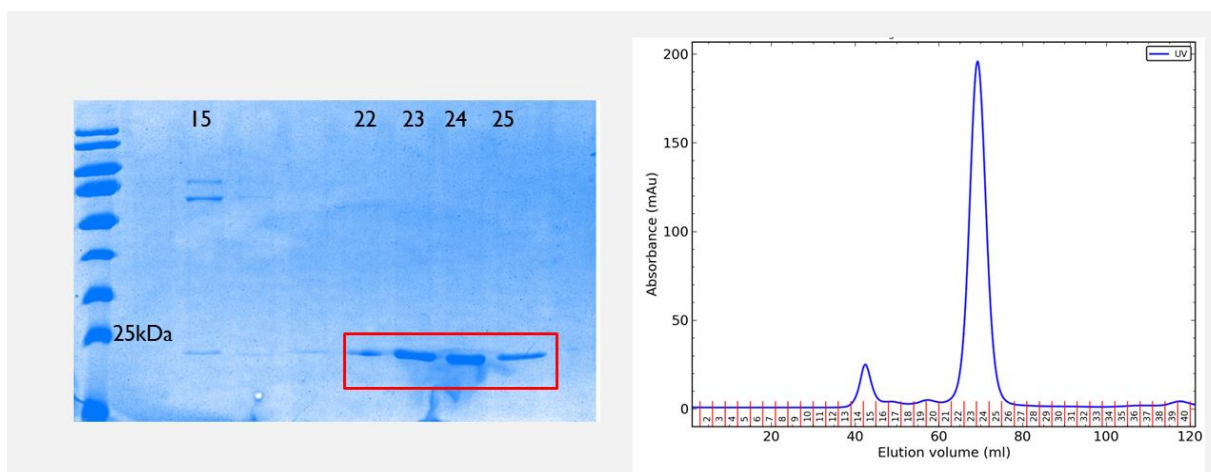


**Figura 33:** **A)** Perfil eletroforético (à esquerda) e cromatograma da purificação (à direita) do fragmento **LmxM.34.3070<sub>652-822</sub>** recombinante (~21 kDa) através de cromatografia de afinidade no Teste 3 (tabela 4.1) em coluna HiTrap carregada com  $\text{Ni}^{2+}$ . A lise ocorreu na presença do detergente TWEEN 20. A proteína foi eluída durante um gradiente linear de 30 a 250 mM de imidazol em 12 volumes de coluna (60 mL). As frações de interesse foram avaliadas por SDS-PAGE. Em vermelho esta destacada a fração solúvel de fragmento **LmxM.34.3070<sub>652-822</sub>** que eluiu na cromatografia de afinidade. INS= Uma amostra da fração insolúvel também aplicada no gel. A banda espessa do fragmento **LmxM.34.3070<sub>652-822</sub>** mostra que, nessas condições, o fragmento encontra-se majoritariamente na porção insolúvel (indicada por uma caixa rosa). **B)** Perfil eletroforético e cromatograma da etapa de purificação do fragmento **LmxM.34.3070<sub>652-822</sub>** recombinante (~21 kDa) por exclusão molecular (experimento GF1. Tabela 5) em coluna Superdex S75 16/600. A proteína foi eluída isocraticamente em tampão de exclusão molecular (Tris 50 mM pH = 8, NaCl 200 mM, Glicerol 1%, B-mercaptoetanol 1 mM, EDTA 1mM). As frações de interesse foram avaliadas por SDS-PAGE. As bandas do fragmento **LmxM.34.3070<sub>652-822</sub>** estão indicadas em vermelho.

No teste 4 (ver Tabela 4.1) fizemos uso conjugado de TWEEN 20 e Triton-X, ambos na concentração de 0.03% na solução de lise. Também fizemos o uso de L-aminoácidos de arginina e glutamina nesta solução. Esses aminoácidos em solução reduzem a formação de agregados na amostra pois tendem a solvatar proteínas (Tsumoto et al., 2004) e com isso, rompem as interações proteína-proteína associadas à formação de corpos de inclusão. Nestas condições, obtivemos um aumento da absorbância demonstrada pelo cromatograma de afinidade (Figura 34) e um melhor rendimento final do fragmento **LmxM.34.3070<sub>652-822</sub>** após a exclusão molecular, equivalente a 1.3 mg a partir de 0,5 L de LB (Figura 35 à direita).



**Figura 34:** Acima: Perfil eletroforético e cromatograma da quarta tentativa de purificação do fragmento **LmxM.34.3070<sub>652-822</sub>** recombinante (~21 kDa) através de cromatografia de afinidade (Teste 4, tabela 5) em coluna HiTrap carregada com Ni<sup>2+</sup>. O tampão de lise continha 0.03% TWEEN 20, 0.03% Triton-X, 50 mM de L-Arginina e 50 mM de L-Glutamina. O fragmento **LmxM.34.3070<sub>652-822</sub>** foi eluído durante um gradiente linear de 30 a 250 mM de imidazol em 10 volumes de coluna (50 mL). As frações de interesse foram avaliadas por SDS-PAGE. Em vermelho está destacada a fração solúvel de **LmxM.34.3070<sub>652-822</sub>** que eluiu na cromatografia de afinidade. Abaixo: INS= Uma amostra da fração insolúvel também analisada SDS-PAGE. Na caixa rosa encontra-se a banda correspondente a **LmxM.34.3070<sub>652-822</sub>** demonstrando que mesmo nessas condições, o fragmento também se encontrava majoritariamente na porção insolúvel.



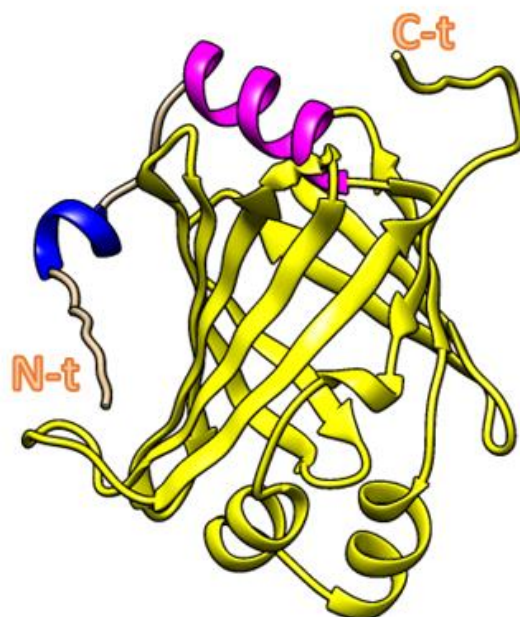
**Figura 35:** Perfil eletroforético (à esquerda) e cromatograma (à direita) da etapa de purificação do fragmento **LmxM.34.3070<sub>652-822</sub>** recombinante (~21 kDa) através de cromatografia de exclusão molecular (experimento GF2, Tabela 5) na coluna Superdex S75 16/600. O fragmento **LmxM.34.3070<sub>652-822</sub>** foi eluído isocraticamente em tampão de exclusão molecular (Tris 50 mM, NaCl 200 mM, Glicerol 1%, B-mercaptoetanol 1 mM, EDTA 1mM, pH= 8) em um volume de coluna (120 mL). As frações de interesse foram avaliadas por SDS-PAGE. Em vermelho esta destacada a fração solúvel de **LmxM.34.3070<sub>652-822</sub>** que eluiu na cromatografia por exclusão molecular. Foi alcançado um bom nível de pureza e uma boa resolução do cromatograma demonstrada por picos monodispersos, a fração 23 representativa da eluição de fragmento **LmxM.34.3070<sub>652-822</sub>** apresentou uma absorbância equivalente a 200 mAu. Na presença de aditivos, alcançamos um maior rendimento do fragmento **LmxM.34.3070<sub>652-822</sub>** na porção solúvel. As frações 22, 23, 24 e 25 foram concentradas em 500  $\mu$ L em membranas Amicon e o rendimento final, avaliado por espectrofotometria a 280 nm, foi de 1.3 mg por 0,5 L de LB.

Mesmo com o uso de vários aditivos no tampão de lise durante as triagens de condições, o fragmento **LmxM.34.3070<sub>652-822</sub>** continuava a ser encontrado abundante na fração insolúvel do lisado clarificado (ver Figura 34). Nossa hipótese para este problema de insolubilidade agora voltava-se para o desenho da construção, uma vez que o fragmento **LmxM.34.3070<sub>652-822</sub>** poderia não estar abarcando todas as regiões necessárias para o *folding* adequado de LeiFABP, gerando instabilidade estrutural e, conseqüentemente, a sua agregação.

Alternativamente, hipotetizamos que a inserção de regiões extras na construção predita para codificar o apenas o domínio FABP-like poderia estar perturbando a aquisição da estrutura nativa. Portanto, resolvemos testar as outras construções. Para isso desenhamos o fragmento **LmxM.34.3070<sub>669-822</sub>** que corresponde a região predita para possuir a topologia clássica de uma FABP uma vez que não contém as  $\alpha$ -hélices na região N-terminal do barril (Figura 17C), e o fragmento **LmxM.34.3070<sub>638-822</sub>** que corresponde ao barril-beta FABP-like e contém as duas  $\alpha$ -hélices N-terminais preditas nas modelagens *in silico* (Figura 17A).

#### 4.11. As hélices N-terminais adicionais previstas por PSIPRED e ColabFold são essenciais para expressão do domínio C-terminal de LmxM34.3070 solúvel

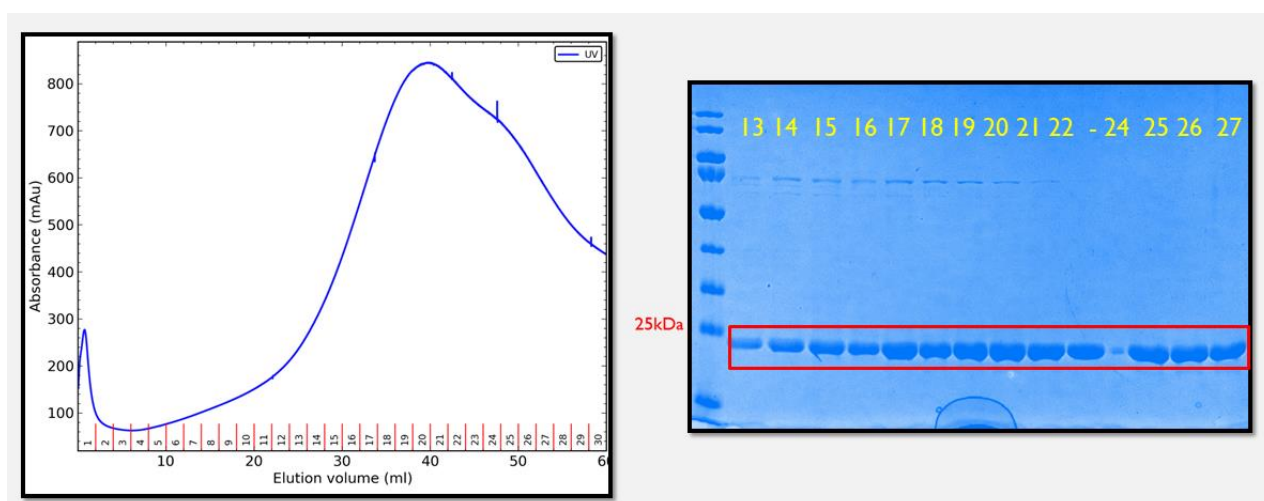
As modelagens *in silico* realizadas com PHYRE2 e Raptor-X e utilizando a sequência de aminoácidos de Lmx.M.34.3070 como *input* sugeriram a presença de um domínio na região C-terminal com estrutura semelhante às FABPs. As FABPs possuem uma topologia conservada consistindo de um barril-beta formado por 10 fitas-beta antiparalelas e um motivo hélice-alça-hélice em uma das extremidades do barril (Chmurzyńska, 2006). Análises da sequência primária de LmxM.34.3070 com ColabFold e PSIPRED sugeriram a presença de duas hélices adicionais na região N-terminal ao domínio FABP-like (Figuras 16, 17 e 36).



