

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA
Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)

FABIANA LAURETTI FERREIRA

Caracterização funcional de proteínas hipotéticas de
***Leptospira interrogans* como adesinas e potenciais**
antígenos vacinais e para diagnóstico

Versão corrigida da Tese conforme Resolução CoPGr 5890
O original se encontra disponível na Secretaria de Pós-Graduação do IQ-USP

São Paulo

Data do Depósito na SPG:

08/11/2021

FABIANA LAURETTI FERREIRA

**Caracterização funcional de proteínas hipotéticas de
Leptospira interrogans como adesinas e potenciais
antígenos vacinais e para diagnóstico**

*Tese apresentada ao Instituto de Química da
Universidade de São Paulo para obtenção do
Título de Doutora em Ciências (Bioquímica)*

Orientador: Dr. Paulo Lee Ho

São Paulo

2021

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

F383c Ferreira, Fabiana Lauretti
 Caracterização funcional de proteínas hipotéticas de *Leptospira interrogans* como adesinas e potenciais antígenos vacinais e para diagnóstico / Fabiana Lauretti Ferreira. - São Paulo, 2021.
 139 p.

Tese (doutorado) - Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica.

Orientador: Ho, Paulo Lee

1. Proteínas. 2. *Leptospira*. 3. Adesinas. 4. Virulência. I. T. II. Ho, Paulo Lee, orientador.

À minha amada família: Reinaldo, Walquiria, Julianne, Irene e Benício Augusto.

Obrigada por estarem sempre presentes e cuidarem tão bem de mim

Com todo o meu amor.

AGRADECIMENTOS

Agradeço primeiramente pela minha saúde e de todos os meus familiares, amigos e colegas nestes tempos tão difíceis.

Agradeço ao meu orientador Dr. Paulo Lee Ho por todos esses oito anos de muito aprendizado, que ajudaram a moldar meu desenvolvimento científico e profissional. Muito obrigada por compartilhar seu conhecimento, pelo apoio e motivação fundamentais para a realização do meu trabalho, pela paciência e compreensão, além da confiança e reconhecimento que me ajudaram a chegar até aqui. Seguimos trabalhando juntos!

Ao apoio de toda a equipe do laboratório de Bacteriologia II: Léo, Jô, Enéas, Eliane, Sandra, Vera, Aline e Ricardo, pelo auxílio nos experimentos e pelas discussões sempre enriquecedoras. Agradeço também às doutoras Angela Barbosa e Patrícia Abreu pelo especial suporte experimental e científico em momentos cruciais do Doutorado. Aos membros dos laboratórios de Bacteriologia e Biotecnologia II, III e IV por conceder o uso de equipamentos complementares.

Aos meus amigos e colegas Marcela, Tasson, Paloma, Juliana, Adriano e Giovanna por tornar os dias de trabalho mais leves, tanto com conversas científicas quanto com momentos de muita descontração. Àquelas amigadas que continuaram fora do laboratório: Júlia e Stefanni, muito obrigada por todo o carinho, apoio e todos os momentos divertidos compartilhados. Agradeço a todos pela troca de conhecimentos e experiências que com certeza levarei comigo.

À toda a equipe VAE por me acolherem e fazerem me sentir tão bem e amparada nesses últimos anos. Agradeço à Milena pela grande oportunidade de contribuir com meu trabalho e aprender cada vez mais nesse incrível mundo de produção de vacinas e soros, além da confiança, compreensão e todo apoio profissional e científico. À Bianca por sempre poder contar com o seu suporte a qualquer momento. Agradeço aos colegas e às amigadas que têm

me motivado nestes últimos momentos antes da tese, além de serem apoio aos desabafos e angústias: Dani, Jú (olha ela aqui de novo), Aline, Thiago, Soraia, Gustavo, Vitor e Tércio. Todos vocês me inspiram a ser cada vez melhor. Muito obrigada por tudo, pessoal!

Aos professores e funcionários do Departamento de Bioquímica no Instituto de Química da USP pela importante contribuição na minha formação, especialmente aos docentes de Bioquímica Avançada e Biologia Estrutural, os quais foram as principais fontes de conhecimento teórico e que fizeram me apaixonar pela bioquímica de proteínas.

À comunidade da *Gordon Research Conference on Biology of Spirochetes* de 2018 pela incrível experiência científica em vários aspectos e pela oportunidade única de poder organizar parte do próximo evento.

Agradeço às instituições de fomento à pesquisa fundamentais para a realização deste projeto: ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pela concessão da bolsa de Doutorado, à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) por todo o suporte adicional e ao longo da minha carreira científica, e ao apoio da Fundação Butantan.

Finalmente, a conclusão deste trabalho não seria possível sem aqueles que são fonte de toda a minha força nos bons e maus momentos. Agradeço pela minha querida e tão amada família por todo amor e carinho incondicionais. Aos meus pais Walquiria e Reinaldo, por estarem presentes em todos os momentos e possibilitarem a minha total dedicação à tese nestes últimos meses, sempre com muito carinho, compreensão, paciência, apoio e incentivo. Essa conquista também é de vocês! À minha querida irmã Julianne que sempre me inspira e seu pedacinho de amor mais lindo, Benício Augusto. À minha vó Irene, obrigada por me amar tanto. À minha vó Amélia, que se foi durante essa caminhada. Amo vocês de todo o meu coração.

*“Nada na vida deve ser temido, somente compreendido.
Agora é hora de compreender mais e temer menos.”*

Marie Curie

RESUMO

Ferreira, F.L. **Caracterização funcional de proteínas hipotéticas de *Leptospira interrogans* como adesinas e potenciais antígenos vacinais e para diagnóstico**. 2021. 139p. Tese de Doutorado - Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

Leptospira spp. constitui um grupo de bactérias espiroquetas gram-negativas englobando espécies saprofíticas, intermediárias e patogênicas, sendo as últimas agentes causadores da leptospirose, doença zoonótica de alcance mundial e endêmica em regiões tropicais em desenvolvimento. O crescente número de espécies identificadas de leptospiros destaca ainda mais sua diversidade genética e mecanismos de virulência únicos, muitos deles com função ainda desconhecida. Esforços para o desenvolvimento de novas vacinas com proteção cruzada e efeito duradouro revelaram possíveis candidatos vacinais que necessitam ser adequadamente validados, sendo assim, há ainda uma urgente necessidade de uma vacina universal contra a leptospirose capaz de controlar e reduzir os surtos cada vez mais frequentes da doença. Adesinas são importantes fatores de virulência em diversos patógenos, constituindo antígenos promissores para o desenvolvimento de vacinas contra a leptospirose, assim como para o desenvolvimento de métodos diagnósticos mais rápidos e precisos. Previamente, foram identificadas três proteínas hipotéticas conservadas em *L. interrogans* pela técnica de *phage display*, denominadas arbitrariamente como LepA069, LepA962 e LepA388. A expressão do gene codificador da proteína LepA069 apresentou aumento de aproximadamente 70 % em animais infectados por leptospiros virulentas, representando a primeira evidência funcional desta proteína ainda desconhecida. Porções recombinantes da lipoproteína hipotética LepA962 (LepA962_Nt e LepA962_Phg) foram obtidos, sendo demonstrada a forte interação da proteína

LepA962_Phg, contendo a sequência identificada por *phage display*, com laminina, fibronectina plasmática, colágeno I e fibrinogênio de maneira dose-dependente. Adicionalmente, LepA962_Phg apresentou ligação às células VERO e à sua matriz extracelular secretada, e o soro obtido a partir desta proteína recombinante foi capaz de se ligar à superfície de leptospiros virulentas, indicando que LepA962_Phg pode representar um importante domínio de interação entre as leptospiros e seu hospedeiro. Finalmente, a proteína LepA388 pertencente a uma extensa família de proteínas modificadoras de virulência com função desconhecida (DUF_61), presente apenas nas leptospiros patogênicas mais virulentas, apresentou aumento na expressão de seu gene codificador em animais infectados por leptospiros virulentas de acordo com dados na literatura. Além disso, porções recombinantes da região N-terminal desta proteína apresentaram ligação a laminina, colágenos I e IV, vitronectina e fibronectinas plasmática e celular, principalmente considerando a sequência identificada por *phage display*. Estes dados reforçam as previsões de modelos tridimensionais da proteína LepA388 e de outros membros da família DUF_61, as quais identificam domínios semelhantes a toxinas (como abrina e CARDS) responsáveis pela ligação e internalização celulares nos hospedeiros. Dados recentes sugerem um possível papel citotóxico desempenhado pelas proteínas desta família em leptospiros, as quais podem também ser consideradas potenciais candidatas vacinais e para diagnóstico da leptospirose, devido à sua distribuição restrita em espécies e cepas patogênicas de importância para saúde humana.

Palavras-chave: *Leptospira*, virulência, adesinas, *phage display*, domínios funcionais

ABSTRACT

Ferreira, F.L. **Functional characterization of hypothetical proteins in *Leptospira interrogans* as adhesins and potential vaccine and diagnostic antigens.** 2021. 139p. PhD Thesis - Graduate Program in Biological Sciences (Biochemistry). Instituto de Química, Universidade de São Paulo, São Paulo.

Leptospira spp. constitutes a group of gram-negative spirochete bacteria comprising saprophytic, intermediate and pathogenic species, the last being causative agents of leptospirosis, a zoonotic disease of worldwide extent and endemic in developing tropical regions. The growing number of identified leptospiral species further highlights their genetic diversity and unique virulence mechanisms, many of them with unknown function. Efforts to develop new vaccines with cross-protection and long-lasting effect have revealed possible vaccine candidates that need to be properly validated. Therefore, there is still an urgent need for a universal vaccine against leptospirosis capable of controlling and reducing the increasing outbreaks of the disease. Adhesins are important virulence factors in several pathogens, constituting promising antigens for the development of vaccines against leptospirosis, as well as for the development of faster and more accurate diagnostic methods. Previously, three conserved hypothetical proteins in *L. interrogans* were identified by phage display technique, arbitrarily named as LepA069, LepA962 and LepA388. Expression of the LepA069 encoding gene showed an increase of approximately 70 % in animals infected by virulent leptospires, representing the first functional evidence of this still unknown protein. Recombinant portions of the hypothetical lipoprotein LepA962 (LepA962_Nt and LepA962_Phg) were obtained, demonstrating the strong interaction of the LepA962_Phg protein, containing the sequence identified by phage display, with laminin, plasma fibronectin, collagen I and fibrinogen in a

dose-dependent manner. Furthermore, LepA962_Phg showed binding to VERO cells and its secreted extracellular matrix, and the serum obtained from this recombinant protein was able to bind to the surface of virulent leptospires, indicating that LepA962_Phg may represent an important domain of interaction between leptospires and its host. Finally, LepA388 protein belonging to an extensive family of virulence modifying proteins with unknown function (DUF_61), present only in the most virulent pathogenic leptospires, showed an increase in the expression of its encoding gene in animals infected by virulent leptospires according to data in literature. Moreover, recombinant portions of the N-terminal region of this protein showed binding to laminin, collagens I and IV, vitronectin and plasma and cell fibronectins, especially considering the sequence identified by phage display. These data support the predictions of three-dimensional models of the LepA388 protein and other members of the DUF_61 family, which identify toxin-like domains (such as abrin and CARDS) responsible for cellular binding and internalization in hosts. Recent data suggest a possible cytotoxic role played by proteins of this family in leptospires, which can also be considered potential vaccine candidates and antigens for diagnosis, due to their restricted distribution in pathogenic species and strains of importance to human health.

Keywords: *Leptospira*, virulence, adhesins, *phage display*, functional domains

LISTA DE ABREVIATURAS E SIGLAS

ADP: adenosina difosfato

ATP: adenosina trifosfato

BRASIL: *Biopanning and Rapid Analysis of Selective Interactive Ligands*

BSA: albumina sérica bovina

CARDS: *Community-acquired respiratory distress syndrome*

cDNA: DNA complementar

ColI: colágeno I

ColIV: colágeno IV

Da: Daltons

dATP: desoxiadenosina trifosfato

DMEM: meio Eagle modificado por Dulbecco

DNA: ácido desoxirribonucleico

DNase: desoxirribonuclease

DO_{600nm}: densidade óptica a 600 nm

DUF: domínio de função desconhecida

EGFP: proteína verde fluorescente melhorada

ELISA: ensaio de imunoabsorção enzimática

EMJH: meio de cultura Ellinghausen-McCullough-Johnson-Harris

FN: fibronectina

IMAC: cromatografia de afinidade por metais imobilizados

IPTG: isopropil- β -D-tiogalactopiranosídeo

kb: quilo base

K_D: constante de dissociação

kDa: quilo Dalton

Lam: laminina

LB: meio de cultura Luria-Bertani

LPS: lipopolissacarídeo

mART: mono ADP ribosiltransferase

MAT: teste de aglutinação microscópica

Mb: mega base

MEC: matriz extracelular

mRNA: RNA mensageiro

OMP: proteína de membrana externa

OPD: dicloridrato de o-fenilenodiamina

PBS: solução tamponada de fosfatos

PCR: reação em cadeia da polimerase

PDB: *Protein Data Bank*

pI: ponto isoeletrico

RT-qPCR: PCR quantitativa com transcriptase reversa

SDS-PAGE: eletroforese em gel de poliacrilamida com dodecil sulfato de sódio

SFB: soro fetal bovino

SOC: meio de crescimento super ótimo,
com glicose

TLR: receptor tipo Toll

tRNA: RNA transportador

VTN: vitronectina

X-gal: 5-bromo-4-cloro-3-indolil- β -D-
galactopiranosídeo

SUMÁRIO

1. INTRODUÇÃO	16
1.1. Gênero <i>Leptospira</i>	16
1.1.1. Taxonomia e classificações filogenéticas	16
1.1.2. Principais características estruturais e fisiológicas	19
1.1.3. Características genômicas das diferentes espécies de <i>Leptospira</i>	21
1.1.4. Fatores de virulência e adesinas	23
1.2. Leptospirose	26
1.2.1. Ciclo de transmissão e diversidade de hospedeiros	26
1.2.2. Incidência e epidemiologia	27
1.2.3. Principais manifestações clínicas	29
1.2.4. Diagnóstico	30
1.2.5. Desenvolvimento de vacinas	31
1.3. Resultados prévios – Identificação de novas adesinas por <i>phage display</i>	34
2. OBJETIVOS	36
2.1. Objetivos geral e específicos	36
3. MATERIAL E MÉTODOS	37
3.1. Organismos, condições gerais de cultivo e vetores	37
3.2. Preparação de bactérias <i>E. coli</i> competentes	38
3.2.1. Bactérias <i>E. coli</i> eletrocompetentes	38
3.2.2. Bactérias <i>E. coli</i> quimiocompetentes	39
3.3. Desenho de oligonucleotídeos para clonagens em vetores de expressão e para reações de qPCR	40
3.4. Avaliação da expressão gênica das proteínas hipotéticas por RT-qPCR	41
3.5. Análises <i>in silico</i> das proteínas hipotéticas	42
3.6. Clonagem e subclonagem das sequências selecionadas aos vetores de expressão	43
3.7. Produção das proteínas recombinantes em diferentes cepas de bactérias <i>E. coli</i>	45
3.7.1. Produção de proteínas recombinantes em maior volume	48
3.8. Isolamento das proteínas recombinantes produzidas em bactérias	49
3.8.1. Extração das proteínas recombinantes das células bacterianas (lise celular)	50
3.8.2. Purificação das proteínas recombinantes (IMAC e outras técnicas)	52
3.8.3. Diálise das proteínas recombinantes de LepA962 purificadas	53
3.9. Análise de amostras da proteína recombinante LepA962_Nt por espectrometria de massas	55
3.10. Avaliação da estrutura secundária de proteínas recombinantes por difração de raios X	55

3.11. Produção de antissoros em camundongos a partir de proteínas recombinantes	56
3.12. Ensaio imunoenzimático e de ligação <i>in vitro</i> baseados em ELISA	58
3.12.1. Titulação de antissoros	58
3.12.2. Ligação de proteínas recombinantes a componentes de matriz extracelular	59
3.12.3. Ligação de proteínas recombinantes a células VERO e à MEC secretada	61
3.12.4. Ligação dos antissoros das proteínas recombinantes na superfície de leptospiros	62
3.13. Detecção da interação de proteínas com componentes da MEC por <i>Far Western Blotting</i>	63
3.14. Migrações eletroforéticas	65
3.14.1. Eletroforese em gel de agarose	65
3.14.2. Eletroforese em gel de poliacrilamida (SDS-PAGE)	66
4. RESULTADOS	67
4.1. Estratégias exploradas para produção e purificação de proteínas recombinantes	67
4.2. Análise da lipoproteína hipotética LepA962	70
4.2.1. Desenho, produção e purificação das proteínas recombinantes LepA962_Nt e LepA962_Phg	70
4.2.2. Análise da estrutura secundária das proteínas recombinantes LepA962_Nt e LepA962_Phg e obtenção de antissoros	74
4.2.3. Análise da interação entre as proteínas recombinantes (LepA962_Nt e LepA962_Phg) e componentes da MEC	77
4.2.4. Análise da interação entre as proteínas recombinantes (LepA962_Nt e LepA962_Phg) a células VERO e sua MEC secretada	80
4.2.5. Análise da interação dos antissoros das proteínas recombinantes LepA962_Nt e LepA962_Phg à superfície de leptospiros	81
4.3. Análise da lipoproteína hipotética LepA388	83
4.3.1. Expressão gênica em camundongos infectados com <i>L. interrogans</i>	83
4.3.2. Predições da estrutura tridimensional da proteína LepA388	84
4.3.2.1. Modelos a partir do servidor Phyre2	84
4.3.2.2. Modelo a partir dos programas SWISS-Model, HHpred e I-TASSER	90
4.3.2.3. Predição de domínios conservados na estrutura tridimensional de LepA388	96
4.3.3. Desenho, produção e purificação de proteínas recombinantes da proteína LepA388	97
4.3.4. Análise da interação entre as proteínas recombinantes de LepA388 com proteínas de MEC	99
4.3.5. Novas modelagens de LepA388 e de todos os membros da família DUF_61	103
4.3.6. Produção e purificação de proteínas fluorescentes com EGFP	105

4.4. Análise da lipoproteína hipotética LepA069	107
5. DISCUSSÃO	109
6. CONCLUSÕES E PERSPECTIVAS	117
7. REFERÊNCIAS	118
APÊNDICES	133
LISTA DE ANEXOS	139