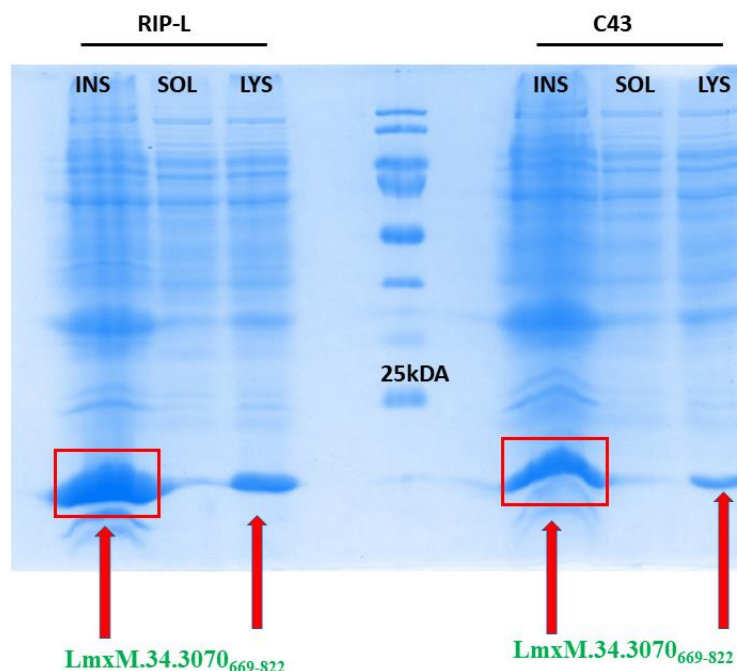
**Figura 36:** Modelo predito por Colab Fold para a região C-terminal de LmxM.34.3070 mostrando em amarelo a presença do motivo hélice-alça-hélice e fitas beta antiparalelas que formam um barril (topologia FABP-like) e as hélices N-terminal adicionais previstas por ColabFold e PSIPRED. **Em magenta:** hélice N-terminal 1. **Em azul:** hélice N-terminal 2. A construção **LmxM.34.3070<sub>669-822</sub>** corresponde apenas ao domínio FABP-like (amarelo); **LmxM.34.3070<sub>652-822</sub>**, expresso de forma parcialmente solúvel, corresponde ao anterior + hélice N-terminal 1; **LmxM.34.3070<sub>638-822</sub>**, expresso solúvel corresponde ao anterior + hélice N-terminal 2.

Testes de expressão e purificação do fragmento **LmxM.34.3070<sub>638-822</sub>** (teste 5, tabela 4.2) , que consiste no barril-beta FABP-like e todas as hélices N-terminal (ver Figura 36), revelaram um aumento expressivo de solubilidade em relação à construção

**LmxM.34.3070<sub>652-822</sub>**. Este aumento de solubilidade foi observado na cromatografia de afinidade, que demonstrou um pico significativo com alta amplitude, equivalente a 800 mAu (Figura 37 à esquerda). Quando amostras das frações que correspondiam a este pico foram preparadas e aplicadas em gel SDS-PAGE, observou-se bandas espessas de tamanho consistente com **LmxM.34.3070<sub>638-822</sub>** (Figura 37 à direita). Em contraste, o fragmento **LmxM.34.3070<sub>669-822</sub>**, que consiste apenas no barril FABP-like sem hélices N-terminais, possui baixa solubilidade (Figura 38).

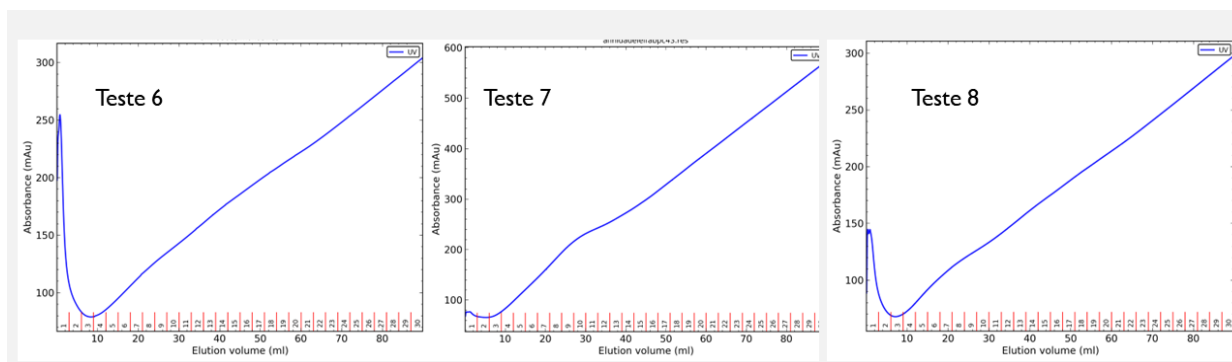


**Figura 37:** Cromatograma (à esquerda) e perfil eletroforético (à direita) da purificação do fragmento **LmxM.34.3070<sub>638-822</sub>** recombinante (~22 kDa) através de cromatografia de afinidade em coluna HiTrap carregada com  $\text{Ni}^{2+}$ . A lise ocorreu na presença de 0.03% de TWEEN 20, 0.03% de Triton-X, 50mM L-Arginina e 50mM L-Glutamina (teste 5, tabela 4.2). O fragmento **LmxM.34.3070<sub>638-822</sub>** foi eluído durante um gradiente linear de 30 a 250 mM de imidazol em 12 volumes de coluna (60 mL). As frações 13-27 que compõem o pico observado no cromatograma foram preparadas e submetidas a SDS-PAGE. Observa-se bandas espessas nas frações 13-27 (caixa vermelha) com massas consistentes com o fragmento **LmxM.34.3070<sub>638-822</sub>**.



**Figura 38:** SDS-PAGE do teste de solubilidade de células RIP-L e C43 transformadas com **LmxM.34.3070<sub>669-822</sub>**+ pET28a (~19 kDa). As células foram induzidas com 300  $\mu$ M de IPTG à 16°C *overnight*. A lise ocorreu na presença de 0.03% de TWEEN 20, 0.03% de Triton-X, 50mM de L-Arginina e 50mM L-Glutamina. As células foram lisadas por *French Press* e clarificadas por centrifugação. As amostras Lys= lisado, Sol= solúvel e Ins= insolúvel foram aplicadas no gel para fracionamento. O fragmento **LmxM.34.3070<sub>669-822</sub>** encontrou-se na porção insolúvel.

A baixa solubilidade do fragmento **LmxM.34.3070<sub>669-822</sub>** verificada no teste de solubilidade (Figura 38) foi corroborada em três experimentos (testes 6, 7 e 8 na tabela 4.2 e na tabela 4.3) independentes de purificação deste fragmento. A ausência de picos com intensidade significativa na cromatografia de afinidade (Figura 39), confirmou que o fragmento **LmxM.34.3070<sub>669-822</sub>** está ausente da fração solúvel do lisado.



**Figura 39:** Cromatogramas das três tentativas de purificação do fragmento **LmxM.34.3070<sub>669-822</sub>** recombinante (~19 kDa) através de cromatografia de afinidade em coluna HiTrap carregada com  $\text{Ni}^{2+}$ . Estes experimentos

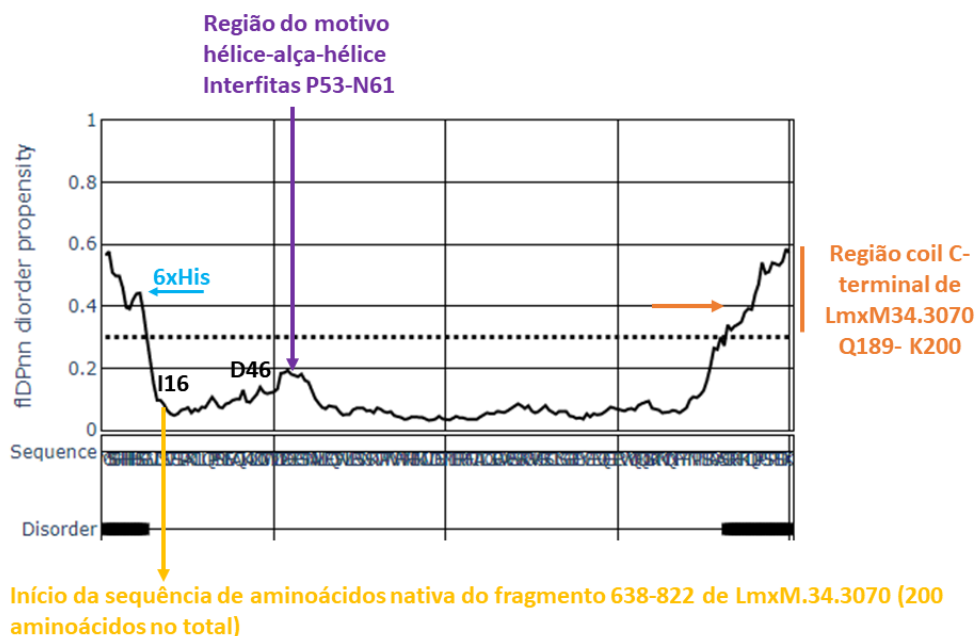


correspondem aos Testes 6, 7 e 8, (Tabela 4.2 e Tabela 4.3). A lise ocorreu na presença de 0.03% de TWEEN 20, 0.03% de Triton-X, 50mM de L-Arginina e 50mM L-Glutamina. Nos testes 6 e 7 a expressão ocorreu em RIP-L. No teste 6 foi realizado um gradiente de 30-250 mM de Imidazol, e no teste 7 o gradiente foi realizado entre 30-500 mM de Imidazol. No Teste 8 a expressão foi realizada em C43, uma cepa de *E. coli* engenheirada para tolerar a expressão de proteínas tóxicas. O gradiente foi realizado entre 30-250 mM. Não foi observado pico no cromatograma em nenhum dos três experimentos.

Os dados de expressão e solubilidade dos três fragmentos sugerem que as duas hélices em N-terminal são importantes para que o domínio LeiFABP de *L. amazonensis* adquira sua estrutura nativa e assim, ganhe solubilidade. A presença de hélices na região N-terminal já foi descrita em membros da família das FABPs (Flower, 1996). Todavia, diversos estudos estruturais demonstram que a topologia das FABPs é fortemente conservada em diferentes organismos, e corresponde apenas ao barril-beta e ao motivo hélice-alça-hélice na alça AB (Chmurzyńska, 2006) (Ono, 2005)(Smathers & Petersen, 2011)(Y. Zhang et al., 2020). FABPs nunca foram descritas em *Leishmania*.