1. INTRODUÇÃO

1.1. Gênero *Leptospira*

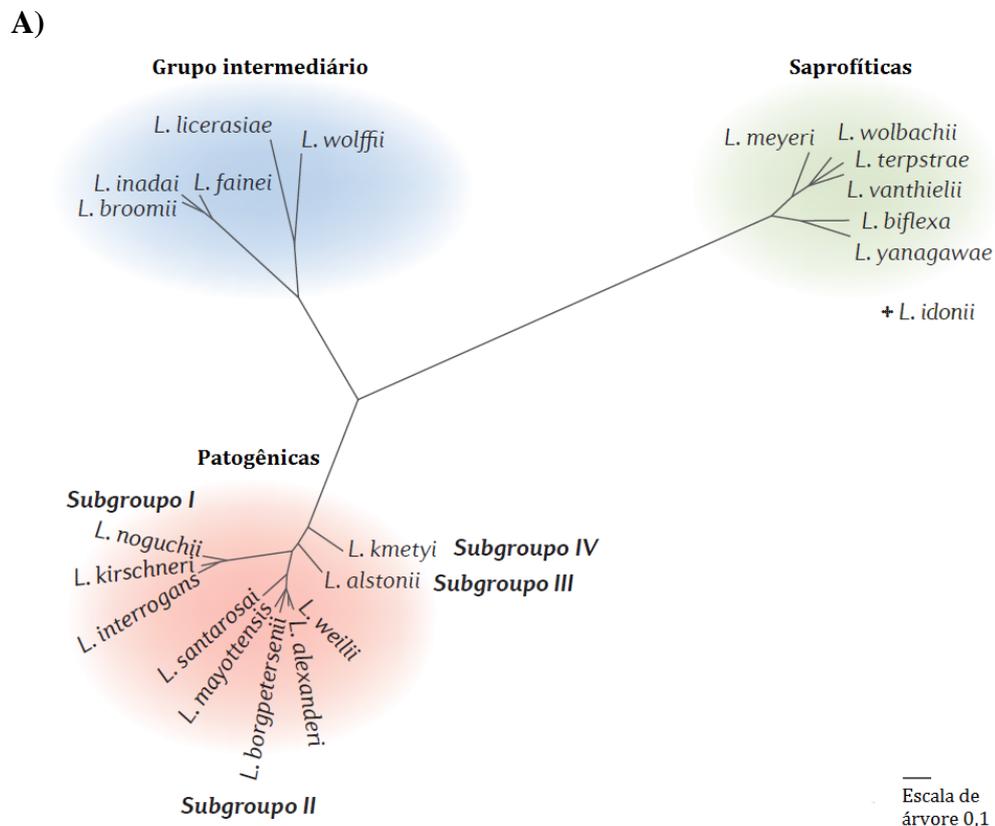
1.1.1. Taxonomia e classificações filogenéticas

As bactérias gram-negativas do gênero *Leptospira* pertencem ao filo Spirochaetes e à classe *Spirochaetia*. Anteriormente, as leptospiras e todas as outras espiroquetas eram classificadas em uma única ordem, *Spirochaetales*, a qual reunia quatro principais famílias: *Spirochaetaceae*, *Brachyspiraceae*, *Leptospiraceae* e *Brevinemataceae* (BUCHANAN, 1917; PASTER, 2010). Entretanto, a recente identificação de marcadores moleculares conservados em cada uma das famílias e novas análises filogenéticas revelaram uma extensa diversidade genética entre as espiroquetas, sendo propostos novos táxons, incluindo ordens e famílias. Desta forma, a família *Leptospiraceae* passou a ser classificada como a única família da nova ordem *Leptospirales*, a qual possui assinaturas conservadas na proteína ribossomal L14 (50S), na proteína ribossomal S2 (30S), na alanina-tRNA sintetase e nas proteínas flagelares FlgG e FlaB (GUPTA et al., 2013a; GUPTA et al., 2013b). A família *Leptospiraceae* engloba o gênero-tipo *Leptospira* e os gêneros *Leptonema* e *Turneriella*.

O gênero *Leptospira* inclui bactérias de vida livre e patogênicas, as últimas sendo agentes etiológicos da leptospirose em animais e humanos. Inicialmente, as espécies patogênicas eram subdivididas em sorovares, baseados na diversidade estrutural do componente carboidrato dos lipopolissacarídeos (LPS) expostos em sua superfície, e definidos pelo teste de aglutinação microscópica (MAT). Atualmente, a classificação sorológica das leptospiras ainda é importante para investigações clínicas e epidemiológicas da leptospirose, sendo identificados mais de 300 sorovares distintos. Entretanto, essa classificação consiste em uma metodologia trabalhosa e que requer laboratórios especializados (PICARDEAU, 2017). Além disso, a divisão das leptospiras em sorovares não possui correlação com análises moleculares das

espécies a partir da crescente disponibilização de genomas completos sequenciados, indicando que os genes envolvidos na diversidade dos sorovares podem ser transmitidos lateralmente entre as diferentes espécies (FAINE et al., 1999; KO et al., 2009; LEVETT, 2015).

Até 2018, o gênero *Leptospira* compreendia 22 espécies identificadas principalmente por análises filogenéticas baseadas em hibridização de DNA e no gene 16S rRNA, as quais eram divididas em três clados: saprofitas (7 espécies), patogênicas (10 espécies) e intermediárias (5 espécies), sendo as últimas causadoras de manifestações leves da leptospirose ou com patogenicidade indefinida (Figura 1A) (KO et al., 2009; PICARDEAU, 2017). As espécies patogênicas eram ainda divididas em subgrupos I-IV (Figura 1A), de acordo com ganhos de genes ocorridos evolutivamente por duplicação gênica ou transferência horizontal (XU et al., 2016).



B)

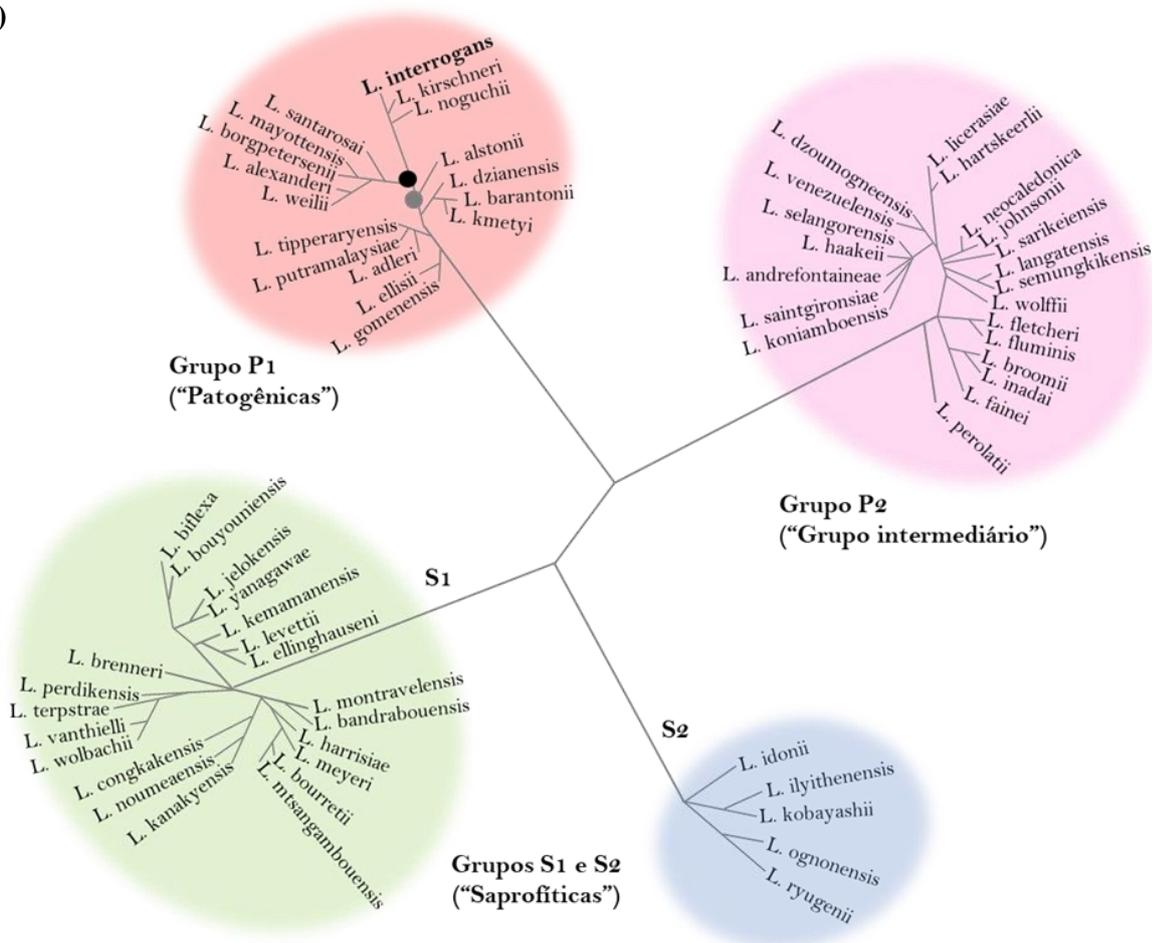


Figura 1. A) Árvore filogenética com a classificação das 22 espécies de *Leptospira* até 2018, divididas em espécies saprofiticas, patogênicas e grupo intermediário. As espécies patogênicas são ainda divididas em subgrupos I-IV. Figura modificada de PICARDEAU, 2017. **B)** Classificação das 64 espécies de *Leptospira* a partir de 2019, com os grupos P1 (patogênicas), P2 (intermediárias), S1 e S2 (saprofiticas), baseados no trabalho de VINCENT et al., 2019. A relação entre as espécies é meramente ilustrativa e não representa as distâncias filogenéticas entre elas. O círculo preto destaca o nodo com as espécies patogênicas mais frequentemente envolvidas em leptospirose humana, enquanto que o círculo cinza destaca o nodo em que as proteínas modificadoras de virulência (VM) estão presentes.

Nos anos seguintes, entretanto, houve um incrível salto no número de espécies devido ao aumento da disponibilidade do sequenciamento de nova geração, permitindo a anotação de genomas completos de maneira mais rápida e relativamente econômica, além do crescente interesse em estudos de metagenômica e de comunidades microbianas presentes em solos (THIBEAUX et al., 2018; VINCENT et al., 2019). Atualmente, estão descritas 64 espécies de *Leptospira*, com a nova proposta de divisão em quatro grupos (Figura 1B): P1 com 17 espécies

patogênicas, P2 com 21 espécies intermediárias cuja patogenicidade ainda é indefinida, e as espécies saprofíticas que foram subdividas nos grupos S1 (21 espécies) e S2 (5 espécies) (PICARDEAU, 2020).

1.1.2. Principais características estruturais e fisiológicas

Leptospira spp. são bactérias espiroquetas, apresentando-se como células espiraladas, flexíveis e com mobilidade em ambientes líquidos e semissólidos. Além do formato em espiral, as leptospiros são morfologicamente finas e alongadas, com diâmetro entre 0,1-0,15 μm e comprimento de 6-20 μm (FAINE et al., 1999).

Seu envelope celular é semelhante a bactérias gram-negativas, apresentando membranas interna e externa envolvendo o espaço periplasmático que contém a parede celular de peptidoglicano, da qual são derivados o formato, força e rigidez das leptospiros. A membrana externa expõe na superfície numerosas proteínas e lipoproteínas, além das moléculas de LPS, as quais são responsáveis pela diversidade antigênica das leptospiros (Figura 2). Entretanto, uma das singularidades que as leptospiros apresentam em sua morfologia é a estreita proximidade da parede de peptidoglicano à membrana citoplasmática (ou interna), resultando numa membrana externa fluida e fracamente conectada ao restante da célula. Devido às características singulares de seu envelope celular, as leptospiros não são coradas por coloração Gram, sendo visualizadas e analisadas sob microscopia de campo escuro (PICARDEAU, 2020). Como a maioria das espiroquetas, as leptospiros possuem dois flagelos periplasmáticos ou endoflagelos, componentes singulares entre as bactérias. Os flagelos estão ancorados próximos às extremidades das células e se estendem em seu comprimento, mas não se sobrepõem, e esta disposição e ancoramento são responsáveis por promover movimentos translacionais, além das rotações intrínsecas dos flagelos determinarem as características extremidades em formato de gancho ou em espiral (Figura 2) (CAMERON, 2015; PICARDEAU, 2017).

Todas as leptospiros crescem em ambientes aeróbicos ou microaerofílicos, e utilizam ácidos graxos de cadeia longa como fonte principal de carbono e energia (PASTER, 2010). As condições ótimas de crescimento *in vitro* tanto para espécies saprofíticas quanto para patogênicas ocorrem em temperatura de 28-30 °C em meio de cultura Ellinghausen-McCullough-Johnson-Harris (EMJH), com pH entre 7,2 e 7,6. Espécies patogênicas também crescem a 37 °C, mas não exibem crescimento em baixas temperaturas (11-13 °C) como as espécies saprofíticas. O tempo de dobramento celular *in vitro* varia entre 6-20 h; o crescimento das espécies patogênicas é mais lento, podendo se estender por dias ou até semanas quando são recentemente isoladas de hospedeiros (CAMERON, 2015; PICARDEAU, 2020).

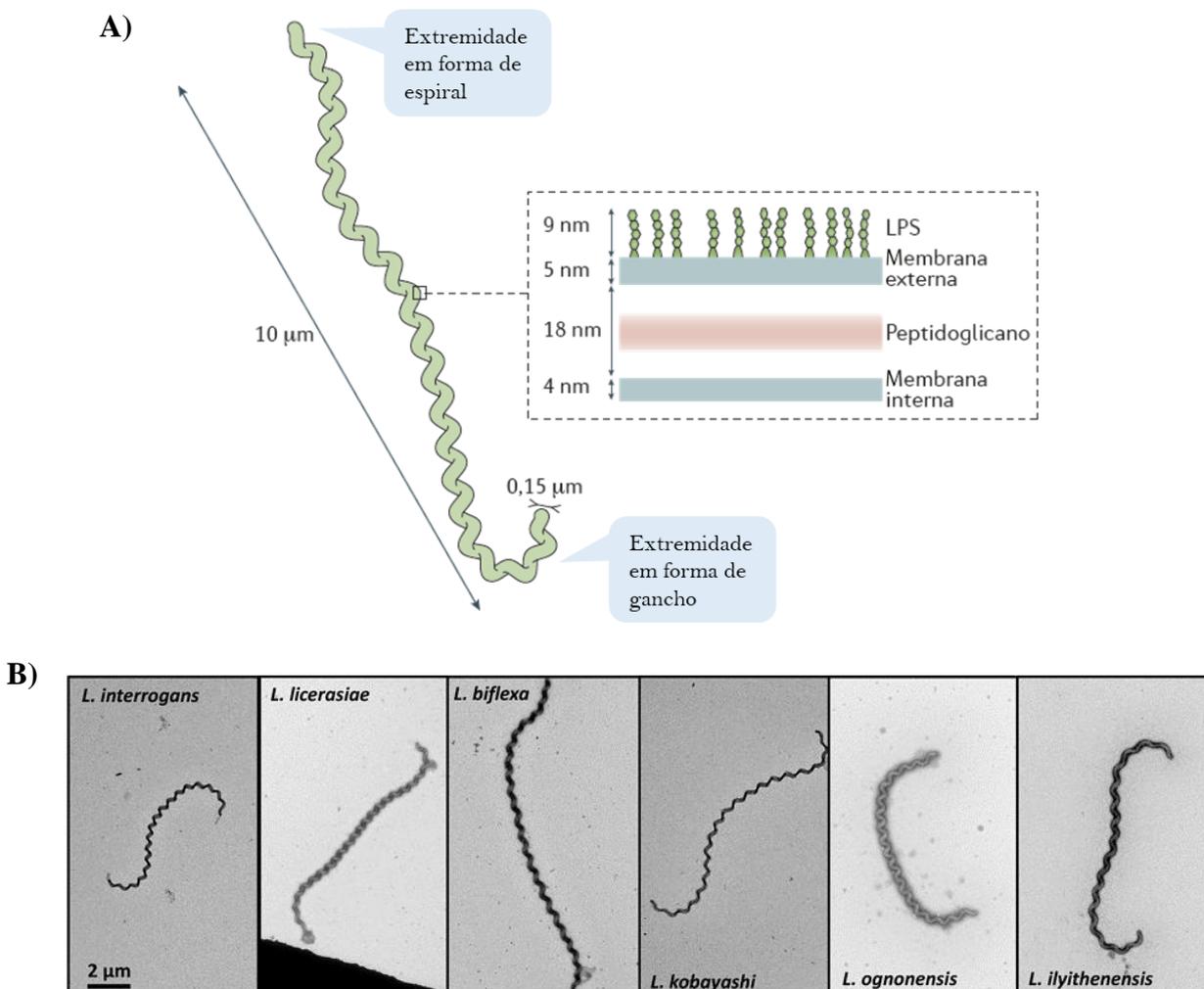


Figura 2. A) Representação da morfologia de *L. interrogans*, onde estão destacados os tipos de extremidades em forma de espiral e em forma de gancho característicos das leptospiros. Em destaque,

encontra-se a arquitetura do envelope celular, com as membranas interna e externa envolvendo a camada de peptidoglicano e a camada externa de LPS. Figura modificada de PICARDEAU, 2017. **B)** Microscopia eletrônica de transmissão das espécies representativas de cada grupo em *Leptospira*: P1 (*L. interrogans*), P2 (*L. licerasiae*), S1 (*L. biflexa*) e S2 (*L. kobayashii*, *L. ognonensis* e *L. ilyithenensis*). Figura modificada de VINCENT et al., 2019.

1.1.3. Características genômicas das diferentes espécies de *Leptospira*

As leptospiros possuem genomas relativamente grandes em relação a outras bactérias espiroquetas. Seus genomas variam de 3 Mb a 5 Mb em tamanho, com 35-45% de conteúdo GC, refletindo a alta variabilidade entre as espécies (PICARDEAU, 2017). Os genomas são geralmente organizados em dois cromossomos circulares com aproximadamente 4 Mb e 300 kb cada, além da presença de um plasmídeo adicional de 74 kb (chamado de p74) em *L. biflexa* (KO; GOARANT; PICARDEAU, 2009; PICARDEAU, 2015).

A variabilidade genética entre diferentes espécies de leptospiros é refletida diretamente na capacidade de sobrevivência em ambientes externos, assim como na adaptação a diversos hospedeiros, estando relacionada à perda de genes ancestrais para adaptação em vida livre e à aquisição de novos genes por duplicação genética ou transferência horizontal para a adaptação à patogenicidade em um ou mais hospedeiros específicos.