Para verificar se o trecho N-terminal do fragmento **LmxM.34.3070<sub>638-822</sub>** poderia se comportar de forma independente da região FABP-like, submetemos a sequência deste fragmento **LmxM.34.3070<sub>638-822</sub>** para uma predição de desordem no servidor fIDPnn (Hu et al., 2021). Esta análise sugeriu alta desordem para a região compreendida entre os resíduos M1-V15 do fragmento recombinante **LmxM.34.3070<sub>638-822</sub>**, o que é esperado pois este trecho é composto pela *tag* de histidinas em N-terminal e os resíduos adicionalmente traduzidos a partir de códons extras inseridos na ORF do fragmento recombinante pelo vetor de expressão (Seta azul na figura 40). Para a cauda C-terminal (resíduos Q189-K200) também foi prevista desordem (seta laranja na figura 40), o que também é esperado, uma vez que nesta região não foi predita nenhum tipo de estrutura secundária em nossas modelagens *in silico*. Todavia, para a região compreendida entre os resíduos I16 (onde de fato se inicia a sequência que pertence a proteína LmxM.34.3070) e D46, que corresponde às hélices N-terminal previstas por ColabFold e PSIPRED, fIDPnn prevê baixa desordem (Figura 40). Devido à proximidade desta região com restante do domínio, provavelmente estas hélices interagem com o barril

FABP-like, contribuindo para o seu enovelamento e aquisição de estrutura nativa. Um leve aumento de desordem é observado no trecho compreendido entre os resíduos P53-N61 (seta violeta na figura 40), que nos modelos *in silico*, é predito para compor o motivo hélice-alça-hélice, sugerindo uma flexibilidade nesta região, possivelmente associada à dinâmica do domínio LeiFABP com ligantes.



**Figura 40:** Predição de desordem por fIDPnn para toda a construção **LmxM34.3070<sub>638-822</sub>(pET28a)**. Para as regiões M1-V1 (seta azul) e Q189-K200 (seta laranja), fIDPnn prevê desordem, o que é esperado, pois estas regiões não apresentam estrutura secundária segundo ColabFold e PSIPRED. Todavia, para a região compreendida entre I16-D46 predita para formar as hélices N-terminal, fIDPnn prevê baixa desordem. Um leve aumento de desordem é observado nos resíduos P53-N61 (seta violeta) preditos para compor o motivo hélice-alça-hélice que está entre as duas primeiras fitas beta ( $\beta$ A e  $\beta$ B) do barril FABP-like.

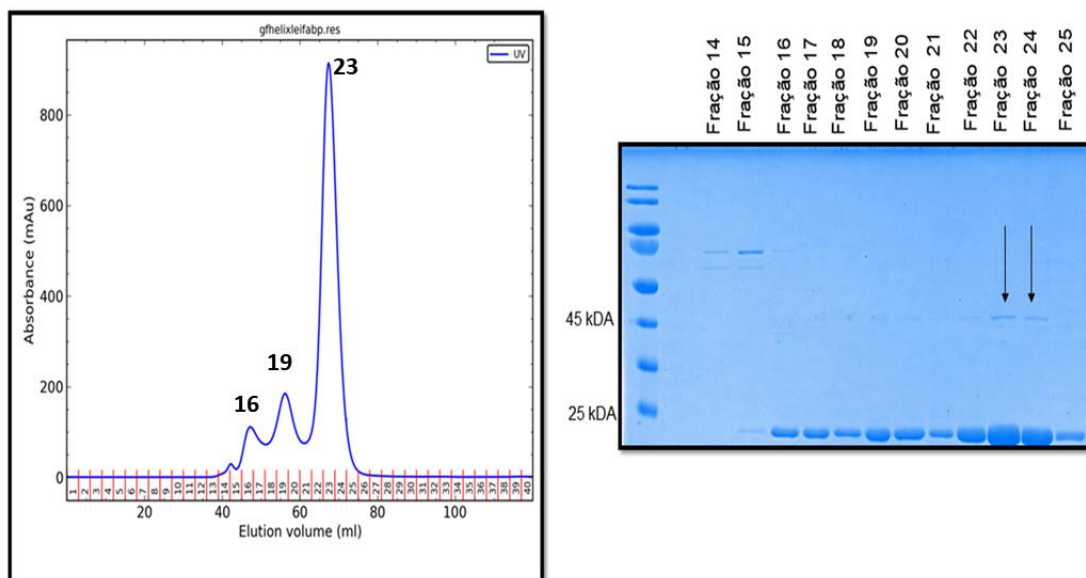
Em resumo, após a fase de testes com os diferentes fragmentos contendo LeiFABP, concluímos que o melhor fragmento para estudar o domínio LeiFABP era **LmxM.34.3070<sub>638-822</sub>** devido à sua alta solubilidade. Isso foi demonstrado por cromatografia de exclusão molecular onde foi observado um pico com uma amplitude de 900 mAu (GF3-tabela 5). A maior amplitude de absorvância obtida nas purificações por exclusão molecular com o fragmento **LmxM.34.3070<sub>652-822</sub>** foi de 200 mAu. **Diante disso, a partir de agora, o fragmento LmxM.34.3070<sub>638-822</sub> será referido como LeiFABP.**

#### 4.12. LeiFABP tende a formar oligômeros

As frações de interesse da cromatografia de afinidade do fragmento **LmxM.34.3070<sub>638-822</sub>** (ver Teste 5 tabela 4.2) foram concentradas até 500 µL com o auxílio de concentradores Amicon® Ultra 0.5 (*cutoff* 10kDa) e aplicadas em uma coluna de exclusão molecular Superdex S75 16/600 (GF3- tabela 5). A eluição foi realizada isocraticamente em tampão de exclusão molecular (Tris 50 mM pH 8, Glicerol 1%, NaCl 200 mM, EDTA 2 mM, Beta-mercaptoetanol 1 mM). Apesar de termos obtido êxito em aumentar o rendimento na porção solúvel, o equivalente a 5.65 mg por 0,5 L de LB, observamos uma baixa resolução de separação, o que é ilustrado pela sobreposição de picos no cromatograma (Figura 41, a esquerda).

A sobreposição de picos é um forte indicativo de agregação. As frações que eluíram da exclusão molecular foram avaliadas por SDS-PAGE (Figura 41 à direita). Curiosamente, os picos de menor amplitude e que eluíram primeiro (frações 16 a 19) quando analisados por SDS-PAGE mostram apenas bandas com tamanhos esperados para LeiFABP (~22kDa).

Além disso, o maior pico representado pelas frações 23 e 24 resultou em uma banda na altura de 22 kDa e em bandas inespecíficas (setas pretas do gel SDS-PAGE da Figura 41) não visualizadas anteriormente no perfil eletroforético da cromatografia de afinidade (Figura 37 à direita), com massa molecular na região de 45 kDa, o que é consistente com dímeros de LeiFABP.



**Figura 41:** Cromatograma de purificação de LeiFABP por exclusão molecular (à esquerda) em coluna Superdex S75 16/600 no experimento GF3 (tabela 5). Perfil eletroforético (à direita) em SDS-PAGE de frações separadas por gel filtração. Observa-se que as frações 16, 19 e 23, que correspondem a volumes de eluição diferentes, mostram bandas com tamanhos esperados para o fragmento LeiFABP (~22kDa). Bandas inespecíficas foram observadas nas frações 23 e 24 e são indicadas pelas setas pretas. Essas bandas não foram visualizadas anteriormente na cromatografia de afinidade.

Começamos a trabalhar com a hipótese de que LeiFABP poderia estar populando diferentes estados oligoméricos. Para testar se a sobreposição dos picos que eluíram anteriormente da coluna de gel filtração e as bandas inespecíficas observadas no gel SDS-PAGE correspondiam a oligômeros, amostras das frações 16, 19 e 23 (representantes de cada um desses picos) (cromatograma da Figura 41) foram direcionadas para experimentos de SEC-MALS.

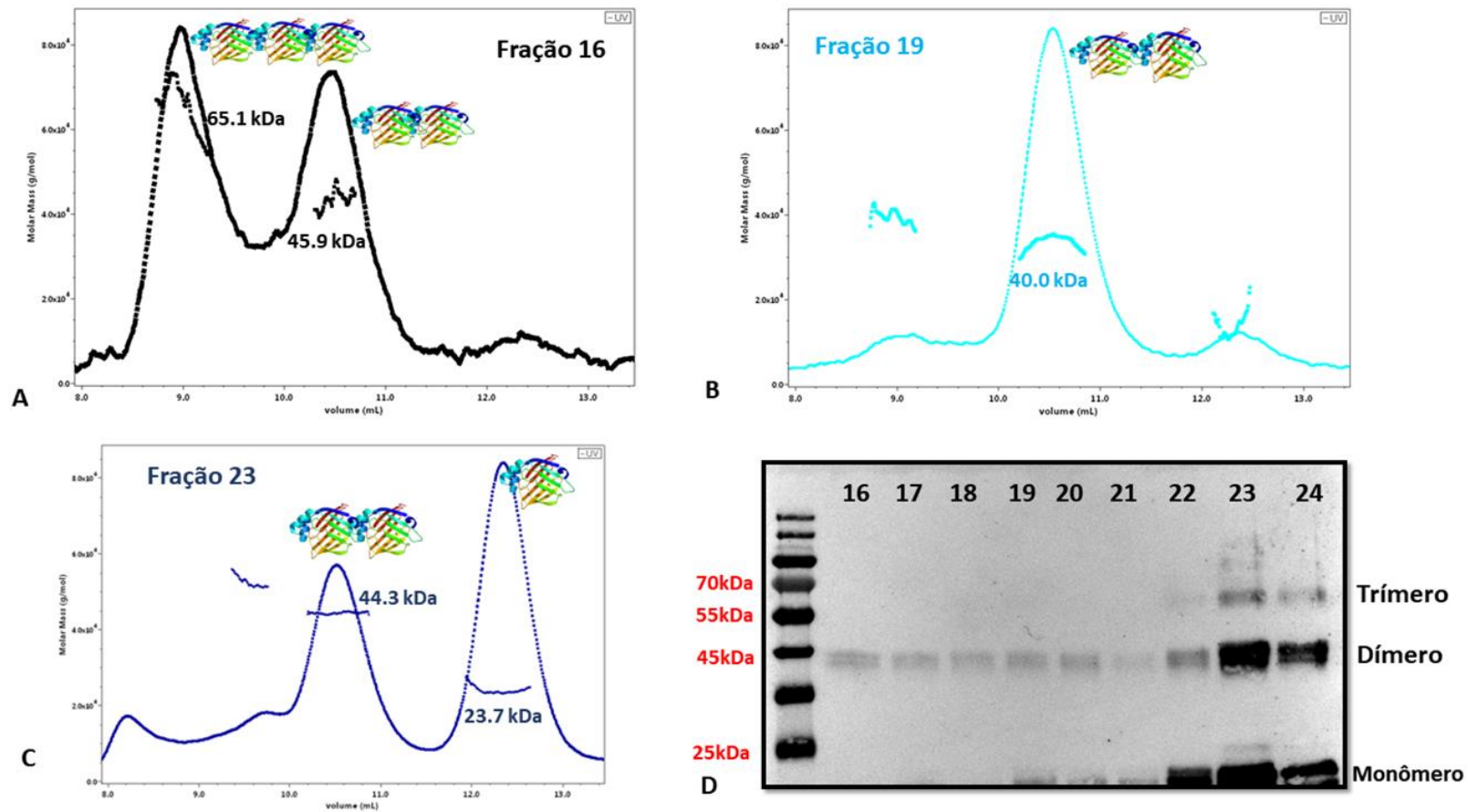
Além disso, para confirmar os dados de SEC-MALS e validar a hipótese de formação de oligômeros, realizamos um experimento de western blot com alíquotas das frações separadas por exclusão molecular. As medidas de SEC-MALS demonstraram que LeiFABP tende a formar oligômeros. As análises de espalhamento de luz a múltiplos ângulos apontaram que a fração 23 contém monômeros e dímeros uma vez que foi detectada a presença de duas espécies de massa igual a 23 e 44 kDa (Figura 42C).

Já as medidas da fração 19 mostram que ela é composta apenas por dímeros (Figura 42B), uma vez que foi detectada a presença de apenas uma espécie de aproximadamente 40

kDa. Finalmente, as medidas de SEC-MALS mostraram que a fração 16 é composta por dímeros e por trímeros, pois foram detectadas espécies de aproximadamente 66 e 45 kDa (Figura 42A). Essas medidas são consistentes com o perfil cromatográfico de exclusão molecular (ver figura 41).

Essa análise foi validada por western blot. A revelação da membrana transferida com amostras das frações deste experimento de exclusão molecular mostrou bandas consistentes com monômeros, dímeros e trímeros de LeiFABP que foram reconhecidas pelo anticorpo (Figura 42D). Com isso, concluímos que LeiFABP recombinante tende a formar oligômeros em solução.

Apesar do experimento de western blot ter sido realizado em condições desnaturantes, acreditamos que a observação de dímeros e trímeros foi possível devido à alta estabilidade apresentada pelos barris-beta (Gessmann et al., 2011). Barris-beta tendem a apresentar alta estabilidade devido a compactação de sua estrutura globular e às extensas interações de hidrogênio formadas entre as fitas betas que compõem o barril (Gessmann et al., 2011). Como LeiFABP é, provavelmente, um barril-beta, acreditamos que a estabilidade da estrutura dificultou a dissociação dos dímeros e trímeros, os quais foram visualizados mesmo em vista das perturbações químicas e térmicas que ocorreram no preparo de amostras.

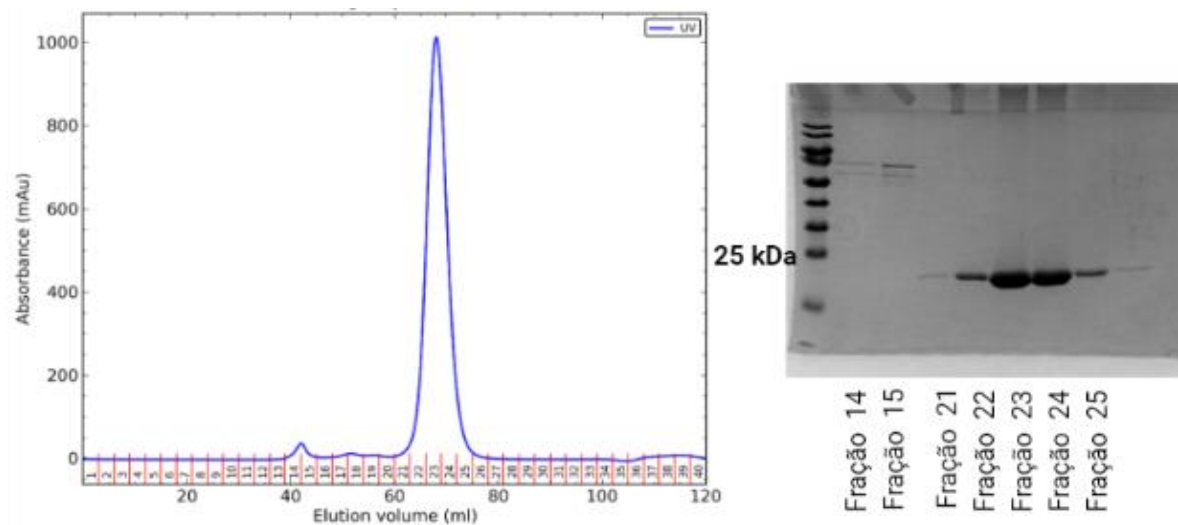


**Figura 42:** Análise do estado de agregação de LeiFABP por SEC-MALS. As frações 16, 19 e 23 que eluíram da coluna de Superdex S75 (Fig. 39) foram reaplicadas em uma coluna de exclusão Superdex 200 10/300, acoplada a um detector de espalhamento de luz a múltiplos ângulos (MALS Wyatt Technology). **A)** Medidas de SEC-MALS da fração 16 demonstrando a presença de espécies com massa de aproximadamente 65 e 46 kDa consistentes com dímeros e trímeros de LeiFABP. **B)** Medidas de SEC-MALS da fração 19 demonstrando a presença de uma espécie com massa de 40 kDa consistente com dímeros de LeiFABP. **C)** Medidas de SEC-MALS da fração 23 demonstrando a presença de espécies com massas de 44,3 e 23,7 kDa consistentes com dímeros e monômeros de LeiFABP respectivamente. **D)** Western blot com anticorpo Anti-His. As frações de interesse (16-24) obtidas na exclusão molecular (Figura 41) foram submetidas a SDS-PAGE e transferidas para membrana de nitrocelulose. Em seguida, a membrana foi incubada com anticorpo monoclonal Anti-His conjugado a fosfatase alcalina e revelada com adição BCIP-NBT. A revelação da membrana mostra bandas com tamanhos consistentes com monômeros, dímeros e trímeros de LeiFABP.

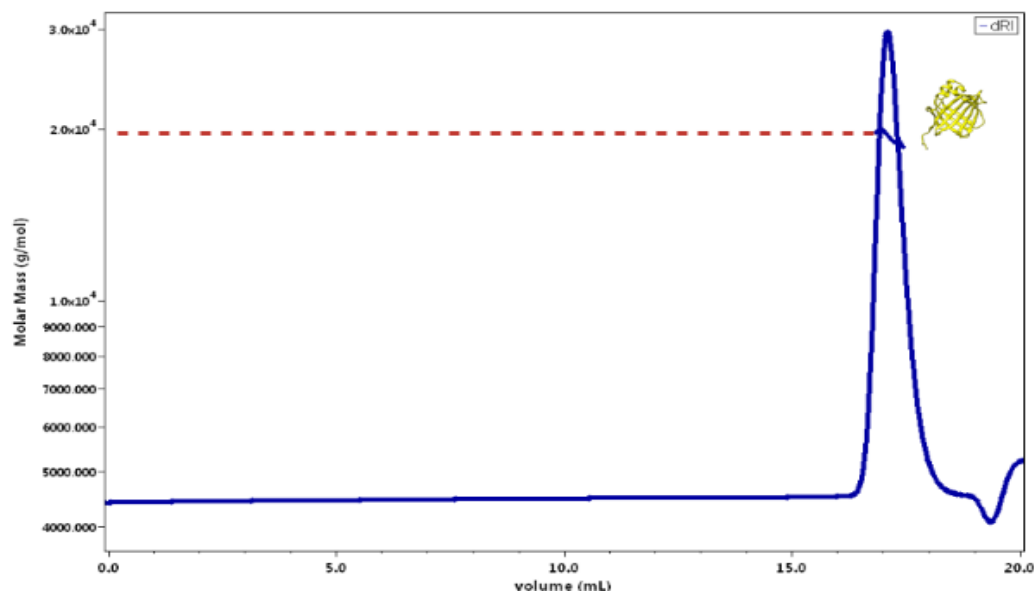
### 4.13. Obtendo amostras monoméricas de LeiFABP

Estávamos diante de um problema, pois amostras polidispersas poderiam atrapalhar a aquisição de dados de RMN e a cristalização dessa construção. Neste caso, para contornar a formação de oligômeros, aumentamos a força iônica e a concentração de agente redutor no tampão de exclusão molecular para evitar a formação de interações intermoleculares de natureza eletrostática e a formação de pontes dissulfeto intermoleculares (Francis & Page, 2010). A concentração final de agente redutor Beta-mercaptoetanol foi aumentada para 10 mM (GF4-tabela 5), e a força iônica, determinada pela concentração final de NaCl, foi alterada de 200 para 400 mM (GF4 - tabela 5).

A purificação por exclusão molecular de LeiFABP nestas novas condições demonstrou a eluição de apenas um pico monodisperso da coluna de gel filtração (cromatograma da Figura 43). As medidas de SEC-MALS confirmaram a presença de uma única espécie com massa molecular correspondente ao monômero da LeiFABP (Figura 44). Além disso, verificamos um bom grau de pureza de LeiFABP por SDS-PAGE (gel SDS-PAGE da Figura 43). O rendimento desta preparação, conforme medidas de absorvância em 280 nm foi de 4.56 mg por 0,5 L de LB. Dessa forma, conseguimos adquirir amostras homogêneas, puras e solúveis, e, portanto, boas candidatas para ensaios de biologia estrutural.



**Figura 43: Cromatograma** e perfil eletroforético da etapa de purificação do domínio LeiFABP por exclusão molecular em coluna Superdex S75 16/600 em GF4 (tabela 5). O pico monodisperso de maior amplitude é representado pela fração 23. O perfil eletroforético demonstra um bom nível de pureza, onde foi observada apenas uma banda de massa molecular consistente com LeiFABP.



**Figura 44:** Análise do estado de agregação de LeiFABP por SEC-MALS. A fração 23 que eluiu da coluna Superdex S75 (Fig. 38) no experimento GF4 (tabela 5) foi aplicada em uma coluna de exclusão por tamanho Superdex 200 10/300, acoplada a um detector de espalhamento de luz a múltiplos ângulos. As medidas de SEC-MALS da fração 23 demonstraram a presença de apenas uma espécie com massa ~20 kDa, que é consistente com o monômero de LeiFABP (~22kDa)



**Tabela 8: Resumo dos resultados envolvendo testes de solubilidade, perfil cromatográfico da exclusão molecular, solubilidade e rendimento final com os três fragmentos construídos para estudos estruturais de LeiFABP: **LmxM.34.3070**<sub>652-822</sub>, **LmxM.34.3070**<sub>638-822</sub> e **LmxM.34.3070**<sub>669-822</sub>.**

Condições de expressão em cada experimento	Clone	Cromatogramas na exclusão molecular	Solubilidade	Rendimento Final por 0,5 L de LB (mg)
300 µM de IPTG OD= 0.6-0.8 Temperatura de indução = 16 °C	<b>LmxM.34.3070</b> <sub>652-822</sub>	Bom	Parcialmente solúvel	0.4
300 µM de IPTG OD= 0.6-0.8 Temperatura de indução = 16 °C	<b>LmxM.34.3070</b> <sub>652-822</sub>	Bom	Parcialmente solúvel	1.3
300 µM de IPTG OD= 0.6-0.8 Temperatura de indução = 16 °C	<b>LmxM.34.3070</b> <sub>638-822</sub>	Parcialmente bom, mas com sobreposição de picos e formação de oligômeros	Solúvel	5.65
300 µM de IPTG OD= 0.6-0.8 Temperatura de indução = 16 °C	<b>LmxM.34.3070</b> <sub>638-822</sub>	<b>Muito bom, com um pico monodisperso de alta amplitude e sem formação de oligômeros</b>	Solúvel	4.56
300 µM de IPTG OD= 0.6-0.8 Temperatura de indução = 16 °C	<b>LmxM.34.3070</b> <sub>669-822</sub>	Não foi possível purificar este fragmento pois ele está ausente na porção solúvel do lisado	Insolúvel	-
300 µM de IPTG OD= 0.6-0.8 Temperatura de indução = 16 °C	<b>LmxM.34.3070</b> <sub>669-822</sub>	Não foi possível purificar este fragmento pois ele está ausente na porção solúvel do lisado	Insolúvel	-
300 µM de IPTG OD= 0.6-0.8 Temperatura de indução = 16 °C	<b>LmxM.34.3070</b> <sub>669-822</sub>	Não foi possível purificar este fragmento pois ele está ausente na porção solúvel do lisado	Insolúvel	-