A espécie saprofítica *L. biflexa* (genoma de 3,96 Mb) sobrevive apenas em ambientes externos e possui mais genes relacionados a sensores ambientais e proteínas metabólicas do que a maioria das espécies patogênicas e intermediárias (PICARDEAU et al., 2008; RICALDI et al., 2012; PICARDEAU, 2017). A espécie patogênica *L. borgpetersenii* possui um dos menores genomas do gênero (entre 2,96-3,62 Mb), indicando que a espécie passou por um processo de perdas genéticas e especialização que a impedem de sobreviver em ambientes externos, dependendo do contato direto entre seus hospedeiros para ser transmitida (BULACH et al., 2006). Por outro lado, a espécie patogênica *L. interrogans* possui um dos maiores genomas (entre 4,40-5,02 Mb), pois além de adquirir genes relacionados a fatores de virulência, ela

também reteve alguns genes do ancestral de vida livre, permitindo sua sobrevivência no ambiente externo e ampliando as chances de transmissão entre hospedeiros (REN et al., 2003; KO; GOARANT; PICARDEAU, 2009; XU et al., 2016).

O genoma central (*core genome*) de *Leptospira spp.* contém aproximadamente 1.100 genes, envolvendo componentes do metabolismo (como enzimas para glicólise e ciclo do ácido tricarboxílico), transportadores do tipo ABC (“*ATP-binding cassette*”), lipoproteínas, genes flagelares e de quimiotaxia (PICARDEAU, 2018). Em geral, as espécies saprofíticas possuem o genoma central relativamente maior, enquanto que as espécies patogênicas possuem os maiores pangenomas (especialmente aquelas pertencentes ao grupo P1), os quais envolvem genes que modulam sua virulência e ainda se encontram abertos, ou seja, parte dos pangenomas ainda não foram caracterizados (VINCENT et al., 2019). Os maiores pangenomas nas espécies patogênicas são reflexos do maior número de genes espécie-específicos presentes nestas espécies (637 ± 129), em comparação às espécies intermediárias (418 ± 126) e às espécies saprofíticas (321 ± 90) (FOUTS et al., 2016).

Os genes estritamente envolvidos com patogênese e virulência permanecem amplamente desconhecidos, indicando que as espécies patogênicas possuem mecanismos únicos de interação com seus hospedeiros (KO; GOARANT; PICARDEAU, 2009; XU et al., 2016; PICARDEAU, 2018). Além disso, as espécies patogênicas abrigam significativamente maior porcentagem de regiões codificadoras não atribuídas às categorias funcionais conhecidas, incluindo os domínios (DUF) e COGs (*Clusters of Orthologous Groups*) com função desconhecida. Estas regiões podem representar pseudogenes, anotações genéticas erradas, assim como proteínas exclusivas do gênero *Leptospira* (VINCENT et al., 2019).

1.1.4. Fatores de virulência e adesinas

A indução das doenças infecciosas por bactérias patogênicas é resultado da ação de um conjunto diverso de fatores de virulência, os quais incluem não somente os componentes que promovem um claro efeito negativo em seus hospedeiros, como as endotoxinas e exotoxinas, mas também incluem fatores que habilitam os patógenos a escapar ou modular as respostas imunes inatas e adaptativas dos seus hospedeiros para sua vantagem replicativa no ambiente intra ou extracelular. Entre os diversos fatores de virulência existentes, os fatores de aderência ou adesinas são essenciais por favorecerem a resistência dos patógenos ao fluxo de fluidos no hospedeiro, permitindo assim sua fixação a células-alvo e a potencial invasão de tecidos e células, promovendo a disseminação e o estabelecimento da infecção nos diferentes hospedeiros (PIZARRO-CERDÁ; COSSART, 2006; KLINE et al., 2009; JOHNSON; JOHNSON, 2018).

Os principais genes associados a virulência em leptospiros incluem as famílias codificadoras das proteínas modificadoras de virulência (VM), proteínas ricas em leucina (LRR) e proteínas com domínio WGR (PF05406), além dos genes necessários para a síntese de vitamina B12 e ácido siálico, as peptidases colagenase, termolisina e esfingomielinases, a catalase KatE envolvida no estresse oxidativo e diversas proteínas de membrana externa (OMPs), como as proteínas Loa22, LigB e LipL32. A maioria destes genes são exclusivos das espécies patogênicas de *Leptospira* (FOUTS et al., 2016; PICARDEAU, 2017).

Assim como para a maioria dos patógenos, a adesão em todas as fases do ciclo infeccioso é essencial para o sucesso no estabelecimento da infecção das leptospiros em seus hospedeiros. As leptospiros patogênicas aderem-se a vários tipos de células, como fibroblastos, macrófagos, células endoteliais e células epiteliais renais. Além disso, foi demonstrada sua interação *in vitro* com componentes da matriz extracelular (MEC), incluindo fibronectina, colágenos, laminina e elastina, e com componentes de fluidos extracelulares, como

fibrinogênio, plasminogênio e reguladores do sistema complemento (MURRAY, 2015; PICARDEAU, 2017). Após a primeira demonstração da ligação entre fibronectina e uma proteína de 36 kDa, isolada da camada externa da variante virulenta *L. interrogans* sorovar Icterohaemorrhagiae (MERIEN et al., 2000), um extenso e crescente número de proteínas de leptospiros têm sido reportadas como potenciais adesinas que interagem a um ou mais componentes da MEC. Os principais fatores de virulência que apresentam afinidades múltiplas de ligação a componentes celulares, de MEC e de plasma do hospedeiro estão descritos a seguir.

As proteínas Lig (LigA e LigB) pertencem à família de proteínas bacterianas semelhantes à imunoglobulina, apresentando ligação a colágenos I e IV, laminina, fibronectina, tropoelastina, fibrinogênio, fator H, FHL-1, FHR-1 e C4bp; a proteína LigB também apresentou interação com colágeno III, elastina e heparina (CHOY et al., 2007; LIN et al., 2009; CASTIBLANCO-VALENCIA et al., 2012; CHING et al., 2012). Adicionalmente, a proteína LigA é uma das únicas moléculas candidatas a vacinas promissoras que conferem proteção contra a fase aguda da leptospirose (ADLER, 2015a).

As proteínas Len (LenA-F) é um grupo de seis proteínas que apresentam similaridade estrutural às endostatinas humanas, e todos os seus membros demonstraram ligação a fibronectina e laminina (STEVENSON et al., 2007). LenB apresenta interação adicional ao fator H, assim como a proteína LenA, previamente descrita como LfhA ou Lsa24, a qual também interage com plasminogênio e FHL-1 (BARBOSA et al., 2006; VERMA et al., 2006, 2010).

LipL32 é a lipoproteína mais abundante presente na membrana externa de leptospiros patogênicos e intermediárias, estando ausente nas espécies saprofitas. Foi demonstrada a ligação da proteína recombinante de LipL32 a laminina, colágenos I, IV, V e XX, fibronectina e plasminogênio (HAUK et al., 2008; HOKE et al., 2008; MURRAY et al., 2009; VIEIRA et al., 2010a; CHAEMCHUEN et al., 2011). Tais características, junto à demonstração de

expressão *in vivo* de seu gene, são fortes indicativos de que a LipL32 é um importante fator de virulência. Entretanto, estudos de Pinne e Haake mostraram que a proteína pode não estar localizada na superfície das leptospiros (MURRAY, 2013; PINNE; HAAKE, 2013).

Outras proteínas que se destacam pela interação com componentes de MEC e de plasma incluem: TlyC (laminina, fibronectina e colágeno IV) (CARVALHO et al., 2009), Lsa20/Lsa25/Lsa33 (laminina, plasminogênio e C4bp) (MENDES et al., 2011; DOMINGOS et al., 2012), Lsa63 (laminina e colágeno IV) (VIEIRA et al., 2010b), Lsa66/OmpL1 (laminina, fibronectina e plasminogênio) (OLIVEIRA et al., 2011; FERNANDES et al., 2012; PINNE; MATSUNAGA; HAAKE, 2012), rLIC12976 (laminina) (LIMA et al., 2013), enolase (plasminogênio) (NOGUEIRA et al., 2013), OmpL37/OmpL47 (laminina, fibronectina, elastina, fibrinogênio, colágeno III) (PINNE; CHOY; HAAKE, 2010).

Além das potenciais adesinas listadas acima, diversas proteínas recombinantes de *Leptospira* tiveram sua ligação a componentes da MEC demonstrada e publicada, revelando a intrigante multiplicidade e redundância das adesinas em leptospiros patogênicas. A influência da maioria destas proteínas na virulência das leptospiros ainda permanece obscura, já que mutantes em genes específicos (*lipL32*, *ligB*, *lenB*, *lenE*) retêm a virulência em modelos de leptospirose aguda e colonização (MURRAY, 2015). Estudos indicam que as leptospiros patogênicas podem se ligar de maneira mais eficiente a células do que a componentes da MEC, pela interação demonstrada *in vitro* a proteoglicanos de superfície e caderinas (BREINER et al., 2009; EVANGELISTA et al., 2014a; ROBBINS et al., 2015). Entretanto, a interação entre as potenciais adesinas a receptores celulares e aos diversos componentes de MEC e plasma do hospedeiro ainda precisa ser melhor elucidada, através de análises de perda e ganho de função em cepas de leptospiros mutantes patogênicas e saprofíticas, respectivamente (PICARDEAU, 2017).

1.2. Leptospirose

1.2.1. Ciclo de transmissão e diversidade de hospedeiros

A leptospirose é uma doença zoonótica negligenciada e globalmente difundida, causada por espécies patogênicas de *Leptospira*. O ciclo de transmissão de leptospiras é iniciado com a sua penetração no hospedeiro por cortes e abrasões na pele ou via membranas mucosas, através dos quais as leptospiras atingem a corrente sanguínea. A sua interação com componentes da MEC e receptores celulares dos hospedeiros adicionada à sua característica motilidade em movimentos espirais permitem a adesão e invasão das células endoteliais por leptospiras, rompendo barreiras teciduais (incluindo os tecidos conectivos mais viscosos) e estabelecendo rapidamente uma infecção sistêmica, atingindo todos os órgãos incluindo o cérebro. Em ratos, um dos principais hospedeiros de manutenção (definitivos), as leptospiras são eliminadas de todos os órgãos após infecção, exceto nos rins, órgãos imunoprivilegiados nos quais as leptospiras colonizam as microvilosidades do epitélio dos túbulos renais proximais, promovendo a liberação persistente de leptospiras ao ambiente através da urina (PICARDEAU, 2017; SAMROT et al., 2021).

Além dos ratos, a colonização de ampla variedade de hospedeiros reservatórios incluindo mamíferos pequenos, principalmente roedores e marsupiais, além de carreadores domésticos como gado, porcos e cachorros, também resulta na liberação de bactérias pela urina, podendo se manter por períodos variados durante sua vida, dependendo da espécie infectada. Em adição aos hospedeiros bastante conhecidos acima, quase todos os mamíferos podem ser carreadores de leptospiras patogênicas, incluindo cavalos, morcegos, primatas e até mamíferos aquáticos como cetáceos, leões marinhos e focas. Outros hospedeiros considerados atípicos incluem espécies aviárias, répteis, anfíbios e peixes (KO; GOARANT; PICARDEAU, 2009; ADLER; DE LA PEÑA MOCTEZUMA, 2010; GOARANT et al., 2019; CILIA et al., 2021).

Humanos são hospedeiros acidentais e sua exposição a leptospiros pode ocorrer através de contato direto com um animal infectado e/ou sua urina, ou através de contato indireto por meio de água ou solo contaminados com a urina de um animal infectado (HAAKE; LEVETT, 2015). Uma das possíveis razões para a susceptibilidade de humanos à leptospirose é a falha do receptor do sistema imune inato TLR4 humano em reconhecer o LPS de leptospiros, enquanto que o mesmo receptor reconhece o LPS de *E. coli* em concentrações extremamente baixas. Este fenômeno acontece devido a diferenças estruturais entre as moléculas de LPS das duas espécies. Curiosamente, o TLR4 de murinos reconhecem o LPS de leptospiros, enquanto que linhagens que não possuem este receptor (como camundongos C3H/HeJ) são suscetíveis à infecção letal por leptospiros patogênicas. Além do camundongo C3H/HeJ, hamsters e cobaias também apresentam leptospirose severa quando infectados e, portanto, constituem modelos importantes para o estudo das manifestações clínicas e danos causados pela leptospirose (EVANGELISTA; COBURN, 2010; HAAKE; ZÜCKERT, 2015).

1.2.2. Incidência e epidemiologia

A grande diversidade de mamíferos e outras classes de animais como reservatórios de leptospiros patogênicos é a principal causa da ampla distribuição geográfica da leptospirose, sendo altamente endêmica em regiões tropicais como Sul e Sudeste Asiáticos, Oceania, Subcontinente Indiano, Américas Latina e Caribenha e África Subsaariana Oriental. Frequentemente, os surtos em tais regiões estão associados com alta pluviosidade devido a mudanças climáticas, gerando enchentes e o consequente escoamento de leptospiros em solos para corpos de água, contribuindo para a dispersão das bactérias em águas contaminadas. Regiões tropicais englobam muitos países em desenvolvimento, onde há grande concentração de populações pobres que frequentemente sofrem com más condições sanitárias e de habitação, exacerbando o risco de exposição às leptospiros em comunidades rurais e urbanas

(EVANGELISTA; COBURN, 2010). A crescente urbanização descontrolada culminando no surgimento de favelas em diversas cidades de países em desenvolvimento também cria condições propícias para a transmissão da leptospirose por roedores (COSTA et al., 2015). Outro fator que contribui para a maior incidência da doença em países de clima tropical é a maior sobrevivência de leptospiras em ambientes úmidos e quentes. O risco de infecções em humanos também está relacionado a profissões de risco, como agricultores, veterinários e ordenhadores, além de exposições recreativas associadas a esportes aquáticos (LEVETT, 2001; HAAKE; LEVETT, 2015).

A cada ano, estima-se que a leptospirose cause 1,03 milhão de casos em todo o mundo, com um número estimado de quase 60 mil mortes, valores que aproximam a leptospirose dos casos de febres hemorrágicas e a colocam como a principal causa zoonótica de morbidade e mortalidade. Entretanto, tais estimativas são possivelmente subestimadas já que muitas das regiões endêmicas possuem subnotificação e dados limitados sobre a doença, além de refletir a incidência de casos severos de leptospirose, os quais representam uma pequena fração de todas os casos clínicos (5-15%). Adicionalmente, a leptospirose é uma das principais causas de febre aguda que são frequentemente e erroneamente diagnosticadas como malária, dengue ou influenza (COSTA et al., 2015). No Brasil, considerando o período entre 2007 e 2019, a média dos casos confirmados por ano é de 3.744, enquanto o número médio de óbitos atribuídos à leptospirose é de 330, correspondendo a 8,9 % dos casos confirmados (BRASIL, 2020).

Assim como outras doenças zoonóticas, a leptospirose afeta tanto a saúde humana quanto o desenvolvimento agrícola e pecuário dos países endêmicos, atingindo alguns de seus principais setores econômicos. As manifestações crônicas da doença no gado podem levar a infertilidade, perdas reprodutivas, redução na produção de leite e abortos, causando grandes perdas econômicas e/ou de recursos de subsistência (KO; GOARANT; PICARDEAU, 2009; GOARANT; DELLAGI; PICARDEAU, 2021). Na Nova Zelândia, o custo da leptospirose

humana e animal foi estimado em US\$ 18,80 milhões por ano (SANHUEZA et al., 2020), enquanto que no Brasil, os custos calculados com hospitalizações em 2007 chegaram a mais de R\$ 800 mil, um grande impacto para o sistema público de saúde no país (DE SOUZA et al., 2011).

1.2.3. Principais manifestações clínicas

O desenvolvimento da leptospirose e a progressão da severidade da doença são influenciados por fatores intrínsecos ao hospedeiro, pela quantidade de bactérias infectantes que penetraram no hospedeiro e pelos fatores de virulência apresentados pela cepa infectante, resultando em uma ampla variedade de manifestações clínicas (EVANGELISTA; COBURN, 2010; HAAKE; LEVETT, 2015).

A doença se inicia com um período de incubação, em que as leptospiras se multiplicam na corrente sanguínea e se espalham para os diversos órgãos, levando à fase de leptospiremia. As infecções por leptospiras são assintomáticas ou provocam sintomas leves e inespecíficos iniciais, como febre aguda associada a mialgia, dor de cabeça, calafrios, dores abdominais e sufusão conjuntival. Em alguns casos, podem provocar meningite asséptica (MARCHIORI et al., 2011).

O aumento no nível de anticorpos a partir de 7 dias da infecção leva à remissão da doença leve em quase 90 % dos casos, entretanto, parte dos pacientes pode desenvolver manifestações clínicas severas e tardias, cuja apresentação clássica é a doença de Weil, caracterizada por icterícia profunda, disfunção renal grave e hemorragias pontuais. Há também uma forma de doença severa emergente, a síndrome pulmonar hemorrágica associada à leptospirose, caracterizada por dificuldade respiratória e hemorragia pulmonar maciça súbita (LEVETT, 2001; KO; GOARANT; PICARDEAU, 2009). A mortalidade dos casos de doença de Weil e da síndrome pulmonar hemorrágica associada à leptospirose é alta (> 10 % e > 50 %,

respectivamente), mesmo quando os pacientes são adequadamente tratados (MARCHIORI et al., 2011; DONG; CHEN, 2021).

1.2.4. Diagnóstico

A inespecificidade dos sintomas iniciais da leptospirose geralmente resulta em diagnósticos errôneos onde a confirmação laboratorial é ausente e em regiões em que o diagnóstico clínico é frequentemente atribuído a doenças mais comuns como influenza, malária e dengue. O diagnóstico da leptospirose é principalmente realizado com o teste de aglutinação microscópica (MAT), considerado padrão ouro dos métodos diagnósticos e possui alta sensibilidade para detecção e identificação de anticorpos contra sorogrupos específicos. O teste é baseado na reação entre soros de pacientes e antígenos em suspensão a partir de cultivos de diferentes sorovares de leptospiros. As desvantagens do MAT incluem o tempo adequado para a realização do teste que precisa coincidir com a fase imune da doença (pelo menos 6 dias após infecção), além da necessidade do cultivo de múltiplos sorovares em cultura, o que requer laboratórios especializados e custos potencialmente altos de manutenção. Devido à complexidade do MAT, outros testes sorológicos e moleculares têm sido desenvolvidos para um diagnóstico mais rápido da leptospirose (HAAKE; LEVETT, 2015; SAMROT et al., 2021).

O ensaio imune baseado em microesferas magnéticas (MIA) utiliza a tecnologia xMap® (Luminex) para a detecção e classificação de anticorpos, incluindo IgM, específicos a antígenos de leptospiros, permitindo a análise de alto número de amostras por dia (WYNWOOD et al., 2015). A detecção de anticorpos IgM e IgG específicos de leptospiros também pode ser realizada por ELISA e ensaio de hemaglutinação indireta (TERPSTRA; LIGTHART; SCHOONE, 1985; LEVETT; WHITTINGTON, 1998). Apesar dos ensaios citados frequentemente apresentarem sensibilidade e especificidade consideráveis, os baixos níveis de

anticorpos nas fases iniciais da doença podem levar ao diagnóstico de falsos negativos (KHAKI, 2016).