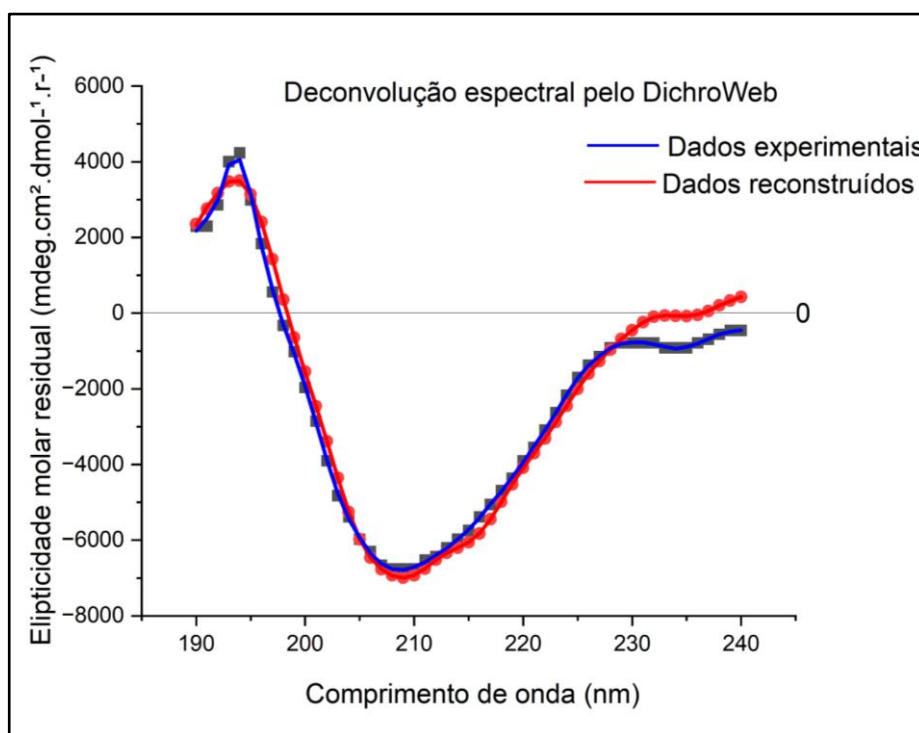
#### 4.14. LeiFABP possui estruturas do tipo Beta

A caracterização do conteúdo de estrutura secundária de LeiFABP foi realizado por espectroscopia de dicroísmo circular. A ligação peptídica absorve radiação ultravioleta na faixa de comprimentos de onda do UV distante (190 – 240 nm) e apresenta atividade óptica resultando em espectros de dicroísmo circular característicos nesta faixa (190-240 nm) (Greenfield, 2007). Por isso, utilizamos a espectroscopia de dicroísmo circular para caracterizar o conteúdo de estrutura secundária e o enovelamento de LeiFABP (Kelly et al., 2005) (Greenfield, 2007). O espectro de CD de LeiFABP apresenta perfil esperado para uma proteína com estruturas do tipo beta com leve contribuição de  $\alpha$ -hélice devido à presença de uma banda positiva em aproximadamente 195 nm, de um mínimo entre 200 e 210 nm, e um leve ombro entre 210 e 220 nm. Para realizar uma interpretação mais precisa desse espectro, realizamos a deconvolução espectral no servidor DichroWeb (Miles et al., 2022).

O DichroWeb reconstruiu os dados experimentais utilizando Continn, um algoritmo de Provencher & Gloeckner, publicado em 1982 (Provencher & Gloeckner, 1981). Resumidamente, o algoritmo interpreta o valor de elipticidade como sendo resultado da somatória dos sinais advindos da contribuição de todas as estruturas secundárias presentes na proteína analisada. Esta análise é realizada para cada ponto do intervalo do comprimento de onda utilizado na medida, uma vez que diferentes estruturas secundárias apresentam comportamento da elipticidade específico em um determinado  $\lambda$  (Provencher & Gloeckner, 1981). Os dados utilizados pelo algoritmo para esta inferência são recuperados de *datasets* contendo modelos experimentais de proteínas cujas estruturas já foram resolvidas e os espectros de CD já determinados. O *dataset* utilizado como referência nesta análise foi o SMP180T (Abdul-Gader et al., 2011) que inclui um conjunto de dados experimentais de 98 proteínas solúveis e 30 proteínas de membrana. Interessantemente, as proteínas que compõem este *dataset* são ricas em fitas-beta e possuem variadas funções, dentre elas: a absorção ativa de

nutrientes e interação com ligantes (Abdul-Gader et al., 2011), sendo esta a função a qual LeiFABP possivelmente está associado.

A análise quantitativa do espectro de CD indicou que 32% dos aminoácidos do LeiFABP estão em estruturas do tipo beta, enquanto 15% estão em hélice. Este resultado é consistente com a previsão do ColabFold, que prevê 38% de resíduos em fita-beta e 19% de resíduos em hélice (Figura 45).



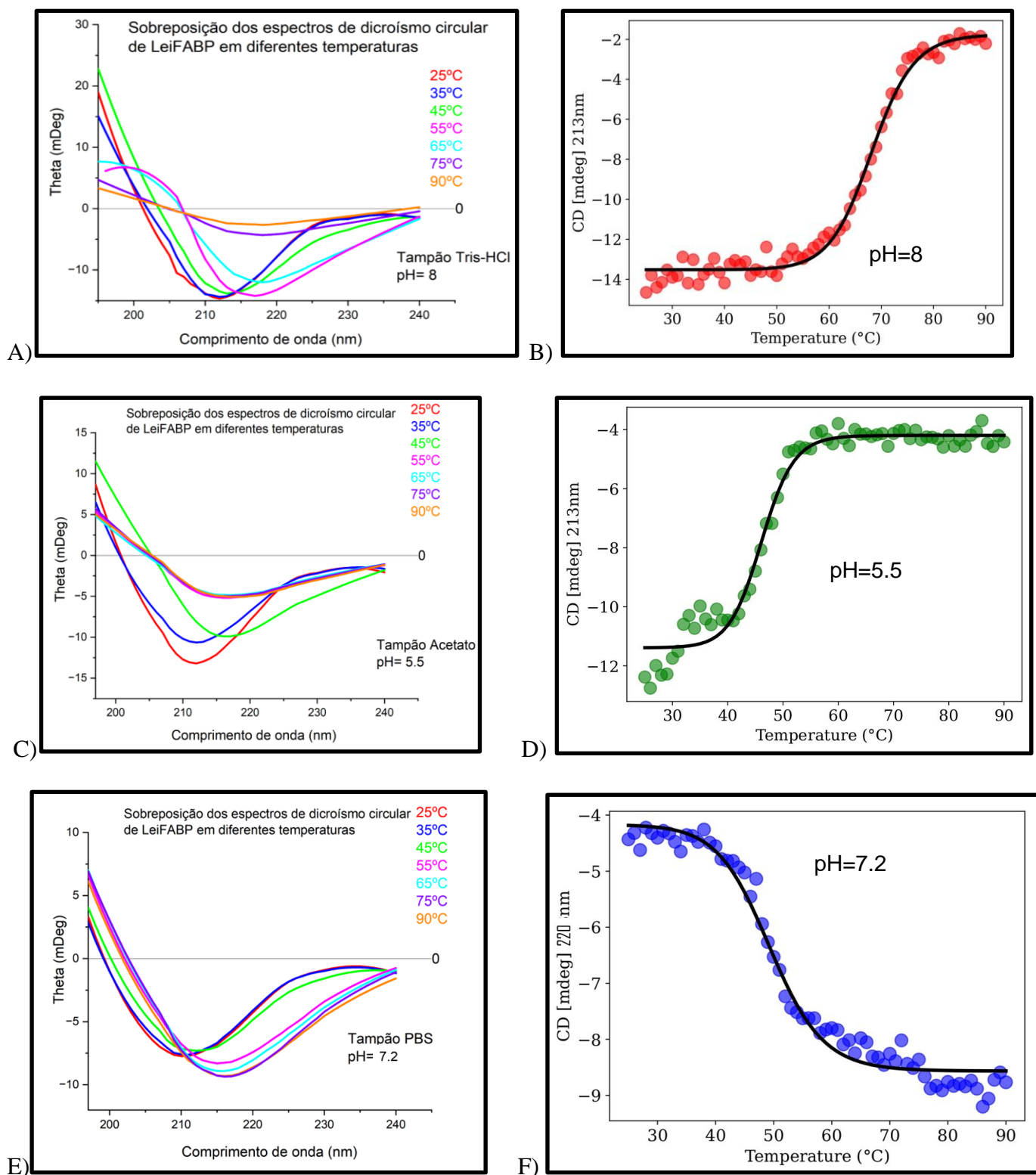
**Figura 45:** Deconvolução do espectro de CD de LeiFABP a 25 °C através do DichroWeb que reconstruiu os dados experimentais através do algoritmo Continl e utilizando o modelo SMP180t. Essa análise encontrou 32 % de resíduos em estrutura do tipo beta, e 15 % de resíduos em estrutura do tipo hélice, enquanto nossas previsões *in silico* sugerem que LeiFABP seja um barril-beta com 38% de resíduos compondo fitas-beta e 19% de resíduos compondo hélices.

#### 4.15. Influência do tampão na estabilidade térmica de LeiFABP

A espectroscopia de dicroísmo circular também foi utilizada para monitorar a desnaturação térmica de LeiFABP em diferentes condições de pH e composições de tampão. Para isso, os experimentos de desnaturação foram realizados em diferentes condições: 1) LeiFABP purificado em tampão Tris-HCl pH= 8, 2) LeiFABP purificado em tampão acetato pH= 5.5, e 3) LeiFABP purificado em tampão PBS pH= 7.2.

A curva de desnaturação de LeiFABP foi obtida através do comportamento dos valores de elipticidade registrados em  $\lambda = 213$  ou  $220$  nm com a temperatura no intervalo entre  $25$  e  $90^\circ\text{C}$ , em passos de  $1^\circ\text{C}$ . Os espectros de dicroísmo circular de LeiFABP obtidos no intervalo entre  $25$  e  $90^\circ\text{C}$  em cada condição de tampão e pH foram sobrepostos para a avaliação da estabilidade térmica e para monitoramento do seu estado de enovelamento em função da temperatura (Figura 46 A, C e E). Nas condições alcalina e ácida, a temperatura de desnaturação foi estimada através do ajuste dos dados de elipticidade registrados em  $213$  nm de acordo com a equação 5. A  $T_m$  foi estimada em  $68^\circ$  (Figura 46B) na condição alcalina, e em  $49^\circ\text{C}$  (Figura 46D) na condição ácida.

A temperatura de desnaturação de LeiFABP em pH 7.2 foi obtida através do ajuste dos dados de elipticidade registrados em  $220$  nm, cujo valor foi estimado em  $50^\circ\text{C}$  (Figura 46F). O comportamento dos espectros de CD em função da temperatura sugere que o mecanismo de desenovelamento de LeiFAB em pH 7.2 segue um equilíbrio entre dois estados. Isso pode ser observado devido à presença de um ponto isodicroico em aproximadamente  $210$  nm (Figura 46E). Em contraste, o mecanismo de desenovelamento nos pHs 5.5 e 8.0 parece ser mais complexo, possivelmente envolvendo a formação de intermediários de desenovelamento ou eventos de agregação (Figuras 46A e 46C). Em nenhum dos casos LeiFABP parece perder completamente a sua estrutura secundária a  $90^\circ\text{C}$ . Os experimentos mostraram que a estabilidade térmica deste domínio é influenciada pela composição e pelo pH do tampão.



**Figura 46:** A) Sobreposição dos espectros de dicroísmo circular de LeiFABP obtidos 25, 35, 45, 65, 75 e 90 °C em tampão alcalino Tris 50 mM pH 8.0, beta-mercaptoetanol 2 mM, NaCl 200 mM, glicerol 1 %, EDTA 2 mM. B) Comportamento da elipticidade em 213 nm em função da temperatura (pH alcalino). A T<sub>m</sub> de LeiFABP nesta condição foi estimada em 68°C. C) Sobreposição dos espectros de dicroísmo circular de LeiFABP obtidos a 25, 35, 45, 65, 75 e 90 °C em tampão Ácido Acético 5 mM, Acetato 45 mM pH 5.5, beta-mercaptoetanol 2 mM, NaCl 200 mM, glicerol 1 %, EDTA 2 mM. D) Comportamento da elipticidade em 213 nm em função da temperatura a pH ácido. A T<sub>m</sub> de LeiFABP nesta condição foi estimada em 68°C. E) Sobreposição dos espectros de dicroísmo circular de LeiFABP obtidos a 25, 35, 45, 65, 75 e 90 °C em tampão neutro PBS: fosfato de sódio monobásico 36 mM, fosfato de sódio dibásico 14 mM pH 7.2, beta-mercaptoetanol 2 mM, NaCl 200 mM,

glicerol 1 %, EDTA 2 mM. F) Comportamento da elipticidade em 220 nm em função da temperatura a pH neutro. A  $T_m$  de LeiFABP nesta condição foi estimada em 50°C.

#### 4.16. Espectro de RMN Unidimensional de LeiFABP

A espectroscopia de Ressonância Magnética Nuclear é uma técnica utilizada para caracterizar a estrutura da matéria (RULE, 2006). Partículas subatômicas possuem uma propriedade quântica denominado *spin*, cuja analogia clássica é entendida como uma grandeza vetorial associada ao momento angular da partícula. Partículas com carga e *spin* diferente de zero possuem momento magnético, o qual pode interagir com campos magnéticos externos. A interação entre o momento magnético nuclear e o campo magnético externo resulta no aparecimento de uma energia denominada energia de acoplamento (Howard, 1998) (RULE, 2006). Devido ao acoplamento entre o momento magnético nuclear e o campo magnético externo os dois estados do spin 1/2, alfa (*spin down*) e beta (*spin up*), adquirem energias diferentes. A variação de energia devido à transição entre os estados de *spin alfa e beta* pode ser calculada como sendo:

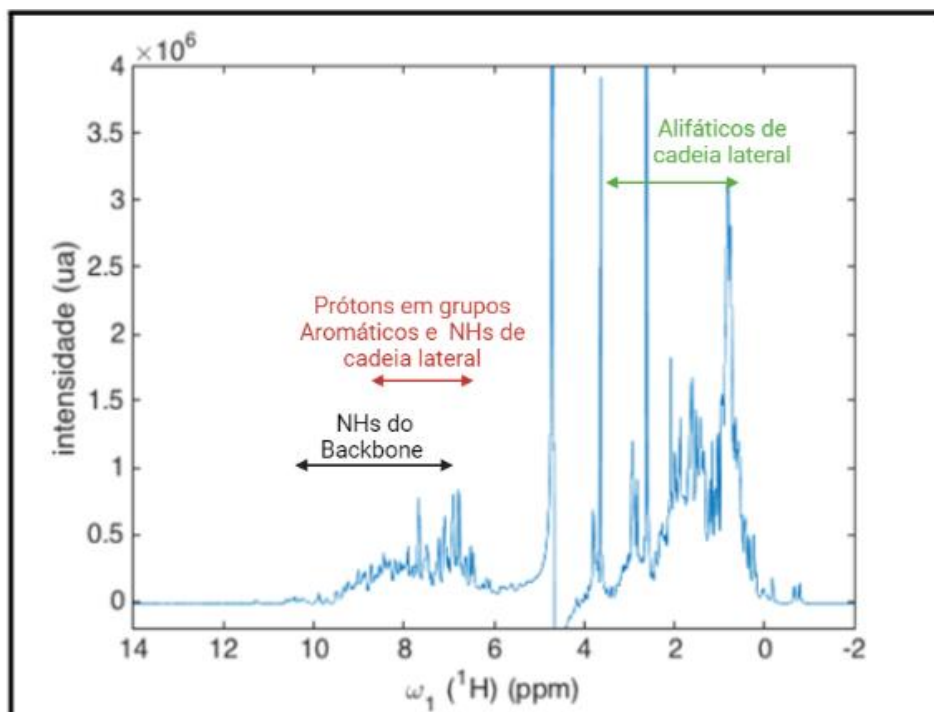
$$\Delta E = h\omega_0 \text{ (Equação 7)}$$

em que a  $h$  é a constante de Planck e  $\omega_0$  a frequência de Larmor. A absorção de um pacote de energia  $\Delta E$  pelo sistema é capaz de promover transições entre os estados de *spin alfa e beta*. A espectroscopia de ressonância magnética interessa-se em obter informações sobre a energia de transição, comumente descrita em termos da frequência de ressonância,  $\omega_0$ , dos spins nucleares presentes em uma molécula de interesse. A frequência de ressonância de um *spin* depende da densidade eletrônica ao redor, e conseqüentemente da estrutura química e tridimensional da molécula (RULE, 2006).

*Spins* apresentam mecanismos de interação por acoplamento escalar, fenômeno que ocorre entre átomos unidos por ligações covalentes, e por efeito dipolar, através de átomos

não ligados, mas próximos no espaço (Efeito dipolar/NOE). Estruturas de proteínas podem ser determinadas por RMN uma vez que acoplamentos escalares entre átomos unidos covalentemente fornecem informações sobre ângulos diedrais (Howard, 1998) (RULE, 2006). Adicionalmente, os efeitos NOE fornecem informações espaciais e estereoquímicas entre átomos próximos, contribuindo para a determinação de estrutura (RULE, 2006)

Medidas de RMN de uma proteína não enriquecida isotopicamente permitem examinar a viabilidade das amostras para estudos estruturais. Proteínas mal enoveladas, com sinais de baixa amplitude e mal definidos na dimensão de prótons poderiam dificultar a determinação da estrutura por RMN. Nossas medidas de RMN de LeiFABP não marcado mostram sinais bem espalhados na região de prótons HN apontando que os hidrogênios da cadeia principal possuem diferentes frequências de ressonância, evidenciadas pelo espalhamento do sinal entre 10 e 6 ppm (Figura 47) (RULE, 2006). Como os prótons HN apresentam frequências de ressonância distintas, infere-se que esses hidrogênios estejam expostos a diferentes ambientes químicos (RULE, 2006). Os diferentes ambientes químicos experimentados pelos prótons HN sugerem que estes núcleos estejam em contato com vizinhanças distintas, provavelmente devido à aquisição de estrutura nativa. Prótons HN em cadeias polipeptídicas não enoveladas apresentariam frequências de ressonância similares, com um único sinal na região de 8 ppm, uma vez que estariam expostos ao mesmo ambiente químico, neste caso, apenas interagiriam com o solvente. Portanto, as diferenças nas frequências de ressonância evidenciadas pelo espalhamento de sinais no espectro (Figura 47), permitem inferir que LeiFABP se encontra enovelado. Além disso, a boa resolução do espectro unidimensional de  $^1\text{H}$  mostra que o domínio LeiFABP recombinante possui tamanho adequado para análises de RMN. Dessa forma, conclui-se que a construção proposta é adequada para os estudos estruturais do domínio LeiFABP.



**Figura 47:** Espectro de Ressonância Magnética Nuclear 1D ( $^1\text{H}$ ) a 800 MHz com pré-saturação da água. A amostra de LeiFABP apresentava volume de 500  $\mu\text{L}$  e concentração de 202  $\mu\text{M}$ . A região dos sinais de NH's da cadeia principal e aromáticos de cadeia lateral, como os sinais de alifáticos na região de 4 a 0 ppm estão indicados. Tampão: fosfato de sódio monobásico 36 mM, fosfato de sódio dibásico 14 mM mM pH 7,2, 400 mM de NaCl, 10 mM de beta-mercaptoetanol, 5 mM de EDTA.



---

## **5. Conclusões e Perspectivas**

---

## 5.1. CONCLUSÕES

Neste projeto clonamos e expressamos quatro construções contendo diferentes segmentos de LmxM.34.3070 na tentativa de analisar o seu domínio C-terminal FABP-like. Observamos que a expressão dos fragmentos recombinantes em fusão com uma *tag* de histidinas na região N-terminal em pET28a levou à melhor taxa de expressão. Além disso, a escolha do início do domínio (resíduos 638, 652 ou 669) influenciou significativamente a sua solubilidade. O segmento 638 - 822 de LmxM.34.3070, clonado no vetor pET28a para expressão em fusão com uma cauda de histidinas N-terminal, apresentou a melhor solubilidade e, portanto, foi utilizado para a caracterização de LeiFABP por SEC-MALS, CD e RMN. O rendimento de uma preparação desta construção purificada a partir de meio litro de LB foi de 4 mg conforme a medida de absorvância a 280 nm. Aparentemente as hélices previstas para a região N-terminal desse fragmento são importantes para a estabilidade apesar de não fazerem parte da topologia clássica de uma FABP. De fato, predições de desordem sugerem que essa região é rígida, e a predição de estrutura sugere que essas hélices formam contatos com o restante do domínio.

Dados de SEC-MALS e de espectroscopia de dicroísmo circular mostraram que **LeiFABP apresenta estruturas secundárias consistente com os modelos obtidos computacionalmente,  $T_m \sim 50$  °C em tampão PBS e acetato, e  $T_m \sim 68$  °C em Tris pH 8, e, além disso, se encontra monomérica nas condições testadas.** O espectro de RMN unidimensional obtido a 800 MHz apresentou boa resolução com sinais bem definidos na região dos prótons HN. A boa resolução do espectro unidimensional de  $^1\text{H}$  permite inferir que o domínio LeiFABP recombinante encontra-se bem enovelado e possui tamanho adequado para análises da estrutura por RMN.

## 5.2. PERSPECTIVAS

FABPs nunca foram anteriormente descritas em *Leishmania spp.* Nossos resultados demonstraram a expressão de um gene codificador de uma proteína contendo um domínio FABP-like, com conteúdo de estrutura secundária consistente com os modelos preditos. Após as otimizações desempenhadas, nossas amostras demonstram-se elegíveis para ensaios de biologia estrutural.

As técnicas de determinação de estrutura são indispensáveis quando se trata do planejamento racional de fármacos. No caso das leishmanioses, um dos grandes interesses é o desenvolvimento de quimioterapias menos nocivas ao paciente e o desenho de moléculas inibidoras de alvos com toxicidade seletiva para o patógeno. Para isso, necessita-se de um conhecimento acurado das características específicas desses alvos, que pode ser obtido através da aquisição de modelos tridimensionais de alta resolução. Portanto, nossos próximos passos envolvem a marcação isotópica da LeiFABP para resolução da estrutura por RMN ou cristalografia.

Adicionalmente, pretendemos obter linhagens celulares de *L. amazonensis* nocautes deste gene para análise de fenótipo dos mutantes e obtenção de informações sobre sua função. Linhagens nocautes também permitirão concluir se *LmxM.34.3070* é um gene essencial.

Considerando-se que o metabolismo de ácidos graxos e lipídios é imprescindível para a sobrevivência do parasito em ambiente hospedeiro esperamos, com este estudo, ampliar o entendimento sobre o metabolismo de ácidos graxos e lipídeos em *Leishmania*.