O ensaio do tipo “*dipstick*” é um método imunocromatográfico para a triagem rápida e diagnóstico de pacientes com suspeita de leptospirose, sem a necessidade de laboratórios e técnicos especializados. O teste “*dipstick*” LEPTO detecta anticorpos IgM específicos e foi testado em diferentes condições epidemiológicas e clínicas, demonstrando alta sensibilidade e comparáveis ao IgM-ELISA (HATTA et al., 2000; KHAKI, 2016). No Brasil, o Ministério da Saúde anunciou recentemente a implantação no Sistema Único de Saúde (SUS) do teste TR DPP® Leptospirose, semelhante ao LEPTO, a partir da tecnologia já testada para a detecção de leishmaniose e desenvolvida pela unidade produtora de imunobiológicos da Fiocruz (Bio-Manguinhos) (TEIXEIRA et al., 2019).

A reação em cadeia da polimerase quantitativa (qPCR) pode detectar a presença de leptospirosas em amostras de urina ou sangue durante os estágios iniciais da doença, quando o título de anticorpos ainda não é alto o suficiente. O teste é altamente sensível e específico, principalmente na aplicação do ensaio de qPCR multiplex, capaz de identificar mais alvos e aumentar a especificidade da reação (BAL et al., 1994; MERIEN et al., 2005; AHMED et al., 2012). Entretanto, o diagnóstico baseado em PCR não identifica o sorovar de *Leptospira* infectante nos pacientes, informação importante para a epidemiologia e saúde pública. Portanto, a confirmação por MAT ainda é essencial para a identificação precisa nesses casos (HAAKE; LEVETT, 2015).

1.2.5. Desenvolvimento de vacinas

Vacinas baseadas em leptospirosas inteiras e inativadas (bacterinas) têm sido rotineiramente utilizadas desde a primeira demonstração de proteção experimental, sendo administradas a animais domésticos e pecuários, além de serem usadas para imunizar

populações humanas (ADLER, 2015b). Entretanto, tais vacinas não foram licenciadas para uso humano generalizado devido a altas taxas de reações adversas e, como a imunidade ativada por essas vacinas é direcionada principalmente ao LPS da leptospira utilizada, a resposta é sorovar-específica e de curta duração, pois o LPS é um antígeno T-independente e, portanto, não induz memória imunológica (ADLER, 2015a). Adicionalmente, as vacinas baseadas em células inativadas são geralmente ineficientes para prevenir a colonização e o transporte das bactérias, limitando seu uso como intervenção para bloquear a transmissão da doença (KO; GOARANT; PICARDEAU, 2009).

Devidos às limitações apresentadas pelas vacinas de bacterina, estudos foram focados no desenvolvimento de vacinas de subunidades, constituídas por proteínas associadas à superfície das leptospiros e conservadas entre vários sorovares. A primeira vacina de subunidades desenvolvida era baseada nas proteínas recombinantes OmpL1 e LipL41, a qual conferiu proteção significativa no modelo letal de leptospirose em hamsters. Entretanto, a demonstração de proteção foi observada em apenas um terço dos experimentos realizados (HAAKE et al., 1999). Mesmo assim, a partir deste primeiro estudo, diversas proteínas recombinantes foram avaliadas como candidatas a vacinas.

As proteínas Lig (LigA e LigB) foram as proteínas mais extensivamente utilizadas como candidatas a vacinas. A região C-terminal da proteína LigA é uma das candidatas a vacina contra leptospirose mais bem aceita, sendo bem sucedida em diversos modelos como vacina de subunidade, de DNA, encapsulada, lipidada, entre outras. Entretanto, a proteína LigA não está presente em todas as espécies patogênicas de *Leptospira*, limitando sua proteção cruzada. Em menor extensão, a região N-terminal da proteína LigB também demonstrou proteção significativa como vacina de subunidade ou de DNA em modelo de hamster (GRASSMANN et al., 2017; SILVEIRA et al., 2017).

LipL32 é a proteína imunopredominante em leptospiros patogênicas, sendo avaliada como candidata vacinal em diversos estudos, mas apenas uma pequena parcela deles demonstraram proteção significativa contra leptospirose, possivelmente devido à natureza subcelular da LipL32 e da ausência de glicosilação em proteínas recombinantes produzidas em sistemas heterólogos (ODIR et al., 2011; ADLER, 2015a).

Métodos bioinformáticos também têm auxiliado na busca de novos candidatos vacinais, como foi o exemplo dos estudos de Faisal e colaboradores utilizando três OMPs hipotéticas (Lp 0607, Lp1118 e Lp1454), as quais somente conferiram proteção significativas contra o desafio em hamsters quando foram combinadas e encapsuladas em lipossomos (FAISAL et al., 2009a, 2009b). A técnica de vacinologia reversa já é aplicada em muitas doenças infecciosas e diversos antígenos promissores de leptospiros foram identificados por esta metodologia, como é o caso da proteína hipotética LemA que induziu proteção parcial, mas bastante significativa em hamsters quando administrada usando DNA e proteína (estratégia de “*prime boost*”) (HARTWIG et al., 2013; DELLAGOSTIN et al., 2017).

De fato, a combinação de diferentes antígenos em vacinas contra a leptospirose tem se mostrado uma tendência entre os diversos estudos realizados, entretanto, os melhores candidatos ainda necessitam ser melhor elucidados. A diversa composição da membrana externa de leptospiros patogênicas foi utilizada como base de uma abordagem realizada recentemente, em que vacinas bivalentes com redução no conteúdo de LPS conferiram proteção contra sorovares virulentos heterólogos, utilizando as cepas virulentas de *L. interrogans* Canicola e Copenhageni (LAURETTI-FERREIRA et al., 2020).

Apesar de alguns resultados promissores, muitos estudos relatados na literatura apresentam doses de desafio inadequadas, ausência de reprodutibilidade e análises estatísticas inapropriadas (ADLER, 2015a; FELIX et al., 2020). É importante que o desenvolvimento de

novas vacinas seja baseado em uma melhor compreensão da membrana externa de leptospiras, assim como na capacidade antigênica e imunogênica de seus componentes.

1.3. Resultados prévios – Identificação de novas adesinas por *phage display*

A tecnologia de *phage display* foi uma das primeiras ferramentas genéticas descritas para o estudo de interações proteína-ligante (SMITH, 1985). A técnica envolve a construção de bibliotecas a partir de fragmentos randômicos de DNA exógenos clonados fusionados a genes codificando proteínas capsidiais de fagos, resultando na expressão da sua proteína quimérica na superfície do fago utilizado. Esta abordagem tem sido aplicada com sucesso no mapeamento de novos epítomos, na construção e desenvolvimento de anticorpos monoclonais e na seleção de agonistas e antagonistas de receptores específicos. O uso do *phage display* é particularmente útil na elucidação das interações infecciosas entre microrganismos patogênicos e seus hospedeiros, revelando-se bem sucedida na identificação de adesinas microbianas e virais, componentes essenciais para a colonização dos patógenos e sucesso no estabelecimento de infecções (NILSSON et al., 1998; WILLIAMS et al., 2002; MULLEN et al., 2006; TAN; HO, 2014; CONNOR et al., 2016; FEVRE; SCHEEPMAKER; HAAS, 2017).

Alguns grupos adotaram esta metodologia no estudo de bactérias espiroquetas, com a identificação e seleção das potenciais adesinas p66, OspC, VIs, OspF e BmpD em *Borrelia burgdorferi* (COBURN et al., 1999; ANTONARA et al., 2007), e na identificação das proteínas imunogênicas TmpA, Ttm e PrrA na cepa V1 proximamente relacionada a *Treponema phagedenis* (ROSANDER et al., 2011). Mais recentemente, foram identificadas lipoproteínas putativas (LIC10508 e LIC13411) e proteínas hipotéticas conservadas (LIC12341 e LIC11574) em *L. interrogans*, das quais as proteínas LIC11574 e LIC13411 demonstraram ligação a VE-caderina (receptor de células endoteliais para *L. interrogans*) (EVANGELISTA et al., 2014b).

Utilizando a técnica de *phage display*, nosso grupo demonstrou a interação da proteína LigB à heparina/heparan sulfato, principalmente de sua porção C-terminal recombinante (rLB-Ct), a qual também apresentou ligação a fibroblastos de camundongos da linhagem A31 e às células epiteliais renais LLC-PK1 (porco) e VERO (macaco) (CHING et al., 2012). Da mesma forma, a lipoproteína LIC12976 foi identificada, e a proteína recombinante rLIC12976 apresentou capacidade de ligação às mesmas células de mamíferos citadas acima e à laminina (LIMA et al., 2013).

Os resultados de *phage display* mais recentes do nosso grupo envolveram a seleção de possíveis adesinas em células epiteliais renais (células VERO), a partir de bibliotecas construídas com o DNA genômico de *L. interrogans*, resultando na identificação de três proteínas denominadas arbitrariamente por LepA962, LepA388 e LepA069. A proteína LepA962 é uma lipoproteína hipotética, a proteína LepA388 é uma proteína hipotética presente apenas nas espécies patogênicas de *Leptospira*, e a proteína hipotética LepA069 foi identificada em dois fagos que apresentaram sequências sobrepostas da proteína, sugerindo um provável indício de enriquecimento da proteína LepA069 durante as rodadas de seleção nas células (FERREIRA, 2015).

Para a presente tese de Doutorado, nosso grupo decidiu dar continuidade à caracterização das proteínas previamente selecionadas pela técnica de *phage display*, por estas representarem importantes alvos de estudo para a determinação de sua possível função como adesinas e como candidatas vacinais, além de serem antígenos potenciais para o diagnóstico da leptospirose.

2. OBJETIVOS

2.1. Objetivos geral e específicos

O objetivo geral do presente trabalho é a caracterização das proteínas hipotéticas conservadas LepA962, LepA388 e LepA069 codificadas por *Leptospira interrogans* e previamente selecionadas através da técnica de *phage display*, avaliando as seguintes possibilidades:

- Atividade das proteínas como adesinas;
- Identificação e delimitação de possíveis domínios funcionais.

3. MATERIAL E MÉTODOS

3.1. Organismos, condições gerais de cultivo e vetores

A bactéria *Leptospira interrogans* sorovar Copenhageni cepa Fiocruz L1-130 (Taxonomy ID: 267671) empregada neste estudo foi originalmente obtida da Faculdade de Medicina Veterinária e Zootecnia (Universidade de São Paulo, Brasil). A manutenção da virulência e as condições de cultivo das leptospiros foram realizadas segundo descrito por Da Silva e colaboradores. Em resumo, bactérias isoladas de hamsters infectados foram cultivadas em meio líquido EMJH (ELLINGHAUSEN; MCCULLOUGH, 1965) suplementado com BSA, sob condições de aerobiose e temperatura de 30 °C. A fim de garantir as características virulentas das células de *L. interrogans*, as bactérias foram utilizadas até a terceira passagem em meio EMJH após isolamento dos animais infectados. A concentração bacteriana de cada cultura foi determinada por contagem das células em câmara de Petroff-Hausser, sob microscópio de campo escuro (DA SILVA et al., 2012).

As cepas de bactérias *Escherichia coli* DH10B competentes foram utilizadas para clonagens, subclonagens e ampliações de vetores. As células foram cultivadas em meio LB (10 g de tripton, 5 g de extrato de levedura, 10 g de NaCl e 200 µL de NaOH 5 N para cada 1 L de meio), contendo os antibióticos adequados, em temperatura de 37°C e sob agitação. As cepas de *E. coli* competentes utilizadas para a produção das proteínas recombinantes foram: BL21 (DE3), BL21 Star (DE3) pLysS, *E. coli* BL21 C43, *E. coli* Origami (DE3) e *E. coli* ArcticExpress (DE3), as quais foram cultivadas de modo semelhante à cepa DH10B. Também utilizada para a produção de proteínas recombinantes, a cepa competente *E. coli* BL21-SI foi cultivada em meio LBON (meio LB sem NaCl) contendo os antibióticos adequados, em temperatura de 30°C e sob agitação (DONAHUE; BEBEE, 1999).

As linhagens de células eucarióticas selecionadas para os experimentos de ligação em superfície celular foram as células epiteliais de rim VERO, originadas de macacos *Cercopithecus aethiops* (SIMIZU et al., 1967).

Finalmente, os vetores utilizados para a clonagem e expressão de sequências codificadoras das proteínas recombinantes incluíram o sistema comercial *pGEM®-T Easy Vector* (Promega), e os vetores de expressão pAE (RAMOS et al., 2004), pAZ (BUTERA et al., 2005) e pSUMO_{Ulp1} (SHIMOKAWA-FALCÃO et al., 2017).

3.2. Preparação de bactérias *E. coli* competentes

Segundo descrito na literatura (SAMBROOK; RUSSELL, 2001), as bactérias *E. coli* DH10B foram preparadas para eletrocompetência, enquanto que as bactérias *E. coli* Origami (DE3) e ArcticExpress (DE3) foram preparadas para quimiocompetência, com algumas modificações.

3.2.1. Bactérias *E. coli* eletrocompetentes

Para cada lote, uma colônia isolada de *E. coli* DH10B foi coletada a partir de uma placa fresca contendo meio LB-ágar (meio LB acrescido de 15 g de bacto-ágar/L de meio), e inoculada em 25 mL de meio LB sem antibióticos para cultivo a 37°C por aproximadamente 16 h, em um agitador orbital (200-250 rpm). Em 500 mL de meio LB, foram transferidos 6,5-10 mL do pré-inóculo inicial para novo cultivo a 37°C sob agitação, até ser atingida uma DO_{600nm} ideal entre 0,5 e 0,7. Os cultivos foram então separados em tubos de centrifugação estéreis de 150 mL e mantidos em gelo por 20 min. A partir deste passo, as bactérias foram sempre mantidas em gelo ou a 4 °C, assim como todas as soluções utilizadas foram previamente resfriadas. Para centrifugação, foi utilizado um rotor basculante a 3.150-3.200 x g por 20 min a 4 °C, e após o descarte dos sobrenadantes, os precipitados celulares foram lavados com água

ultrapurificada estéril (125 mL/tubo). Em seguida, as suspensões foram novamente incubadas em gelo por 20 min, centrifugadas e os precipitados celulares resultantes foram lavados por mais três vezes com uma solução de glicerol 10% estéril. Durante a última lavagem, o conteúdo de células foi reunido para um tubo cônico de 50 mL estéril e o precipitado celular final foi homogeneizado em 0,6-1 mL da solução de glicerol 10% estéril para separação em alíquotas de 60 µL, as quais foram imediatamente congeladas em contato com gelo seco e posteriormente armazenadas a -80 °C.

3.2.2. Bactérias *E. coli* quimiocompetentes

A fim de manter seus plasmídeos característicos, os meios de cultura para as bactérias *E. coli* Origami (DE3) foram acrescidos de 15 µg/mL de canamicina e 12,5 µg/mL de tetraciclina, enquanto que os meios de cultura para as bactérias *E. coli* ArcticExpress (DE3) foram acrescidos de 20 µg/mL de gentamicina e 75 µg/mL de estreptomicina. Seguindo etapas semelhantes às descritas no item 3.2.1, pré-inóculos de 30 mL de meio LB contendo os antibióticos adequados foram obtidos e utilizados para novos cultivos (adição de 16,5 mL para cada 500 mL de meio LB com os mesmos antibióticos).

Após ser atingida a DO_{600nm} ideal de 0,5-0,7, cada cultivo foi separado em tubos de centrifugação estéreis e mantido em gelo por 25 min. Assim como descrito anteriormente, a temperatura foi mantida em torno de 4 °C a partir deste passo para a melhor preservação das bactérias, assim como todas as soluções utilizadas foram previamente resfriadas. Cada cultivo foi então centrifugado em rotor basculante a 3.000 x g por 15 min a 4 °C, e após o descarte dos sobrenadantes, os precipitados celulares foram homogeneizados em 150 mL de solução 0,1 M $MgCl_2$ estéril. Em seguida, as suspensões foram centrifugadas a 2.000 x g por 15 min a 4 °C, e os precipitados celulares foram homogeneizados em 150 mL de solução 0,1 M $CaCl_2$ estéril. Após incubação em gelo por 20 min, as suspensões foram centrifugadas e os precipitados

celulares resultantes foram homogeneizados em 45 mL de solução 0,085 M CaCl₂ e 20% de glicerol estéril. Após centrifugação a 1.000 x g por 15 min a 4°C, o precipitado celular final foi homogeneizado em 1 mL da solução 0,085 M CaCl₂ e 20% de glicerol estéril para separação em alíquotas de 55 µL, as quais foram imediatamente congeladas em contato com gelo seco e posteriormente armazenadas a -80 °C.

Outras bactérias quimiocompetentes utilizadas neste estudo foram adquiridas comercialmente: *MAX Efficiency™ DH5α Competent Cells* (Invitrogen).

3.3. Desenho de oligonucleotídeos para clonagens em vetores de expressão e para reações de qPCR

A fim de melhor analisar as proteínas hipotéticas selecionadas previamente pelas técnicas de *phage display* e *biopanning*-BRASIL: LepA069, LepA962 e LepA388 (FERREIRA, 2015), novos oligonucleotídeos foram desenhados para a clonagem e expressão de suas sequências codificadoras. Para cada proteína, foi clonada a porção referente à sequência apresentada pelo fago (identificadas com sufixo “_Phg”), além de outras sequências (detalhadas no item 4. RESULTADOS). Os vetores de expressão escolhidos (pAE, pAZ e pSUMO_{U1p1}) possuem um promotor derivado do fago T7 e produzem fenótipos de resistência à ampicilina (pAE e pAZ) ou ao cloranfenicol (pSUMO_{U1p1}) nas bactérias eficientemente transformadas por estes vetores. As construções nestes vetores promovem a adição de uma sequência de seis histidinas à porção N-terminal (pAE e pAZ) ou à porção C-terminal (pSUMO_{U1p1}) das proteínas recombinantes, possibilitando sua purificação por meio da cromatografia de afinidade por íons metálicos imobilizados (IMAC), após a sua expressão induzida em bactérias. O vetor pAZ é derivado do vetor pAE e, além da etiqueta de histidinas, o vetor possibilita a fusão de proteínas recombinantes à EGFP (BUTERA et al., 2005) (Figuras S1 a S3).

Adicionalmente, pares de oligonucleotídeos também foram desenhados para a realização de reações em cadeia da polimerase quantitativas utilizando transcriptase reversa (RT-qPCR), a fim de avaliar a expressão gênica das proteínas hipotéticas (ver item 3.4).