---

## **6. Referências**

---

- Abdul-Gader, A., Miles, A. J., & Wallace, B. A. (2011). A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics*, 27(12), 1630–1636. <https://doi.org/10.1093/bioinformatics/btr234>
- Alcolea, P. J., Alonso, A., Molina, R., Jiménez, M., Myler, P. J., & Larraga, V. (2019). Functional genomics in sand fly–derived *Leishmania* promastigotes. In *PLoS Neglected Tropical Diseases* (Vol. 13, Issue 5). Public Library of Science. <https://doi.org/10.1371/journal.pntd.0007288>
- Altschup, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic Local Alignment Search Tool. In *J. Mol. Biol* (Vol. 215).
- Aslett, M., Aurrecochea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., Gardner, M. J., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Hertz-Fowler, C., Houston, R., Innamorato, F., Iodice, J., ... Wang, H. (2009). TriTrypDB: A functional genomic resource for the Trypanosomatidae. *Nucleic Acids Research*, 38(SUPPL.1). <https://doi.org/10.1093/nar/gkp851>
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B., Böhme, U., Hannick, L., Aslett, M. A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U. C. M., Arrowsmith, C., ... El-Sayed, N. M. (2005). The Genome of the African Trypanosome *Trypanosoma brucei*. *Science*, 309(5733), 416–422. <https://doi.org/10.1126/science.1112642>
- Birnboim, H. C., & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research*, 7(6), 1513–1523. <https://doi.org/10.1093/nar/7.6.1513>
- Block, H., Maertens, B., Priestersbach, A., Brinker, N., Kubicek, J., Fabis, R., Labahn, J., & Schäfer, F. (2009). Chapter 27 Immobilized-Metal Affinity Chromatography (IMAC). A Review. In *Methods in Enzymology* (Vol. 463, Issue C, pp. 439–473). Academic Press Inc. [https://doi.org/10.1016/S0076-6879\(09\)63027-5](https://doi.org/10.1016/S0076-6879(09)63027-5)
- Booth, W. T., Schlachter, C. R., Pote, S., Ussin, N., Mank, N. J., Klapper, V., Offermann, L. R., Tang, C., Hurlburt, B. K., & Chruszcz, M. (2018). Impact of an N-terminal polyhistidine tag on protein thermal stability. *ACS Omega*, 3(1), 760–768. <https://doi.org/10.1021/acsomega.7b01598>
- Burton Bogitsh, C. C. T. O. (2005). *Human Parasitology* (3rd ed.). Academic Press.
- Calderone, V., Folli, C., Marchesani, A., Berni, R., & Zanotti, G. (2002). Identification and structural analysis of a zebrafish apo and holo cellular retinol-binding protein. *Journal of Molecular Biology*, 321(3), 527–535. [https://doi.org/10.1016/S0022-2836\(02\)00628-9](https://doi.org/10.1016/S0022-2836(02)00628-9)
- Casali, N. (n.d.). *Escherichia coli* Host Strains. In *E. coli Plasmid Vectors* (Vol. 235, pp. 27–48). Humana Press. <https://doi.org/10.1385/1-59259-409-3:27>
- Chiu, H. J., Bakolitsa, C., Skerra, A., Lomize, A., Carlton, D., Miller, M. D., Krishna, S. S., Abdubek, P., Astakhova, T., Axelrod, H. L., Clayton, T., Deller, M. C., Duan, L., Feuerhelm, J., Grant, J. C., Grzechnik, S. K., Han, G. W., Jaroszewski, L., Jin, K. K., ... Wilson, I. A. (2010). Structure of the first representative of Pfam family PF09410 (DUF2006) reveals a structural signature of the calycin superfamily that suggests a role in lipid metabolism. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications*, 66(10), 1153–1159. <https://doi.org/10.1107/S1744309109037749>
- Chmurzyńska, A. (2006). The multigene family of fatty acid-binding proteins (FABPs): Function, structure and polymorphism. *Journal of Applied Genetics*, 47(1), 39–48. <https://doi.org/10.1007/BF03194597>
- Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research*, 14(6), 1188–1190. <https://doi.org/10.1101/gr.849004>
- Cunningham, A. C. (2002). Parasitic adaptive mechanisms in infection by *Leishmania*. *Experimental and Molecular Pathology*, 72(2), 132–141. <https://doi.org/10.1006/exmp.2002.2418>

- de Aquino, G. P., Gomes, M. A. M., Salinas, R. K., & Laranjeira-Silva, M. F. (2021). Lipid and fatty acid metabolism in trypanosomatids. In *Microbial Cell* (Vol. 8, Issue 11, pp. 262–275). Shared Science Publishers OG. <https://doi.org/10.15698/mic2021.11.764>
- Desjeux, P. (1996). Leishmaniasis. *Clinics in Dermatology*, 14(5), 417–423. [https://doi.org/10.1016/0738-081X\(96\)00057-0](https://doi.org/10.1016/0738-081X(96)00057-0)
- Flower, D. R. (1996). The lipocalin protein family : structure and function. In *Biochem. J* (Vol. 318).
- Flower, D. R., North, A. C. T., & Sansom, C. E. (2000). The lipocalin protein family: structural and sequence overview. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1482(1–2), 9–24. [https://doi.org/10.1016/S0167-4838\(00\)00148-5](https://doi.org/10.1016/S0167-4838(00)00148-5)
- Francis, D. M., & Page, R. (2010). Strategies to optimize protein expression in E. coli. In *Current Protocols in Protein Science* (Issue SUPPL. 61). <https://doi.org/10.1002/0471140864.ps0524s61>
- Furuhashi, M., & Hotamisligil, G. S. (2008). Fatty acid-binding proteins: Role in metabolic diseases and potential as drug targets. In *Nature Reviews Drug Discovery* (Vol. 7, Issue 6, pp. 489–503). <https://doi.org/10.1038/nrd2589>
- Gessmann, D., Mager, F., Naveed, H., Arnold, T., Weirich, S., Linke, D., Liang, J., & Nussberger, S. (2011). Improving the resistance of a eukaryotic  $\beta$ -barrel protein to thermal and chemical perturbations. *Journal of Molecular Biology*, 413(1), 150–161. <https://doi.org/10.1016/j.jmb.2011.07.054>
- Ghorbani, M., & Farhoudi, R. (2018). Leishmaniasis in humans: Drug or vaccine therapy? In *Drug Design, Development and Therapy* (Vol. 12, pp. 25–40). Dove Medical Press Ltd. <https://doi.org/10.2147/DDDT.S146521>
- Gibson, D. G. (2011). Enzymatic assembly of overlapping DNA fragments. In *Methods in Enzymology* (Vol. 498, pp. 349–361). Academic Press Inc. <https://doi.org/10.1016/B978-0-12-385120-8.00015-2>
- Gibson, E. G., Oviatt, A. A., & Osheroff, N. (2020). Two-Dimensional Gel Electrophoresis to Resolve DNA Topoisomers. In *Methods in Molecular Biology* (Vol. 2119, pp. 15–24). Humana Press Inc. [https://doi.org/10.1007/978-1-0716-0323-9\\_2](https://doi.org/10.1007/978-1-0716-0323-9_2)
- Glatz, J. F. C., Luiken, J. J. F. P., & Bonen, A. (2010). *Membrane Fatty Acid Transporters as Regulators of Lipid Metabolism: Implications for Metabolic Disease*. <https://doi.org/10.1152/physrev.00003.2009.-Long-chain>
- Gomes, L., Monteiro, G., & Mergulhão, F. (2020). The impact of IPTG induction on plasmid stability and heterologous protein expression by escherichia coli biofilms. *International Journal of Molecular Sciences*, 21(2). <https://doi.org/10.3390/ijms21020576>
- Green, M. R., & Sambrook, J. (2021). Cloning polymerase chain reaction (PCR) products: TOPO TA cloning. *Cold Spring Harbor Protocols*, 2021(9), 352–357. <https://doi.org/10.1101/pdb.prot101311>
- Greenfield, N. J. (2007). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, 1(6), 2876–2890. <https://doi.org/10.1038/nprot.2006.202>
- Hagel, L. (2011). Gel Filtration: Size Exclusion Chromatography. In *Protein Purification: Principles, High Resolution Methods, and Applications* (3rd ed., pp. 51–91). <https://doi.org/10.1002/9780470939932.ch3>
- Handman, E. (1999). *Cell Biology of Leishmania* (pp. 1–39). [https://doi.org/10.1016/S0065-308X\(08\)60229-8](https://doi.org/10.1016/S0065-308X(08)60229-8)
- Henriques, C., Atella, G. C., Bonilha, V. L., & De Souza, W. (2003). Biochemical analysis of proteins and lipids found in parasitophorous vacuoles containing *Leishmania amazonensis*. *Parasitology Research*, 89(2), 123–133. <https://doi.org/10.1007/s00436-002-0728-y>
- Holm, L. (2020). DALI and the persistence of protein shape. *Protein Science*, 29(1), 128–140. <https://doi.org/10.1002/pro.3749>
- Hong, A., Zampieri, R. A., Shaw, J. J., Floeter-Winter, L. M., & Laranjeira-Silva, M. F. (2020). One health approach to leishmaniasis: Understanding the disease dynamics through diagnostic tools. In *Pathogens* (Vol. 9, Issue 10, pp. 1–24). MDPI AG. <https://doi.org/10.3390/pathogens9100809>

- Howard, M. J. (1998). Protein NMR spectroscopy. *Current Biology*, 8(10), R331–R333. [https://doi.org/10.1016/S0960-9822\(98\)70214-3](https://doi.org/10.1016/S0960-9822(98)70214-3)
- Hu, G., Katuwawala, A., Wang, K., Wu, Z., Ghadermarzi, S., Gao, J., & Kurgan, L. (2021). fIDPnn: Accurate intrinsic disorder prediction with putative propensities of disorder functions. *Nature Communications*, 12(1). <https://doi.org/10.1038/s41467-021-24773-7>
- Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices 1 Edited by G. Von Heijne. *Journal of Molecular Biology*, 292(2), 195–202. <https://doi.org/10.1006/jmbi.1999.3091>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Kaye, P., & Scott, P. (2011). Leishmaniasis: Complexity at the host-pathogen interface. In *Nature Reviews Microbiology* (Vol. 9, Issue 8, pp. 604–615). <https://doi.org/10.1038/nrmicro2608>
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols*, 10(6), 845–858. <https://doi.org/10.1038/nprot.2015.053>
- Kelly, S. M., Jess, T. J., & Price, N. C. (2005). How to study proteins by circular dichroism. In *Biochimica et Biophysica Acta - Proteins and Proteomics* (Vol. 1751, Issue 2, pp. 119–139). <https://doi.org/10.1016/j.bbapap.2005.06.005>
- LAEMMLI, U. K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature*, 227(5259), 680–685. <https://doi.org/10.1038/227680a0>
- Lee, S. H., Stephens, J. L., Paul, K. S., & Englund, P. T. (2006). Fatty Acid Synthesis by Elongases in Trypanosomes. *Cell*, 126(4), 691–699. <https://doi.org/10.1016/j.cell.2006.06.045>
- Lemke, A., Kiderlen, A. F., & Kayser, O. (2005). Amphotericin B. In *Applied Microbiology and Biotechnology* (Vol. 68, Issue 2, pp. 151–162). <https://doi.org/10.1007/s00253-005-1955-9>
- Livore, V. I., Tripodi, K. E. J., & Uttaro, A. D. (2007). Elongation of polyunsaturated fatty acids in trypanosomatids. *FEBS Journal*, 274(1), 264–274. <https://doi.org/10.1111/j.1742-4658.2006.05581.x>
- Longmire, J., Albright, K., Lewis, A., Meincke, L., Hildebrand, C., Leadon, S., Cerutti, P., Dykes, D. D., Fondell, J., Watkins, P., & Polesky, H. F. (1987). A simple salting out procedure for extracting DNA from human nucleated cells. In *Nucleic Acids Research* (Vol. 15, Issue 2). C I R L Press Limited.
- Machado, G. U., Prates, F. V., & Machado, P. R. L. (2019). Disseminated leishmaniasis: Clinical, pathogenic, and therapeutic aspects. *Anais Brasileiros de Dermatologia*, 94(1), 9–16. <https://doi.org/10.1590/abd1806-4841.20198775>
- Madden, T. L., Busby, B., & Ye, J. (2019). Reply to the paper: Misunderstood parameters of NCBI BLAST impacts the correctness of bioinformatics workflows. In *Bioinformatics* (Vol. 35, Issue 15, pp. 2699–2700). Oxford University Press. <https://doi.org/10.1093/bioinformatics/bty1026>
- Miles, A. J., Ramalli, S. G., & Wallace, B. A. (2022). DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. *Protein Science*, 31(1), 37–46. <https://doi.org/10.1002/pro.4153>
- Murray, H. W., Berman, J. D., Davies, C. R., & Saravia, N. G. (2005). Advances in leishmaniasis. *The Lancet*, 366(9496), 1561–1577. [https://doi.org/10.1016/S0140-6736\(05\)67629-5](https://doi.org/10.1016/S0140-6736(05)67629-5)
- Naderer, T., Ellis, M. A., Sernee, M. F., De Souza, D. P., Curtis, J., Handman, E., & Mcconville, M. J. (2006). *Virulence of Leishmania major in macrophages and mice requires the gluconeogenic enzyme fructose-1,6-bisphosphatase*. [www.genedb.org](http://www.genedb.org)
- Ono, T. (2005). Studies of the FABP family: A retrospective. *Molecular and Cellular Biochemistry*, 277(1–2), 1–6. <https://doi.org/10.1007/s11010-005-4816-z>