3.4. Avaliação da expressão gênica das proteínas hipotéticas por RT-qPCR

A expressão dos genes que expressam as proteínas hipotéticas LepA069 e LepA388 foi avaliada em amostras de fígado de camundongos pertencentes à linhagem C3H/HeJ após três dias de infecção por bactérias virulentas de *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130. Previamente, as amostras de fígado foram imediatamente congeladas em nitrogênio líquido após coleta dos animais e armazenadas a -80 °C até a extração do RNA total, utilizando o reagente Trizol (Invitrogen). Posteriormente, os RNAs foram tratados com a enzima DNase (Fermentas) a 37 °C por 30 min, seguida pela reação de transcrição reversa utilizando a enzima *SuperScript III* (Invitrogen) para a obtenção de cDNAs. Como controle negativo, a mesma cepa de *L. interrogans* cultivada por 5 dias em meio EMJH a 30 °C teve seu RNA extraído, purificado e retrotranscrito em cDNA. As amostras de cDNA foram mantidas a -80 °C até seu uso nas reações de qPCR, compostas pelo reagente *SYBR® Green Master Mix* diluído 1X, cDNA (2 ng do controle extraído de *L. interrogans* em cultura, ou 96 ng das amostras extraídas de fígado) e 0,42 µM de cada oligonucleotídeo específico senso (*forward*) e antisense (*reverse*) (Tabela S1), referentes aos genes que expressam as proteínas hipotéticas LepA069 e LepA388, além da proteína LipL32 e do rRNA 16S como controles. As reações de qPCR foram realizadas em triplicata no equipamento *Applied Biosystems 7300 Real-Time PCR System* (Applied Biosystems), utilizando as seguintes etapas de reação: etapa 1 de 50 °C/20 min, etapa 2 de 95 °C/10 min, etapa 3 com 40 ciclos de 95 °C/15 s – 60 °C/1 min – 72 °C/1 min, e etapa 4 para análise das curvas de dissociação com 95 °C/15 s – 60 °C/15 s. Após o término das reações e extração dos dados, os níveis de expressão dos genes selecionados foram analisados a partir dos

valores de limiar de ciclo (C_T , ou *cycle threshold*), comparados aos valores de C_T obtidos com o gene de expressão basal 16S e utilizando o método delta-delta expresso na equação $2^{-\Delta\Delta C_T}$ (LIVAK; SCHMITTGEN, 2001; PFAFFL, 2001).

3.5. Análises *in silico* das proteínas hipotéticas

As proteínas hipotéticas LepA069, LepA962 e LepA388 também tiveram suas sequências aminoacídicas analisadas por programas de predição de domínios e estruturas tridimensionais (Tabela 1), com o objetivo de melhor guiar os desenhos de novas proteínas recombinantes para posteriores análises funcionais, ou para comparar os modelos encontrados com os resultados já obtidos. As sequências das proteínas recombinantes desenhadas a partir dos modelos gerados foram clonadas de acordo com o item 3.6.

Tabela 1

Ferramentas de predição e análise de domínios e estruturas tridimensionais utilizadas.

Ferramenta	Website	Referências
SWISS-MODEL	swissmodel.expasy.org	(GUEX et al., 2009; BERTONI et al., 2017; BIENERT et al., 2017; WATERHOUSE et al., 2018; STUDER et al., 2020)
I-TASSER	zhanglab.ccmb.med.umich.edu/I-TASSER	(ROY; KUCUKURAL; ZHANG, 2010; YANG et al., 2015; YANG; ZHANG, 2015)
Phyre2	sbg.bio.ic.ac.uk/phyre2	(KELLEY et al., 2015)
ThreaDom	zhanglab.ccmb.med.umich.edu/ThreaDom	(XUE et al., 2013)
HHpred	toolkit.tuebingen.mpg.de/tools/hhpred	(HILDEBRAND et al., 2009; ZIMMERMANN et al., 2018)
PyMol	pymol.org/2	(DELANO, 2002)
UCSF Chimera	www.cgl.ucsf.edu/chimera	(PETTERSEN et al., 2004)
Robetta	rosetta.bakerlab.org	(BAEK et al., 2021)

3.6. Clonagem e subclonagem das sequências selecionadas aos vetores de expressão

Sequências selecionadas de cada proteína foram amplificadas por PCR convencional, utilizando um par de oligonucleotídeos específicos (Tabela S2), um vetor recombinante contendo a sequência completa de cada proteína anteriormente clonada e verificada por sequenciamento de DNA (pAE-LepA962F e pAE-LepA388F; FERREIRA, 2015), e seguindo os protocolos das enzimas *Phusion™ High-Fidelity DNA Polymerase* (Thermo Scientific), *Platinum™ Taq DNA Polymerase High Fidelity* ou do kit *Platinum™ PCR SuperMix High Fidelity* (Invitrogen). Cada produto amplificado foi analisado por eletroforese em gel de agarose (ver item 3.14.1) e posteriormente extraído do gel, segundo os protocolos dos kits *QIAquick Gel Extraction Spin Protocol* (QIAGEN), *illustra™ GFX™ PCR DNA and Gel Band* (GE Healthcare) ou *Wizard® SV Gel and PCR Clean-Up System* (Promega). Devido à reação com a polimerase de DNA *Phusion* gerar extremidades cegas nas amplificações, os produtos amplificados por esta enzima foram também submetidos a uma reação de adenilação com a enzima *Taq DNA Polymerase* recombinante (Invitrogen e Thermo Scientific) e 0,2 mM de dATP, a 72°C por 20 min.

Os produtos amplificados, purificados e tratados foram então utilizados para ligação ao vetor *pGEM®-T Easy* (Promega), com relação molar vetor:inserto de 1:3 e segundo o protocolo do fabricante. As amostras ligadas foram adicionadas a alíquotas de bactérias *E. coli* DH10B eletrocompetentes para a transformação por eletroporação a 2,5 kV (configuração pré-programada no eletroporador *MicroPulser* - Bio-Rad), seguida pela recuperação das amostras pulsadas em meio SOC (2% triptona, 0,5% extrato de levedura, 10 mM NaCl, 2,5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄ e 20 mM glicose), a 37 °C e sob agitação de 180-200 rpm por 45-60 min. Bactérias *E. coli* DH5α quimiocompetentes adquiridas comercialmente (*MAX Efficiency™ DH5α Competent Cells*; Invitrogen) também foram utilizadas para a

transformação com as amostras ligadas ao vetor pGEM®-T, seguindo o protocolo de transformação por choque térmico recomendado pelo fabricante.

As bactérias transformadas foram adicionadas a placas LB-ágar contendo 100 µg/mL de ampicilina (resistência presente no vetor pGEM®-T), 40 µg/mL de X-gal e 0,1 mM de IPTG, as quais foram incubadas a 37 °C por aproximadamente 16 h. Colônias brancas representam ligações bem sucedidas devido à quebra do gene LacZ (presente no vetor pGEM®-T), impossibilitando a expressão do gene codificador da β-galactosidase e a geração da coloração azul, característica de colônias negativas (SAMBROOK; RUSSELL, 2001). Algumas colônias brancas foram então selecionadas para amplificação em meio LB acrescido de ampicilina, sendo também adicionadas a reações de PCR para verificação da inserção dos fragmentos no vetor, seguindo o protocolo de enzimas *Taq DNA Polymerase* recombinantes (Invitrogen e Thermo Scientific) e com a adição do par de oligonucleotídeos específico para cada sequência clonada (Tabela S2). As amostras positivas foram confirmadas pela análise de eletroforese em gel de agarose e os cultivos correspondentes foram submetidos à extração e purificação dos vetores recombinantes, de acordo com os protocolos dos kits *QIAprep Spin Miniprep* (QIAGEN) ou *Wizard® Plus SV Minipreps DNA Purification System* (Promega). A fim de confirmar a integridade das sequências amplificadas e clonadas, as amostras purificadas também foram analisadas por sequenciamento de DNA pelo sequenciador *ABI 3500 Genetic Analyser* (Applied Biosystems), além de reações de digestão combinadas com enzimas de restrição e análise em gel de agarose.

Os vetores de expressão pAE, pAZ e pSUMO_{U1p1} também foram amplificados e purificados de acordo com os protocolos dos kits *QIAfilter Plasmid Purification Maxi* (QIAGEN) e *Wizard® Plus SV Minipreps DNA Purification System* (Promega). De forma semelhante às sequências clonadas, a integridade dos vetores foi avaliada através de reações de

digestão com enzimas de restrição, e posterior análise das amostras digeridas por eletroforese em gel de agarose.

A partir da análise dos vetores de expressão e das construções em pGEM®-T, os melhores clones foram escolhidos e submetidos à digestão dupla com duas das seguintes enzimas de restrição: XhoI, BamHI, KpnI e HindIII, de acordo com os sítios de restrição clonados (Tabela S2). Após a digestão, os fragmentos resultantes foram separados por eletroforese em gel de agarose e as bandas correspondentes foram excisadas para a extração e purificação do gel, como descrito acima neste item.

As novas reações de ligação foram realizadas com os fragmentos e vetores digeridos na relação molar vetor:inserto de 1:3, utilizando o protocolo da enzima *T4 DNA ligase* (Thermo Scientific). As amostras resultantes foram usadas para transformação de bactérias *E. coli* DH10B eletrocompetentes ou bactérias *E. coli* DH5 α quimiocompetentes e as colônias resultantes foram avaliadas por PCR. Como descrito acima neste item, os clones positivos foram cultivados para sua amplificação em meio LB com antibiótico adequado e posterior purificação de seus vetores recombinantes, os quais foram avaliados por sequenciamento de DNA e digestões com enzimas de restrição.

3.7. Produção das proteínas recombinantes em diferentes cepas de bactérias *E. coli*

Após a avaliação de cada construção nos vetores de expressão escolhidos, a produção das proteínas recombinantes foi inicialmente testada em pequena escala, utilizando diferentes cepas de bactérias *E. coli* a fim de obter as melhores condições para altos níveis de rendimento das proteínas, mantendo o dobramento correto das estruturas e, conseqüentemente, sua solubilidade dentro das células bacterianas antes de sua extração e purificação. As seguintes cepas bacterianas quimiocompetentes foram utilizadas para este fim: BL21 (DE3), BL21 Star (DE3) pLysS, BL21-SI, BL21 C43 (DE3), Origami (DE3) e ArcticExpress (DE3).

Cada construção em pAE, pAZ ou pSUMO_{Ulp1} foi adicionada a uma alíquota das células competentes e, em seguida, as amostras foram incubadas em gelo durante 30 min, passando por um choque térmico a 42 °C por 2 min e uma nova incubação em gelo por mais 5 min. Para a recuperação das bactérias após a transformação, foram adicionados 350 µL de meio LBON (BL21-SI) ou LB (restante das cepas citadas acima) e as amostras foram incubadas a 30 °C (BL21-SI) ou 37 °C (restante das cepas), sob agitação (180-200 rpm) durante 45-60 min. Alíquotas de cada cultivo resultante foram adicionadas a placas de LBON-ágar (BL21-SI) ou LB-ágar (restante das cepas), contendo os antibióticos adequados de acordo com a cepa bacteriana utilizada e o vetor de expressão (Tabela 2), as quais foram incubadas a 30 °C (BL21-SI) ou 37 °C (restante das cepas) durante aproximadamente 16 h.

Tabela 2

Cepas bacterianas utilizadas e condições de cultivo correspondentes para amplificação dos vetores e produção das proteínas recombinantes.

Cepa de <i>E. coli</i>	Antibiótico(s) para cultivo	Indutor para expressão de proteínas recombinantes
BL21 Star (DE3) pLysS	Cloranfenicol (30 µg/mL) + resistência do vetor de expressão	0,5-1 mM IPTG
BL21-SI	Apenas a resistência do vetor de expressão	200-300 mM NaCl
BL21 C43 (DE3)	Apenas a resistência do vetor de expressão	1 mM IPTG
Origami (DE3)	Tetraciclina (12,5 µg/mL) + resistência do vetor de expressão	1 mM IPTG
ArcticExpress (DE3)	Gentamicina (20 µg/mL) Estreptomicina (75 µg/mL) + resistência do vetor de expressão	0,05-0,5 mM IPTG

A partir das transformações em cada cepa hospedeira, algumas colônias foram selecionadas e adicionadas a 5-10 mL de meio de cultura (LBON ou LB) acrescidos dos antibióticos adequados para incubação a 30 °C ou 37 °C, sob agitação (180-200 rpm) por

aproximadamente 16 h. Os pré-inóculos resultantes foram então adicionados na proporção 1:20 a 10-20 mL de novo meio de cultura (LBON ou LB) acrescidos dos antibióticos adequados, seguido por incubação a 30 °C ou 37 °C sob agitação, até ser atingida uma DO_{600nm} de 0,6-0,8. Estoques em glicerol dos pré-inóculos também foram preparados, reunindo uma alíquota de cada pré-inóculo a um volume equivalente de solução de glicerol 50% esterilizada, e as amostras foram imediatamente armazenadas a -80 °C para conservação das células transformadas. Após ser atingida a DO_{600nm} ideal, a expressão das proteínas recombinantes foi induzida com a adição de NaCl ou IPTG, de acordo com a cepa hospedeira (Tabela 2), seguida por incubação a 30 °C ou 37 °C sob agitação por mais 3-4 h. As bactérias ArcticExpress (DE3) possuem chaperoninas com atividades máximas a baixas temperaturas (FERRER et al., 2003) e, portanto, os cultivos com esta cepa foram resfriados antes da indução com IPTG, sendo então incubados a 10-12 °C, 180 rpm, durante 18 h, 21 h ou 24 h. Amostras de 1 mL de todos os cultivos foram coletadas e centrifugadas (16.500 x g, 4 °C, 5 min) antes da adição do indutor e após as incubações finais, e o restante dos cultivos também foi coletado por centrifugação (3.500 x g, 4 °C, 20 min) para posterior avaliação do rendimento e solubilidade das proteínas recombinantes.

Os precipitados celulares resultantes das amostras de cultivo antes e após indução em cada cepa bacteriana foram homogeneizados diretamente em 100-150 μ L de tampão de amostra para análise em SDS-PAGE (composição do tampão detalhada no item 3.14.2). Em paralelo, os precipitados celulares de maior volume referentes ao restante dos cultivos induzidos foram homogeneizados em 600 μ L do tampão de lise (50 mM Tris-HCl, 50mM NaCl, 5 mM EDTA e 1 mg/mL de lisozima; pH maior ou menor do que uma unidade do pI predito para cada proteína recombinante [ver item 3.8]), encaminhados para sonicação (30-60 s com pulsos de 1,0 s, com intervalos de 1,0 s e 20% de amplitude) e, em seguida, as amostras lisadas foram centrifugadas (16100 x g por 5 min a 4 °C). Os sobrenadantes resultantes foram transferidos

para novos tubos, enquanto que os precipitados foram homogeneizados em 200 µL de solução contendo tampão 8 M de ureia (50 mM Tris-HCl, 500 mM NaCl e 8 M de ureia; pH ajustado como na solução acima), e posteriormente alíquotas foram preparadas com tampão de amostra para análise em SDS-PAGE. Os sobrenadantes foram ainda precipitados com a adição de acetona na proporção 1:2, seguida por incubação a -20 °C por pelo menos 1 h, novamente centrifugados (16100 x g por 10 min a 4 °C) e, após descarte dos novos sobrenadantes, os precipitados finais foram homogeneizados em 50-200 µL de tampão de amostra para análise em SDS-PAGE.

3.7.1. Produção de proteínas recombinantes em maior volume

Após a análise da produção das proteínas recombinantes nas diferentes cepas de *E. coli* testadas, os clones que obtiveram maior rendimento foram selecionados para amplificação em maiores volumes como detalhado a seguir, a partir dos estoques em glicerol previamente preparados. Os clones escolhidos possuíam construções a partir dos vetores pAE ou pAZ.

A maioria das proteínas recombinantes foram obtidas a partir da expressão induzida de suas sequências codificadoras em bactérias BL21-SI, incluindo LepA962_Nt e LepA962_Phg provenientes da lipoproteína hipotética LepA962, além das proteínas LepA388_LecB, LepA388_Phg, LepA388_D1, LepA388_Cards e LepA388_Ct provenientes da proteína hipotética LepA388. O volume total de meio de cultura LBON utilizado para as expressões variou de 250 mL (derivadas de LepA388) a 1 L (derivadas de LepA962), aos quais foram adicionados os pré-inóculos na proporção de 1:20 a 1:40. Após ser atingida uma DO_{600nm} de 0,6-0,8, foi adicionado NaCl equivalente a 200-300 mM aos cultivos para a indução da expressão das proteínas a 25-30 °C, com agitação de 200 rpm durante 4h. Ao final das incubações, os cultivos foram distribuídos em tubos estéreis e centrifugados (3.200 x g por 20

min a 4 °C) para a coleta dos precipitados celulares, os quais foram mantidos a -20 °C ou -80 °C até serem processados para a purificação das proteínas recombinantes.

De forma semelhante à descrita acima, algumas proteínas recombinantes tiveram a sua expressão induzida com IPTG em bactérias BL21 Star (DE3) pLysS ou ArcticExpress (DE3). As proteínas recombinantes LepA402F, LepA271F e LepA388F (FERREIRA, 2015) foram produzidas em 250 mL de cultivo de bactérias BL21 Star (DE3) pLysS induzidas com 1 mM de IPTG e incubadas a 28 °C, com agitação de 200 rpm durante 4h. Adicionalmente, a proteína recombinante LepA388F também foi obtida em 2 L de cultivo nas mesmas condições. Por outro lado, as proteínas recombinantes EGFP e a fusão EGFP-LepA388_Cards obtiveram melhor rendimento em bactérias ArcticExpress (DE3), sendo possível avaliar sua fluorescência precocemente nos precipitados celulares durante os testes iniciais de produção das proteínas recombinantes descritos no item 3.7. A indução de cada proteína recombinante foi realizada em 1 L de cultivo de bactérias ArcticExpress (DE3) induzidas com 0,05 mM de IPTG e incubadas a 12 °C, com agitação de 120 rpm durante 27h. Ao término das incubações dos cultivos tanto em BL21 Star (DE3) pLysS quanto em ArcticExpress (DE3), os precipitados celulares de cada proteína recombinante foram coletados e congelados como descrito no parágrafo anterior.

3.8. Isolamento das proteínas recombinantes produzidas em bactérias

O isolamento das proteínas recombinantes produzidas envolveu as etapas de extração ou lise celular, purificação e, em alguns casos, diálise para posteriores análises funcionais. Antes do processamento das células bacterianas, foi preparado um tampão base com pH maior ou menor do que uma unidade do pI predito para cada proteína, acrescido de alguns reagentes necessários para cada etapa e detalhados ao longo do texto (Tabela 3).

Tabela 3

Relação da massa molecular e do ponto isoelétrico (pI) preditos para cada proteína recombinante, segundo a ferramenta ProtParam (Expasy – SIB), e composição dos tampões base utilizados para seu isolamento a partir das células bacterianas.