- Pacheco, B., Crombet, L., Loppnau, P., & Cossar, D. (2012). A screening strategy for heterologous protein expression in *Escherichia coli* with the highest return of investment. *Protein Expression and Purification*, *81*(1), 33–41. <https://doi.org/10.1016/j.pep.2011.08.030>
- Paloque, L., Perez-Berezo, T., Abot, A., Dalloux-Chioccioli, J., Bourgeade-Delmas, S., Le Faouder, P., Pujo, J., Teste, M.-A., François, J.-M., Schebb, N. H., Mainka, M., Rolland, C., Blanpied, C., Dietrich, G., Bertrand-Michel, J., Deraison, C., Valentin, A., & Cenac, N. (2019). Polyunsaturated fatty acid metabolites: biosynthesis in *Leishmania* and role in parasite/host interaction. *Journal of Lipid Research*, *60*(3), 636–647. <https://doi.org/10.1194/jlr.M091736>
- Pereira, M. G., Nakayasu, E. S., Sant'Anna, C., de Cicco, N. N. T., Atella, G. C., de Souza, W., Almeida, I. C., & Cunha-e-Silva, N. (2011). *Trypanosoma cruzi* epimastigotes are able to store and mobilize high amounts of cholesterol in reservosome lipid inclusions. *PLoS ONE*, *6*(7). <https://doi.org/10.1371/journal.pone.0022359>
- Ponte-Sucre, A., Gamarro, F., Dujardin, J. C., Barrett, M. P., López-Vélez, R., García-Hernández, R., Pountain, A. W., Mwenechanya, R., & Papadopoulou, B. (2017). Drug resistance and treatment failure in leishmaniasis: A 21st century challenge. In *PLoS Neglected Tropical Diseases* (Vol. 11, Issue 12). Public Library of Science. <https://doi.org/10.1371/journal.pntd.0006052>
- Provencher, S. W., & Gloeckner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry*, *20*(1), 33–37. <https://doi.org/10.1021/bi00504a006>
- Rabhi, S., Rabhi, I., Trentin, B., Piquemal, D., Regnault, B., Goyard, S., Lang, T., Descoteaux, A., Enninga, J., & Guizani-Tabbane, L. (2016). Lipid droplet formation, their localization and dynamics during *leishmania major* macrophage infection. *PLoS ONE*, *11*(2). <https://doi.org/10.1371/journal.pone.0148640>
- Real, F., Vidal, R. O., Carazzolle, M. F., Mondego, J. M. C., Costa, G. G. L., Herai, R. H., Würtele, M., De Carvalho, L. M., E Ferreira, R. C., Mortara, R. A., Barbiéri, C. L., Mieczkowski, P., Da Silveira, J. F., Briones, M. R. D. S., Pereira, G. A. G., & Bahia, D. (2013). The genome sequence of *leishmania (Leishmania) amazonensis*: Functional annotation and extended analysis of gene models. *DNA Research*, *20*(6), 567–581. <https://doi.org/10.1093/dnares/dst031>
- Rogers, M. B., Hillely, J. D., Dickens, N. J., Wilkes, J., Bates, P. A., Depledge, D. P., Harris, D., Her, Y., Herzyk, P., Imamura, H., Otto, T. D., Sanders, M., Seeger, K., Dujardin, J. C., Berriman, M., Smith, D. F., Hertz-Fowler, C., & Mottram, J. C. (2011). Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. *Genome Research*, *21*(12), 2129–2142. <https://doi.org/10.1101/gr.122945.111>
- Roux, K. H. (2009). Optimization and troubleshooting in PCR. *Cold Spring Harbor Protocols*, *4*(4). <https://doi.org/10.1101/pdb.ip66>
- RULE, G. S. ; H. T. K. (2006). *Fundamentals of Protein NMR Spectroscopy* (1st ed., Vol. 5). Springer Science & Business Media.
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). *DNA sequencing with chain-terminating inhibitors (DNA polymerase/nucleotide sequences/bacteriophage 4X174)* (Vol. 74, Issue 12).
- Saunders, E. C., Ng, W. W., Kloehn, J., Chambers, J. M., Ng, M., & McConville, M. J. (2014). Induction of a Stringent Metabolic Response in Intracellular Stages of *Leishmania mexicana* Leads to Increased Dependence on Mitochondrial Metabolism. *PLoS Pathogens*, *10*(1). <https://doi.org/10.1371/journal.ppat.1003888>
- Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., & Panda, A. K. (2015). Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. In *Microbial Cell Factories* (Vol. 14, Issue 1). BioMed Central Ltd. <https://doi.org/10.1186/s12934-015-0222-8>
- Singh, S. M., & Panda, A. K. (2005). Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering*, *99*(4), 303–310. <https://doi.org/10.1263/jbb.99.303>



- Smathers, R. L., & Petersen, D. R. (2011). The human fatty acid-binding protein family: Evolutionary divergences and functions. *Human Genomics*, *5*(3), 170. <https://doi.org/10.1186/1479-7364-5-3-170>
- Soares-Bezerra, R. J., Leon, L., & Genestra, M. (2004). Recentes avanços da quimioterapia das leishmanioses: moléculas intracelulares como alvo de fármacos. In *Revista Brasileira de Ciências Farmacêuticas Brazilian Journal of Pharmaceutical Sciences* (Vol. 40, Issue 2).
- Some, D., Amartely, H., Tsadok, A., & Lebendiker, M. (2019). Characterization of proteins by size-exclusion chromatography coupled to multi-angle light scattering (Sec-mals). *Journal of Visualized Experiments*, *2019*(148). <https://doi.org/10.3791/59615>
- Stathopoulos, P. B., Scholz, G. A., Hwang, Y.-M., Rumfeldt, J. A. O., Lepock, J. R., & Meiering, E. M. (2008). Sonication of proteins causes formation of aggregates that resemble amyloid. *Protein Science*, *13*(11), 3017–3027. <https://doi.org/10.1110/ps.04831804>
- Stryjewska, A., Kiepara, K., Librowski, T., & Lochyński, S. (2013). Biotechnology and genetic engineering in the new drug development. Part I. DNA technology and recombinant proteins. *Pharmacological Reports*, *65*(5), 1075–1085. [https://doi.org/10.1016/S1734-1140\(13\)71466-X](https://doi.org/10.1016/S1734-1140(13)71466-X)
- Tavares, V. D. S., De Castro, M. V., Souza, R. D. S. O., Gonçalves, I. K. A., Lima, J. B., Borges, V. D. M., & Araújo-Santos, T. (2021). Lipid droplets from protozoan parasites: Survival and pathogenicity. *Memorias Do Instituto Oswaldo Cruz*, *116*. <https://doi.org/10.1590/0074-02760210270>
- Tessarollo, N. G., Andrade, J. M., Moreira, D. S., & Murta, S. M. F. (2015). Functional analysis of iron superoxide dismutase-A in wild-type and antimony-resistant *Leishmania braziliensis* and *Leishmania infantum* lines. *Parasitology International*, *64*(2), 125–129. <https://doi.org/10.1016/j.parint.2014.11.001>
- Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S., & Arakawa, T. (2004). Role of arginine in protein refolding, solubilization, and purification. In *Biotechnology Progress* (Vol. 20, Issue 5, pp. 1301–1308). American Chemical Society. <https://doi.org/10.1021/bp0498793>
- Uttaro, A. D. (2014). Acquisition and biosynthesis of saturated and unsaturated fatty acids by trypanosomatids. In *Molecular and Biochemical Parasitology* (Vol. 196, Issue 1, pp. 61–70). Elsevier. <https://doi.org/10.1016/j.molbiopara.2014.04.001>
- Waller, R. F., & McConville, M. J. (2002). Developmental changes in lysosome morphology and function *Leishmania* parasites. *International Journal for Parasitology*, *32*(12), 1435–1445. [https://doi.org/10.1016/S0020-7519\(02\)00140-6](https://doi.org/10.1016/S0020-7519(02)00140-6)
- Wang, S., Sun, S., Li, Z., Zhang, R., & Xu, J. (2017). Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLoS Computational Biology*, *13*(1). <https://doi.org/10.1371/journal.pcbi.1005324>
- Waugh, D. S. (2005). Making the most of affinity tags. In *Trends in Biotechnology* (Vol. 23, Issue 6, pp. 316–320). <https://doi.org/10.1016/j.tibtech.2005.03.012>
- WHO. (2021, April 15). *World Health Organization (WHO). Leishmaniases*. [https://www.who.int/health-topics/Leishmaniasis#tab=tab\\_1](https://www.who.int/health-topics/Leishmaniasis#tab=tab_1).
- Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J.-C., Williams, K. L., Appel, R. D., & Hochstrasser, D. F. (n.d.). Protein Identification and Analysis Tools in the ExPASy Server. In *2-D Proteome Analysis Protocols* (Vol. 112, Issue 2, pp. 531–552). Humana Press. <https://doi.org/10.1385/1-59259-584-7:531>
- Xu, J., & Wang, S. (2019). Analysis of distance-based protein structure prediction by deep learning in CASP13. *Proteins: Structure, Function and Bioinformatics*, *87*(12), 1069–1081. <https://doi.org/10.1002/prot.25810>
- Zhang, K., & Beverley, S. M. (2010). Phospholipid and sphingolipid metabolism in *Leishmania*. In *Molecular and Biochemical Parasitology* (Vol. 170, Issue 2, pp. 55–64). <https://doi.org/10.1016/j.molbiopara.2009.12.004>
- Zhang, Y., Zhang, J., Ren, Y., Lu, R., Yang, L., & Nie, G. (2020). Tracing the evolution of fatty acid-binding proteins (FABPs) in organisms with a heterogeneous fat distribution. *FEBS Open Bio*, *10*(5), 861–872. <https://doi.org/10.1002/2211-5463.12840>