Gene/Proteína de origem	Proteína recombinante	Parâmetros preditos (massa molecular e pI)	Tampões base
<u>LepA962</u>	LepA962_Nt	23,4 kDa; pI = 6,43	50 mM Tris + 500 mM NaCl + 5 mM imidazol, pH 7,8
	LepA962_Phg	15,2 kDa; pI = 5,46	50 mM Tris + 500 mM NaCl + 5 mM imidazol, pH 7,5
<u>LepA388</u>	LepA388_Blec	14,8 kDa; pI = 7,75	20 mM Fosfato de sódio + 500 mM NaCl, pH 6,5
	LepA388_Phg	10,7 kDa; pI = 6,41	20 mM Fosfato de sódio + 500 mM NaCl, pH 8,0
	LepA388_D1	31,5 kDa; pI = 6,54	20 mM Fosfato de sódio + 500 mM NaCl, pH 8,0
	LepA388_Cards	35,4 kDa; pI = 7,64	20 mM Fosfato de sódio + 500 mM NaCl, pH 6,5
	LepA388_Ct	23,7 kDa; pI = 9,05	20 mM Fosfato de sódio + 500 mM NaCl, pH 6,5
	EGFP-LepA388_Cards (vetor pAZ)	62,5 kDa; pI = 6,26	20 mM Fosfato de sódio + 500 mM NaCl + 5 mM imidazol, pH 6,5
	LepA388F	69,7 kDa; pI = 8,44	20 mM Fosfato de sódio + 500 mM NaCl, pH 6,5
<u>LepA271</u>	LepA271F	70,3 kDa; pI = 6,96	20 mM Fosfato de sódio + 500 mM NaCl, pH 8,0
<u>LepA402</u>	LepA402F	70,2 kDa; pI = 8,84	20 mM Fosfato de sódio + 500 mM NaCl, pH 6,5

3.8.1. Extração das proteínas recombinantes das células bacterianas (lise celular)

Inicialmente, a extração das proteínas recombinantes foi realizada pela homogeneização dos precipitados celulares correspondentes a cada cultivo bacteriano com tampão de lise composto pelo tampão base previamente resfriado e acrescido do inibidor de serino protease PMSF (concentração final de 1 mM); em alguns casos também foi adicionado um coquetel inibidor de protease comercial na proporção 1:1.000 (Sigma-Aldrich). Após a homogeneização

completa, as células bacterianas foram submetidas a 4-5 ciclos de lise mecânica sob pressão de 1000-1300 bar (*Lab Homogenizer PandaPLUS 2000*, GEA), sendo incubadas em gelo ao final de cada ciclo para evitar o aquecimento excessivo das amostras. As células lisadas foram em seguida centrifugadas a 7.800 x g por 25-35 min a 4 °C, e os sobrenadantes e precipitados resultantes foram processados de acordo com a localização subcelular das proteínas recombinantes.

Os sobrenadantes contendo as proteínas recombinantes produzidas na região citoplasmática das bactérias foram imediatamente clarificados em filtro 0,45 µm ou 0,22 µm e encaminhados para a purificação, a fim de evitar a degradação das proteínas. Já os precipitados contendo as proteínas recombinantes em corpúsculos de inclusão foram submetidos a uma série de lavagens, a fim de isolar os corpúsculos de outros contaminantes celulares e solubilizar as proteínas para posterior purificação e renaturação (SINGH et al., 2015). Para o isolamento dos corpúsculos, os mesmos foram homogeneizados em tampão base contendo 2% de Triton X-100, seguida por centrifugação (7.800 x g por 20 min a 4 °C). Os precipitados contendo os corpúsculos isolados foram então homogeneizados, incubados (de 1 h a ~16 h) e novamente centrifugados em uma sequência de tampões base contendo concentrações crescentes de ureia (tampões de solubilização com 2 M, 4 M ou 8 M ureia). A cada etapa foram coletadas amostras para avaliar a solubilização das proteínas recombinantes de interesse por SDS-PAGE; em alguns casos, os precipitados finais foram ainda homogeneizados em tampão base contendo 6 M de guanidina-HCl e/ou 1 mM de β-mercaptoetanol, a fim de garantir a completa solubilização das proteínas recombinantes.

Após a avaliação das amostras onde se encontravam a maior porção solubilizada de cada proteína recombinante, as mesmas foram clarificadas em filtro 0,45 µm ou 0,22 µm e encaminhadas para a purificação.

3.8.2. Purificação das proteínas recombinantes (IMAC e outras técnicas)

A etapa de purificação foi realizada em sistemas de cromatografia ÄKTA (ÄKTAprime plus ou ÄKTA Pure, GE Healthcare/Cytiva) e envolveu principalmente a técnica de cromatografia de afinidade do tipo IMAC com colunas pré-empacotadas HisTrap HP (GE Healthcare/Cytiva), compostas pela resina *Ni Sepharose High Performance* (resina em pérolas pequenas de 34 µm), já associadas a íons de níquel. Os tampões de ligação utilizados para as purificações possuíam a mesma composição das soluções onde as proteínas recombinantes se encontravam solubilizadas, enquanto que os tampões de eluição eram acrescidos de 500 mM de imidazol; adicionalmente, em algumas purificações foi utilizado um tampão de lavagem contendo 40-50 mM imidazol, para ser usado antes da eluição das proteínas de interesse. A coleta das proteínas recombinantes purificadas foi baseada nos picos de absorbância de luz UV a 280 nm expressa em mAu (miliunidades arbitrárias), durante a análise em tempo real dos cromatogramas; a eluição da maioria das proteínas recombinantes ocorreu na faixa de concentração entre 100 mM e 200 mM de imidazol.

Após ser purificada por IMAC, a proteína recombinante LepA962_Nt foi submetida a duas técnicas de purificação adicionais: troca aniônica e interação hidrofóbica. Antes da troca aniônica, a amostra em tampão inicial composto por: 50 mM Tris, 0,5 M NaCl, 4 M ureia e aproximadamente 250 mM imidazol (pH 8,0) foi dialisada em tampão 20 mM Tris e 4 M ureia (pH 8,0) para a retirada do NaCl (concentração final de aproximadamente 0,3 mM), sendo posteriormente aplicada a uma coluna preparada manualmente com resina *Q Sepharose Fast Flow* (GE Healthcare/Cytiva) no sistema de cromatografia ÄKTAprime plus (GE Healthcare/Cytiva). A concentração crescente e em gradiente de NaCl resultou na eluição da proteína a partir de 50 mM de NaCl, atingindo o pico de absorbância da luz UV a 280 nm com 100 mM de NaCl. Em seguida, as amostras purificadas foram analisadas por SDS-PAGE, e optou-se por realizar a purificação por interação hidrofóbica, desta vez utilizando o método *in*

batch. Uma amostra proveniente da purificação por troca aniônica com aproximadamente 200 mM de NaCl foi adicionada a 8 tubos contendo a resina *Phenyl Sepharose 6 Fast Flow High Sub* (GE Healthcare/Cytiva) e concentrações variadas de NaCl (2,5 M - 2 M - 1,5 M - 1 M - 0,75 M - 0,5 M - 0,25 M - 0 M). Após homogeneização e incubação sob agitação durante 1 h, as amostras foram centrifugadas (16.500 x g, 1 min) e os sobrenadantes resultantes foram analisados por SDS-PAGE.

3.8.3. Diálise das proteínas recombinantes de LepA962 purificadas

As proteínas recombinantes LepA962_Phg e LepA962_Nt purificadas em tampão contendo 4 M ureia (50 mM Tris-HCl, 500 mM NaCl, 4 M ureia) foram submetidas a renaturação através de diálise gradual, a fim de diminuir gradualmente a concentração de ureia nas amostras e evitar possíveis precipitações. A concentração inicial de cada proteína foi determinada por ensaio colorimétrico baseado no método Bradford (BRADFORD, 1976), utilizando o reagente *Protein Assay Dye Reagent Concentrate* (Bio-Rad) e seguindo o protocolo do fabricante. Em seguida, diluições de aproximadamente 400 µg/mL e 200 µg/mL foram preparadas em 5 mL para cada proteína, as quais foram centrifugadas (7.800 x g por 15 min a 4 °C) e os sobrenadantes resultantes foram posteriormente acondicionados em sacos de diálise *SnakeSkin™ Dialysis Tubing* (Thermo Scientific) com corte molecular de 3,5 kDa (para LepA962_Phg) ou 10 kDa (para LepA962_Nt).

A primeira diálise (4 M ureia → 3 M ureia) foi realizada em 900 mL de tampão de diálise (50 mM Tris-HCl, 500 mM NaCl e 50 mM Glicina, pH 8,0) contendo 3 M de ureia, sob agitação a 4 °C. Após aproximadamente 15 h de incubação, o tampão para a segunda diálise (3 M ureia → 2 M ureia) foi preparado a partir do tampão resultante da primeira diálise diluído em novo tampão de diálise sem ureia, e as amostras foram incubadas a 4 °C sob agitação durante aproximadamente 8 h. O mesmo procedimento foi realizado para a terceira e quarta diálises (2

M ureia → 1 M ureia; 1 M ureia → 0,5 M ureia), enquanto que as duas últimas diálises (0,5 M ureia → 12,2 mM ureia; 12,2 mM ureia → 0,3 mM ureia) foram realizadas diretamente em 800 mL de tampão de diálise contendo 10% de glicerol cada. As incubações em cada diálise variaram entre 8 h a 24 h de duração e, a partir do final da terceira diálise (2 M ureia → 1 M ureia), as amostras foram coletadas e centrifugadas para a retirada de possíveis precipitados, e cada sobrenadante resultante foi novamente acondicionado em novos sacos de diálise com os mesmos cortes moleculares descritos acima para as diálises seguintes. Ao final da sexta e última diálise, as amostras foram coletadas e novamente centrifugadas para separação de possíveis precipitados. As amostras de precipitado após cada centrifugação foram homogeneizadas em tampão contendo 4 M ureia, enquanto que os sobrenadantes finais foram filtrados a 0,22 µm; todas as amostras foram quantificadas por ensaio colorimétrico baseado no método Bradford e também analisadas por SDS-PAGE.

Devido ao baixo rendimento final da proteína recombinante LepA962_Ph_g e posterior degradação durante seu armazenamento a 4 °C, duas novas diluições foram preparadas com aproximadamente 300 µg/mL (5 mL) e 500 µg/mL (2 mL) a partir de amostras purificadas em 4 M de ureia. Antes da primeira diálise e após cada diálise, as amostras foram centrifugadas e acondicionadas em sacos de diálise *SnakeSkin™ Dialysis Tubing* (Thermo Scientific) com corte molecular de 3,5 kDa. A composição e preparação dos tampões de diálise foram realizadas de forma semelhante à descrita anteriormente. As quatro primeiras diálises (4 M → 3 M ureia; 3 M → 2 M ureia; 2 M → 1 M ureia; 1 M → 0,5 M ureia) foram realizadas em 800 mL de tampão de diálise, sendo reaproveitada parte do tampão da diálise anterior. A partir da quinta diálise (0,5 M → 0,1 M ureia), os tampões passaram a conter 10% de glicerol em 650 mL cada e, ao final das sete diálises com incubações entre 4 h e 18 h de duração, as amostras apresentavam concentração de ureia a 0,013 mM. Como descrito acima, os sobrenadantes finais foram filtrados a 0,22 µm e todas as amostras de precipitado após cada centrifugação foram

homogeneizadas em tampão contendo 4 M ureia para posterior quantificação por ensaio colorimétrico baseado no método Bradford e análise por SDS-PAGE. As amostras filtradas foram então armazenadas a -20 °C para evitar a degradação da proteína.

3.9. Análise de amostras da proteína recombinante LepA962_Nt por espectrometria de massas

Análises da proteína LepA962_Nt por SDS-PAGE após as purificações por IMAC, troca aniônica e interação hidrofóbica revelaram a presença de duas bandas de tamanho muito próximo: uma menor e mais intensa com aproximadamente 23,4 kDa, tamanho mais próximo ao tamanho predito para a proteína, e uma maior e menos intensa, com pouco mais de 30 kDa. A fim de analisá-las separadamente por espectrometria de massas, cada banda foi excisada do SDS-PAGE já corado com Coomassie Blue e enviada ao Laboratório de Toxinologia Aplicada, localizado no Centro de Toxinas, Resposta Imune e Sinalização Celular (CeTICS) do Instituto Butantan, sob os cuidados da Dra. Solange Maria de Toledo Serrano. Após tratamento com tripsina, as amostras foram submetidas à ionização por *electrospray* (ESI) e analisadas por espectrometria de massas em tandem, utilizando os analisadores de massas por quadrupolo e por tempo de voo (QUAD-TOF). Os íons resultantes foram então comparados ao banco de dados de *Leptospira interrogans* para identificação de alvos significativos e determinação da proteína de origem de cada amostra.

3.10. Avaliação da estrutura secundária de proteínas recombinantes por dicroísmo circular

A técnica de dicroísmo circular foi utilizada para avaliar as proteínas recombinantes LepA962_Phg e LepA962_Nt e estimar a proporção de suas estruturas secundárias (GREENFIELD, 2006). Amostras das proteínas purificadas e sem ureia foram novamente

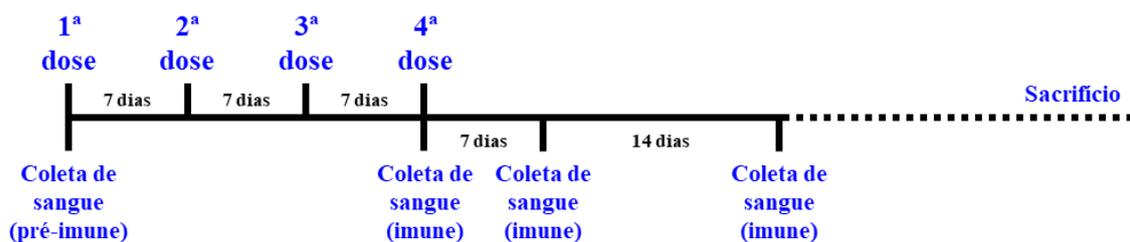
submetidas à diálise em tampão 10 mM de fosfato de potássio realizada em dois ciclos de incubação sob agitação a 4 °C, para a remoção do excesso dos sais de sódio. Em seguida, as amostras dialisadas foram adicionadas a cubetas com caminho óptico de 0,1 cm e analisadas a 25 °C no espectropolarímetro Jasco J-810. Os espectros resultantes foram gerados a partir de dez leituras da polarização de luz medida no intervalo de 188 a 260 nm, com incrementos de 0,1 nm. Posteriormente, os dados em milligrados coletados para cada proteína foram convertidos em valores de elipticidade molar (ERRINGTON et al, 2006), considerando o tamanho das proteínas em Da, a concentração em mg/mL o número de aminoácidos e o caminho óptico da cubeta utilizada. Adicionalmente, as curvas obtidas foram interpretadas usando o algoritmo CDSSTR disponível no servidor Dichroweb (SREERAMA; WOODY, 2000; WHITMORE; WALLACE, 2004, 2008), a fim de revelar a proporção das principais estruturas secundárias em cada proteína analisada.

3.11. Produção de antissoros em camundongos a partir de proteínas recombinantes

A fim de obter os antissoros das proteínas recombinantes LepA962_Phg e LepA962_Nt, seis camundongos BALB/c fêmeas com 5 a 7 semanas de idade foram imunizados pela via intraperitoneal com uma solução contendo 10 µg de uma das proteínas recombinantes diluída em salina estéril e na presença de hidróxido de alumínio como adjuvante (100 µg de Al³⁺/dose). O esquema de imunização incluiu mais três doses das mesmas preparações, as quais foram administradas em intervalos de uma semana (Figura 3). Amostras de sangue foram coletadas via plexo retro orbital nos animais previamente anestesiados com uma solução anestésica de uso oftálmico (cloridrato de proximetacaína); as sangrias foram realizadas antes da primeira dose (soros pré-ímmunes) e após completar uma semana da terceira e da quarta doses (Figura 3). Após cada coleta, as amostras de sangue foram incubadas a 37 °C durante 20-30 min, seguida por incubação a 4 °C por 15-60 min, a fim de permitir a formação e a concentração de coágulo

em cada amostra, respectivamente. Posteriormente, as amostras foram centrifugadas a 360 x g, 4 °C, 10 min. Os sobrenadantes resultantes provenientes dos mesmos grupos de imunização foram reunidos em um único tubo (“pool” dos soros individuais) para nova centrifugação, e os sobrenadantes finais foram separados em alíquotas de 50 µL e armazenados a -20 °C. Os antissoros foram posteriormente titulados por ensaio imunoenzimático do tipo ELISA (ver item 3.12.1), além da análise por *Western Blotting*.

A) Imunização com LepA962 Nt



B) Imunização com LepA962 Phg

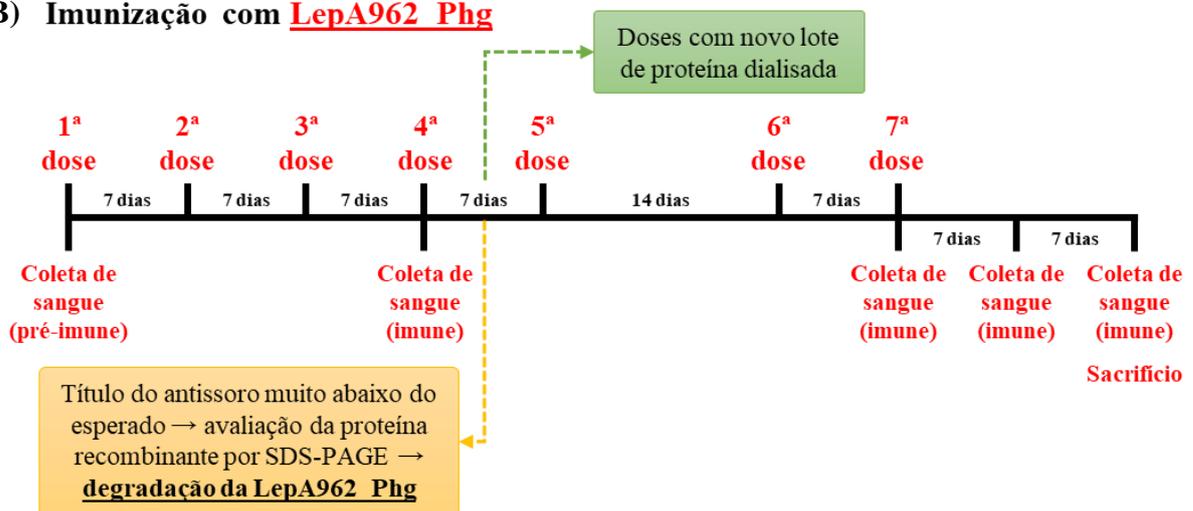


Figura 3. Esquema das imunizações de camundongos com as proteínas recombinantes LepA962_Nt (A) e LepA962_Phg (B) para obtenção dos respectivos antissoros. Doses administradas em intervalos de 7 ou 14 dias, com coletas de sangue antes das primeiras doses (soros pré-ímmunes) e após 7, 14 e 21 dias das últimas doses (soros ímmunes). Em B, a observação de degradação da proteína LepA962_Phg ocorreu após o primeiro antissoro ímmune (após 3ª dose) apresentar título muito abaixo do esperado, seguida pela avaliação da proteína recombinante por SDS-PAGE. A partir disso, imunizações adicionais foram realizadas com novo lote da mesma proteína dialisada. Todos os animais foram sacrificados no final dos esquemas de imunização, de acordo com o protocolo CEUAIB 3834060616.

A análise dos antissoros e das proteínas recombinantes ao longo do esquema de imunização inicialmente planejado revelou um baixo título dos soros contra a proteína LepA962_Phg, a qual sofreu degradação durante seu armazenamento a 4 °C, observada por avaliação da proteína por SDS-PAGE. Por esta razão, o mesmo grupo de camundongos seguiu no esquema de imunização, sendo administradas mais três doses de novas preparações (preservando a mesma composição das primeiras doses, exceto pelo novo lote de proteína recombinante), durante as quatro semanas seguintes (Figura 3B). Coletas de sangue foram realizadas após uma ou duas semanas das últimas doses para a obtenção dos antissoros, como descrito acima.

O uso dos animais nos experimentos descritos está detalhado no protocolo de número 3834060616, aprovado pelo Comitê de Ética no Uso de Animais do Instituto Butantan (CEUAIB) (ANEXO A), respeitando os Princípios Éticos para Experimentação Animal adotados pelo Colégio Brasileiro de Experimentação Animal.

3.12. Ensaio imunoenzimático e de ligação *in vitro* baseados em ELISA

3.12.1. Titulação de antissoros

Os antissoros das proteínas recombinantes LepA962_Phg e LepA962_Nt foram titulados com base em protocolos clássicos de ELISA indireto (HORNBECK et al, 1991; CROWTHER, 1995).

Placas de 96 poços de alta ligação foram previamente sensibilizadas com 0,1 µg/poço (1 µg/mL) de cada proteína recombinante diluída em tampão carbonato-bicarbonato pH 7,6 (50 mM de Na₂CO₃ e 50 mM de NaHCO₃) e incubadas a 4 °C por 16-20 h. Posteriormente, as placas foram lavadas três vezes com 200 µL/poço de solução PBS acrescida de 0,1% de Tween 20 (PBS-T) e, em seguida, os poços foram bloqueados com a adição de 100 µL/poço de solução de PBS-T contendo 10% de leite em pó desnatado, seguido por incubação a 37 °C por 1 h. As

placas foram novamente lavadas e os antissoros foram adicionados e diluídos seriadamente em 100 µL/poço com solução PBS-T contendo 1% de BSA, considerando as sequências de diluição 1:20 a 1:20.480 e/ou 1:1.000 a 1:1.024.000 para cada amostra de soro. Após incubação a 37 °C por 1h30-2h, as placas foram novamente lavadas e incubadas com 100 µL/poço de solução contendo anticorpo anti-IgG de camundongo conjugado à peroxidase diluído 1:10.000 em PBS-T contendo 1% de BSA, durante 1 h a 37 °C. Finalmente, as placas foram lavadas e reveladas com 100 µL/poço de tampão citrato-fosfato pH 5,0 (100 mM de citrato de sódio e 300 mM fosfato de sódio monobásico) acrescido de 0,5 mg/mL de OPD e 0,015% de peróxido de hidrogênio, sob proteção da luz. As reações foram interrompidas após 6 a 22 min com 50 µL/poço de solução 4 N de H₂SO₄ e as placas foram posteriormente analisadas por leitura de absorbância a 492 nm em espectrofotômetro.

Os títulos dos antissoros foram calculados a partir dos valores de absorbância, subtraindo os valores obtidos nos poços de controle (sem adição de soro); as diluições com os valores resultantes mais próximos a 0,1 foram definidos como o título de cada antissoros.

3.12.2. Ligação de proteínas recombinantes a componentes de matriz extracelular

A interação das proteínas recombinantes a macromoléculas presentes na matriz extracelular (MEC) foi avaliada de acordo com protocolos baseados em ELISA previamente publicados (BARBOSA et al., 2006; ATZINGEN et al., 2008). Placas de 96 poços de alta ligação foram sensibilizadas com 1 µg/poço (100 µL/poço) das seguintes proteínas de MEC e outras macromoléculas diluídas em PBS: colágeno tipo I de cauda de rato (Corning), laminina e colágeno tipo IV de membrana basal de sarcoma murino Engelbreth-Holm-Swarm, fibronectina celular de fibroblastos humanos, fibronectina e fibrinogênio de plasma humano, decorina de cartilagem articular bovina e BSA como controle negativo (Sigma-Aldrich). Após incubação das placas a 4 °C por aproximadamente 16 h, as placas foram lavadas três vezes com

200 μL /poço de PBS-T (PBS acrescido de 0,05% de Tween 20) e, em seguida, foram bloqueadas com PBS com 3% de gelatina durante 2 h a 37 °C. As placas foram novamente lavadas e 1 μg de cada proteína recombinante (LepA962_Phg e LepA962_Nt) foi adicionada em 100 μL de PBS em triplicata. Utilizando as mesmas concentrações, uma proteína recombinante do domínio C-terminal da proteína LigB (LigBC) e BSA também foram adicionadas em triplicata como controles positivo e negativo, respectivamente. As placas foram incubadas com as proteínas por 2 h a 37 °C, seguida por lavagem e adição de antissoros específicos a cada proteína recombinante: antissoros produzidos em camundongos (a partir de LepA962_Phg ou LepA962_Nt) e hamsters (a partir de LigBC), na proporção 1:5.000 em PBS (100 μL /poço), e incubação a 37 °C por 1 h. Após nova lavagem, foram adicionados 100 μL /poço de anticorpo secundário anti-IgG de camundongo ou anti-IgG de hamster, ambos conjugados à peroxidase e também diluídos 1:5.000 em PBS, seguida pela incubação a 37 °C por 50 min. A etapa de revelação foi conduzida após lavagem das placas para remoção do excesso dos anticorpos secundários, sendo posteriormente adicionados 100 μL /poço da mesma solução de revelação descrita no item 3.12.1 (tampão citrato-fosfato acrescido de OPD e peróxido de hidrogênio). Após incubação no escuro por 10 min à temperatura ambiente, as reações foram interrompidas pela adição de 50 μL /poço de 4 N H_2SO_4 e a absorbância foi imediatamente analisada a 492 nm em espectrofotômetro.

Para a determinação das curvas de ligação dependente de dose entre as proteínas recombinantes LepA962_Phg e LepA962_Nt e as proteínas da MEC laminina, fibronectina plasmática, fibrinogênio e colágeno tipo I, três experimentos independentes foram realizados seguindo o mesmo protocolo acima, utilizando concentrações de proteína recombinante variando de 0 a 1.000 nM em PBS. Os cálculos dos coeficientes de dissociação (K_D) foram determinados com auxílio do software GraphPad Prism.

3.12.3. Ligação de proteínas recombinantes a células VERO e à MEC secretada

Além das proteínas da MEC, a interação das proteínas recombinantes LepA962_Phg e LepA962_Nt também foi avaliada em cultura de células eucarióticas, assim como descrito na literatura com algumas modificações (EVANGELISTA et al., 2014a).

Células epiteliais de rim VERO foram descongeladas e estabelecidas em cultura utilizando frascos de 75 cm². A partir de uma cultura confluenta, as células foram tripsinizadas, homogeneizadas e contadas para serem diluídas a 5×10^5 células/mL em meio de cultura DMEM suplementado com 10% de soro fetal bovino (SFB) e antibióticos penicilina e estreptomicina (100 unidades e 100 µg/mL, respectivamente - Sigma). Em uma microplaca tratada para cultura de células (Corning), foram adicionados 100 µL/poço das células diluídas (5×10^4 células/poço), a qual foi posteriormente incubada a 37 °C, sob atmosfera de 5% de CO₂. Alguns dos poços na microplaca foram deixados vazios.

A fim de avaliar a ligação de proteínas recombinantes à ECM secretada por células VERO, após 24 h de incubação, metade dos poços contendo as células foram lavados 1x com PBS estéril (200 µL/poço), sendo posteriormente adicionados 75 µL/poço de solução 10 mM EDTA estéril para incubação por 10 min à temperatura ambiente em um agitador de placas. Em seguida, as células foram aspiradas e os poços foram observados ao microscópio invertido para avaliar a remoção completa das células; os mesmos poços foram adicionalmente lavados 3x com 200 µL/poço de tampão HBSC (25 mM de HEPES, pH 7,4; 150 mM de NaCl; 1 mM de MnCl₂; 1 mM de MgCl₂; 250 mM de CaCl₂). A outra metade dos poços contendo as células intactas e confluentes foi também lavada 1x com 200 µL/poço de PBS estéril. Aos poços com células, com a MEC secretada e vazios foram adicionadas duas concentrações (0,25 µM e 1 µM) das proteínas recombinantes LepA962_Phg e LepA962_Nt em 100 µL/poço de meio DMEM suplementado com 1% de BSA, 25 mM de HEPES e inibidores de protease (1 mM de benzamidina e 0,01 unidade de inibição de tripsina (TIU)/mL de inibidor de tripsina de soja),

seguida por incubação durante 1 h e 15 min a 37 °C, sob atmosfera de 5% de CO₂. Também foram adicionadas as mesmas concentrações de uma porção recombinante da proteína A de superfície pneumocócica (PspA4Pro), como um controle negativo de proteína marcada com 6x-His e produzida em bactérias BL21-SI (MORENO et al., 2010). Após a incubação, as proteínas não ligadas foram lavadas com 200 µL/poço de tampão HBSC e os poços foram fixados com solução de paraformaldeído a 3% em PBS durante 30 min à temperatura ambiente. Após uma etapa de lavagem com PBS (200 µL/poço), os poços foram bloqueados em solução de PBS com 3% de BSA (200 µL/poço) durante 1 h a 37° C e o ensaio prosseguiu em um formato de ELISA indireto como descrito no item 3.12.1, utilizando as mesmas concentrações dos antissoros específicos das proteínas recombinantes utilizadas (incluindo o anticorpo anti-PspA4Pro) e do anticorpo secundário anti-IgG de camundongo conjugado à peroxidase. Na etapa de revelação, a placa foi incubada durante 30 min sob proteção de luz, antes da parada da reação e leitura de absorbância a 492 nm em espectrofotômetro.

3.12.4. Ligação dos antissoros das proteínas recombinantes na superfície de leptospiros

A ligação dos antissoros das proteínas recombinantes LepA962_Phg, LepA962_Nt e LepA388_Phg à superfície de leptospiros foi avaliada por ensaio de ELISA sanduíche. Microplacas de alta ligação foram revestidas com 100 µL/poço de um dos anticorpos de captura diluídos 1:2.000 em tampão carbonato-bicarbonato pH 9,6 (ver item 3.12.1): soro produzido em coelho contra extrato proteico de *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130, ou soro de coelho virgem (ou *naïve*). As placas foram incubadas a 4 °C por aproximadamente 16 h, seguida por três lavagens com 300 µL/poço de PBS-T (0,1% Tween 20) e etapa de bloqueio com 200 µL/poço de PBS-T acrescido de 5% de BSA durante 1 h e 15 min a 37 °C. Durante a incubação, células virulentas de *L. interrogans* sorovar Copenhageni cepa Fiocruz

L1-130 cultivadas em EMJH (ver item 3.1) foram coletadas por centrifugação (7.000 x g, 15 min, 25 °C), lavadas com PBS estéril e homogeneizadas em PBS a uma concentração de 10^8 células/mL. Ao término da incubação, as placas foram novamente lavadas e, em seguida, foram adicionados 100 µL/poço das leptospiras em suspensão (10^7 células/poço) para nova incubação a 37 °C por 1 h e 40 min. A remoção das bactérias não ligadas ocorreu com três lavagens das placas manualmente com 300 µL/poço de PBS-T, seguida pela adição de 100 µL/poço dos antissoros de camundongo referentes às proteínas recombinantes (antissoro da proteína LepA388_Phg obtido durante o período do Mestrado; FERREIRA, 2015), os quais foram diluídos 1:200 em PBS. Soros pré-ímmunes dos mesmos animais usados para produzir os antissoros foram reunidos e adicionados na mesma diluição (1:200), enquanto um antissoro contra a proteína LipL32 recombinante foi aplicado como controle positivo (diluído 1:1.000). Após 1 h de incubação a 37 °C, as placas foram novamente lavadas e incubadas com 100 µL/poço do anticorpo secundário anti-camundongo conjugado à peroxidase diluído 1:5.000 em PBS durante 1 h a 37 °C. A etapa de revelação colorimétrica com OPD consistiu na incubação das placas por aproximadamente 40 min sob proteção de luz, antes da parada da reação e leitura de absorbância a 492 nm em espectrofotômetro.

3.13. Detecção da interação de proteínas com componentes da MEC por *Far Western Blotting*

A interação das proteínas recombinantes referentes a porções da proteína LepA388 e outras proteínas da família DUF_61 com componentes da MEC foi avaliada pela técnica de *Far Western Blotting* (WU; LI; CHEN, 2007; HALL, 2015). Esta metodologia é bastante similar ao tradicional *Western Blotting*, exceto pela incubação adicional por proteínas de interesse após a etapa de bloqueio, a fim de avaliar interações proteína-proteína.

Primeiramente, as proteínas recombinantes *LepA402_Full*, *LepA271_Full*, *LepA388_Full*, *LepA388_LecB*, *LepA388_Phg*, *LepA388_D1*, *LepA388_Cards* e *LepA388_Ct* purificadas por IMAC em 8 M ureia foram diluídas em tampão de amostra para SDS-PAGE e aplicadas em géis preparados com duas concentrações de poliacrilamida (9% e 16%) para separação sob voltagem de 80-130 V. As proteínas recombinantes LigA e LIC10301 também foram aplicadas aos géis como controles positivo e negativo, respectivamente.

Ao término das migrações eletroforéticas, as proteínas foram transferidas para membranas de nitrocelulose (*Hybond ECL*, GE Healthcare Life Sciences) a 10 V durante 1 h e 15 min, utilizando um sistema de transferência horizontal semi-seca (*Trans-Blot® SD Semi-Dry Transfer Cell*, Bio-Rad). A avaliação da transferência das proteínas foi realizada após incubação das membranas com o corante Ponceau S, seguida por lavagem com água destilada e registro em fotodocumentador (Carestream). Posteriormente, o corante foi removido com 3 ciclos de lavagem com PBS-T (0,05% de Tween 20), sob agitação durante pelo menos 5 min cada, e as membranas foram incubadas com a solução de bloqueio (PBS-T com 5% de BSA) a 4 °C durante aproximadamente 16 h. Antes do descarte do bloqueio, as membranas foram incubadas a T.A. e sob agitação durante 20-30 min, seguida por duas lavagens com PBS-T e pela adição de 5 mL de solução contendo um dos seguintes componentes da MEC ou plasma, diluídos a 10 µg/mL em PBS-T: colágeno tipo I de cauda de rato (Corning), laminina e colágeno tipo IV de membrana basal de sarcoma murino Engelbreth-Holm-Swarm, fibronectina celular de fibroblastos humanos, fibronectina de plasma humano, fragmentos proteolíticos de 45 kDa e 30 kDa da fibronectina de plasma humano e vitronectina (Sigma). Após incubação sob agitação a T.A. durante 1 h e 30 min, as membranas foram lavadas 4 vezes com PBS-T e incubadas com os anticorpos primários contra as proteínas de MEC ou plasma (Sigma), diluídos a 1:2.500 (ou 1:3.000, no caso do anticorpo anti-laminina) em PBS-T, durante 1 h e 15 min a T.A. e sob agitação. As membranas foram novamente lavadas 3 vezes com PBS-T, seguida pela

incubação do anticorpo secundário anti-IgG de coelho conjugado à peroxidase (Sigma) diluído 1:7.000 em PBS-T, durante 1 h a T.A. e sob agitação. Finalmente, as membranas foram novamente lavadas 3 vezes com PBS-T e tratadas com volumes equivalentes das soluções A e B do kit de detecção (*Amersham® ECL Prime WB Detection Reagent*, GE Healthcare Life Sciences), com posterior incubação por 5 min. Em seguida, a luminescência resultante nas membranas foi detectada e documentada através do fotodocumentador Kodak Gel Logic 2200 Imaging System (Carestream).

3.14. Migrações eletroforéticas

3.14.1. Eletroforese em gel de agarose

A metodologia de eletroforese em gel foi principalmente utilizada para avaliar amostras após as etapas de clonagem e/ou subclonagem, como digestões com enzimas de restrição, amplificação por PCR e ligação de fragmentos aos vetores plasmidiais. As amostras foram diluídas em tampão para aplicação em gel de agarose (0,02% de azul de bromofenol, 0,042% de xileno cianol e 5% de glicerol), ou em tampão de aplicação em gel *BlueJuice®* (Invitrogen) diluído para 1X e adicionadas a géis com 1% de agarose em tampão TAE (40 mM de Tris base, 1 mM de EDTA e 20 mM de ácido acético glacial), ou tampão TBE comercial (Bio-Rad). Na maioria dos géis, foi acrescentado o brometo de etídio (agente intercalante de DNA) durante sua preparação para a detecção das moléculas de DNA; em alguns casos, foi utilizado o reagente *SYBR™ Safe DNA Gel Stain* (Invitrogen). As migrações eletroforéticas foram realizadas sob voltagem de 80-100 V e, posteriormente, as bandas de DNA resultantes foram visualizadas através de transiluminação com luz ultravioleta, seguida pelo registro em fotodocumentador. O marcador de tamanho molecular utilizado para as análises foi o *1 Kb plus DNA Ladder* (Invitrogen).

3.14.2. Eletroforese em gel de poliacrilamida (SDS-PAGE)

Amostras proteicas em geral para análise em SDS-PAGE foram preparadas em tampão de amostra diluído 1X em salina ou água ultrapurificada, com a seguinte composição: 50 mM de Tris-HCl pH 6,8, 2% de SDS, 0,1% de azul de bromofenol, 10% de glicerol e 200 mM de β -mercaptoetanol; em alguns casos, foi utilizado o tampão comercial *4x Laemmli Sample Buffer* (Bio-Rad), acrescido de 355 mM de β -mercaptoetanol. Antes da aplicação nos géis, as amostras foram incubadas a 95 °C por 5-10 min. Os géis foram principalmente preparados com 15% de poliacrilamida (SAMBROOK; RUSSELL, 2001), exceto nos géis para *Far Western Blotting* (ver item 3.13), em que foram usadas as concentrações de 9% e 16% de poliacrilamida, além da utilização de géis pré-fabricados comercialmente com 12% de poliacrilamida (*Mini-PROTEAN TGX Precast Gels*, Bio-Rad). As migrações eletroforéticas foram realizadas sob voltagem de 80-200 V em tampão Tris-glicina (25 mM de Tris base, 192 mM de glicina e 0,1% de SDS) e, em seguida, os géis foram incubados em corante azul de Coomassie (3 mM de *Coomassie Brilliant Blue R-250*, 10% de ácido acético glacial e 45% de etanol absoluto), ou transferidos para membranas de nitrocelulose para a realização de ensaios de *Western Blotting*. Para a descoloração de géis em Coomassie, foram realizadas lavagens com solução descorante (30% de etanol absoluto e 10% de ácido acético glacial), a fim de analisar as bandas resultantes das migrações. O principal marcador de tamanho molecular utilizado foi o *Amersham Low Molecular Weight Calibration Kit for SDS Electrophoresis* (GE Healthcare Life Sciences), diluído em tampão de amostra 1X; em alguns casos, foi utilizado o marcador *Amersham ECL Full-Range Rainbow Molecular Weight Marker* (GE Healthcare Life Sciences), diluído em tampão *Laemmli* 1X.

4. RESULTADOS

4.1. Estratégias exploradas para produção e purificação de proteínas recombinantes

Previamente, foi realizada a identificação de três proteínas hipotéticas de *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130 (LepA388, LepA962 e LepA069) pela técnica de *phage display*, selecionadas após enriquecimento de bacteriófagos recombinantes ligados a células VERO em suspensão pela técnica BRASIL (Figura 4). Ainda durante o período do Mestrado, a caracterização funcional da proteína LepA388 foi iniciada, entretanto, a expressão dos fragmentos recombinantes resultou em seu acúmulo em corpúsculos de inclusão, além do grande número de cisteínas ao longo de suas sequências dificultarem sua correta renaturação (FERREIRA, 2015). Por esta razão, novas proteínas recombinantes foram desenhadas, assim como algumas estratégias foram testadas para produção dessas proteínas de forma solúvel, ou para a renaturação a partir da sua expressão em corpúsculos de inclusão.

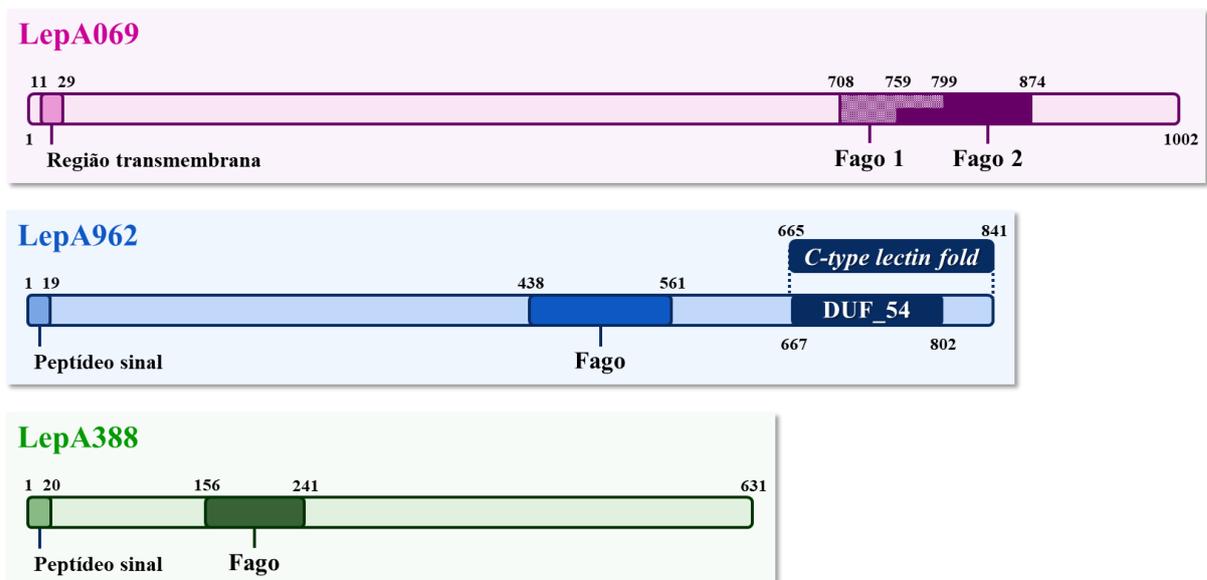


Figura 4. Esquema das três proteínas hipotéticas LepA069, LepA962 e LepA388 identificadas previamente pelo ensaio de *phage display*, seguido por enriquecimento dos fagos e seleção em células VERO (FERREIRA, 2015). Os trechos referentes à região transmembrana ou peptídeo sinal estão indicados, assim como domínios já anotados e a localização das sequências identificadas nos fagos.

As cepas bacterianas BL21 (DE3), BL21 Star (DE3) pLysS, BL21-SI, BL21 C43 (DE3), Origami (DE3) e ArcticExpress (DE3) foram escolhidas para os testes iniciais de produção das proteínas recombinantes. Bactérias BL21 (DE3) são amplamente utilizadas para a expressão de genes heterólogos não tóxicos; a presença do prófago DE3 permite a expressão da polimerase de RNA T7 induzida por IPTG, resultando em altos níveis de expressão dos genes clonados (JEONG et al., 2009). Também recomendada para altos níveis de expressão de genes não-tóxicos a partir de sistemas com promotor T7, a bactéria BL21 Star (DE3) pLysS oferece mais estabilidade ao mRNA tornando-o mais abundante e disponível para a expressão de proteínas de interesse, além do plasmídeo pLysS prevenir a expressão de proteínas recombinantes potencialmente inibidoras de crescimento em células não induzidas (GRUNBERG-MANAGO, 1999). A cepa BL21-SI utiliza a osmolaridade para a indução de expressão de proteínas recombinantes sob o promotor T7; o meio de cultura LBON (sem adição de NaCl) previne a expressão de proteínas em células não-induzidas e o choque osmótico gerado após a adição de NaCl não só promove a indução de expressão, como também pode aumentar a solubilidade das proteínas geradas em alguns casos (DONAHUE & BEBEE, 1999).

Além das cepas bacterianas de uso generalizado citadas acima, também foram escolhidas cepas específicas capazes de potencialmente superar as dificuldades envolvidas na produção das proteínas em estudo, já que são possíveis proteínas de membrana externa e possuem grandes quantidades de cisteínas ao longo de suas sequências. As bactérias BL21 C43 (DE3) permitem a superexpressão de genes codificadores de proteínas de membrana, as quais são muitas vezes tóxicas ao hospedeiro (KWON et al., 2015). Já as bactérias Origami (DE3) possuem mutações nos genes das redutases de glutatona e tiorredoxina, o que aumenta consideravelmente a formação de pontes dissulfeto no citoplasma (PRINZ et al., 1997). Finalmente, a cepa ArcticExpress (DE3) permite a produção de proteínas recombinantes a baixas temperaturas, graças à coexpressão das chaperoninas Cpn10 e Cpn60 de *Oleispira*

antarctica, as quais apresentam alta atividade de dobramento de proteínas a temperaturas entre 4 e 12°C, aumentando o rendimento de proteínas ativas e solúveis. Adicionalmente, as bactérias ArcticExpress (DE3) contêm cópias excedentes de genes que codificam tRNAs geralmente raros na tradução de proteínas heterólogas em *E. coli*, promovendo altos níveis de expressão de muitas proteínas que possuem baixo rendimento em cepas BL21 convencionais (FERRER et al., 2003).

A fim de melhorar a solubilização das proteínas recombinantes durante o processo de expressão nas cepas de bactérias *E. coli* descritas acima, também foi testada a clonagem de suas sequências nucleotídicas no vetor pSUMO_{Ulp1}. A vantagem do uso deste vetor bicistrônico é que ele possui os genes para a codificação da proteína SUMO e da protease Ulp1. A SUMO é uma proteína pequena semelhante à ubiquitina comumente utilizada em fusão a proteínas de difícil expressão (por exemplo, toxinas). A SUMO pode ser removida das proteínas fusionadas pela ação da protease Ulp1, a qual somente reconhece a estrutura terciária da SUMO, liberando as proteínas de interesse em sua forma nativa e com uma cauda de poli-histidinas que permite sua posterior purificação por IMAC (SHIMOKAWA-FALCÃO et al., 2017). Outra estratégia testada foi o cultivo de bactérias transformadas para a produção das proteínas recombinantes na presença de sorbitol, o que promove a solubilidade das proteínas recombinantes produzidas como a proteína fluorescente verde (GFP) e a porção N-terminal da proteína HypF-N, formadora de agregados amiloides (PRASAD et al., 2011).

Mesmo com as estratégias testadas acima, algumas proteínas recombinantes continuaram a ser produzidas apenas em corpúsculos de inclusão. Nestes casos, foram realizadas diálises graduais, a fim de remover gradativamente as altas concentrações de ureia (4 M ou 8 M) usadas para solubilizar as proteínas dos corpúsculos de inclusão, na ausência e presença do redutor beta-mercaptoetanol. De acordo com os experimentos de Christian Anfinsen com a RNase A, traços de beta-mercaptoetanol podem catalisar a formação das pontes

dissulfeto corretas até que a estrutura nativa e mais termodinamicamente estável das proteínas recombinantes seja atingida, o que raramente acontece apenas com a remoção da ureia (SELA et al., 1957; ANFINSEN et al., 1961; ANFINSEN, 1973). Além das diálises, a renaturação das proteínas obtidas a partir de corpúsculos de inclusão também foi realizada em colunas de cromatografia, segundo protocolos de *on-column refolding* (ROGL et al., 1998; LEMERCIER; BAKALARA; SANTARELLI, 2003; GUO et al., 2006), utilizando os sistemas ÄKTA.

A seguir, estão descritos os resultados a partir dos métodos mais eficazes para a obtenção das proteínas recombinantes que possibilitaram a análise das proteínas hipotéticas previamente identificadas.

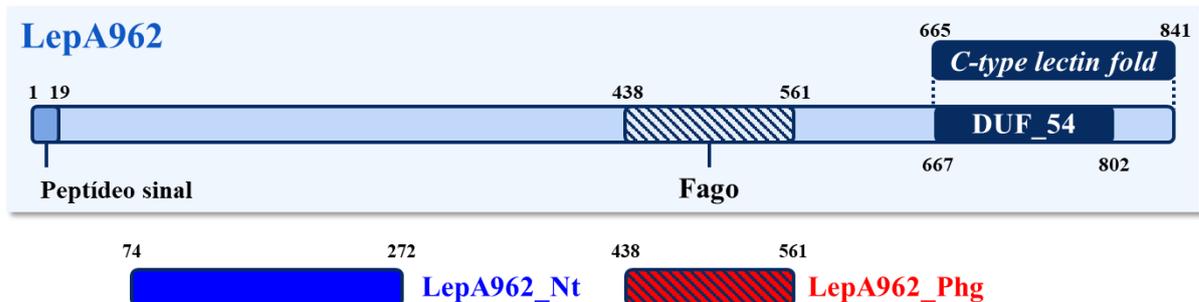
4.2. Análise da lipoproteína hipotética LepA962

4.2.1. Desenho, produção e purificação das proteínas recombinantes LepA962_Nt e LepA962_Phg

LepA962 mostrou-se uma proteína de interesse por se tratar de uma lipoproteína hipotética, possuindo um peptídeo sinal N-terminal contendo a sequência L-V-L-G-C, um possível motivo de lipoproteínas chamado *lipobox* (ZÜCKERT, 2014; EL RAYES; RODRÍGUEZ-ALONSO; COLLET, 2021). Esta proteína contém 841 resíduos de aminoácidos, e sua análise por programas de predição de peptídeos sinais, de regiões transmembrana e de localização subcelular confirmou a presença de um peptídeo sinal típico de lipoproteínas secretadas pela via Sec/SPII, sendo provavelmente ancorada na porção extracelular da membrana externa de leptospiros. Adicionalmente, a lipoproteína possui um domínio na extremidade C-terminal correspondente ao dobramento de lectina do tipo C (“*C-type lectin fold*” na Figura 5A), o qual inclui o domínio de função desconhecida referido ao longo da tese como DUF_54 (Pfam: PF07588). Segundo a ferramenta InterPro, o domínio DUF_54 está presente em 1.153 proteínas, incluindo algumas unidades nos clados eucarióticos

Cryptomonadales, *Cyathomonadacea*, Alveolata e Stramenopiles; entre as bactérias, também está presente no clado Proteobacteria, mas a grande maioria dessas proteínas (>80%) é encontrada dentro do gênero *Leptospira*. Em *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130, o domínio DUF_54 é encontrado em 9 proteínas, das quais 6 são denominadas proteínas semelhantes à endostatina de leptospiros (LenA-F), contendo uma ou duas cópias do domínio DUF_54 conhecido como “*Len-motif*” (STEVENSON et al., 2007).

A)



B)

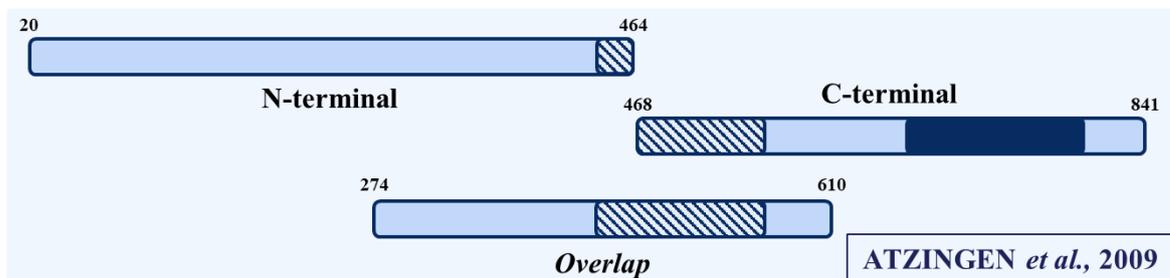


Figura 5. A) Esquema da lipoproteína hipotética LepA962, destacando a região contendo o peptídeo sinal, a sequência identificada por *phage display* (Fago) e o domínio de lectina do tipo C (*C-type lectin fold*) que inclui o domínio de função desconhecida DUF_54. Também estão representadas as duas proteínas recombinantes LepA962_Nt e LepA962_Phg, nas posições correspondentes em relação à proteína LepA962. B) Esquema com os três fragmentos estudados no trabalho de Atzingen e colaboradores (ATZINGEN et al., 2009), destacando a localização da sequência identificada por nós usando a técnica de *phage display* (áreas rajadas).

Na literatura, a lipoproteína LepA962 é conhecida como Lp95, uma proteína de 95 kDa primeiramente descrita no trabalho de Atzingen e colaboradores (ATZINGEN et al., 2009). Neste trabalho, os domínios funcionais da lipoproteína foram inicialmente explorados na

construção de fragmentos recombinantes correspondentes à proteína Lp95 inteira e às porções N-terminal, C-terminal e uma sequência de sobreposição às duas anteriores (“*Overlap*”) (Figura 5B). Utilizando a técnica de *Far Western Blotting*, foi possível observar que o fragmento C-terminal contendo o domínio de lectina/DUF_54, e em menor intensidade o fragmento “*Overlap*”, apresentaram afinidade de ligação à laminina e às fibronectinas plasmáticas e celular (ATZINGEN et al., 2009).

A partir dos dados obtidos na literatura, foi possível comparar a localização da sequência encontrada nos ensaios funcionais de *phage display* em combinação com a técnica BRASIL na proteína LepA962. A Figura 5 mostra que a sequência apresentada pelo fago (entre os resíduos de aminoácidos 438 e 561) encontra-se na intersecção dos três fragmentos recombinantes estudados no trabalho de Atzingen e colaboradores (ATZINGEN et al., 2009), estando principalmente representada no fragmento “*Overlap*”. A fim de melhor elucidar o papel desta sequência específica na função da proteína LepA962, a proteína recombinante LepA962_Ph_g foi desenhada para clonagem no vetor pAE; como controle, a proteína recombinante LepA962_N_t correspondente a uma porção N-terminal sem cisteínas da proteína também foi desenhada para clonagem em pAE (Figura 5A).

A expressão das proteínas foi induzida por choque osmótico em culturas de bactérias BL21-SI transformadas. As proteínas recombinantes LepA962_N_t e LepA962_Ph_g resultantes apresentaram massa molecular correspondente ao valor predito de 23,4 kDa e 15,2 kDa, respectivamente (Figuras 6A e 6B). Ambas as proteínas foram produzidas de forma insolúvel em corpúsculos de inclusão, mas foi possível solubilizar grande parte delas com solução contendo 4 M de ureia e, após sua solubilização, as proteínas recombinantes foram purificadas por cromatografia de afinidade a níquel, seguida por diálise gradual para a remoção da ureia (Figuras 6A e 6B).

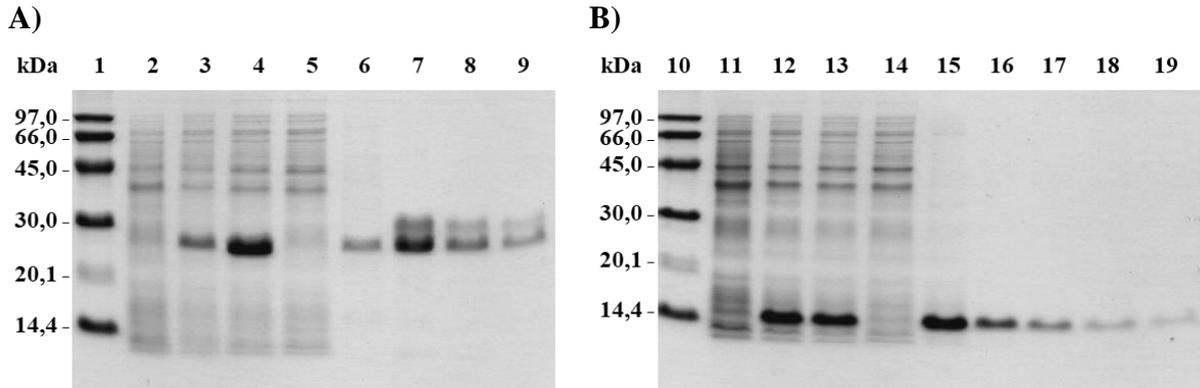


Figura 6. SDS-PAGE com amostras referentes às proteínas recombinantes LepA962_Nt (**A**) e LepA962_Phg (**B**), desde a sua indução em bactérias BL21-SI até sua purificação por IMAC e posterior diálise. Canaletas 1 e 10: marcador de massa molecular; canaletas 2 e 11: culturas de bactéria não induzida; canaletas 3 e 12: culturas de bactéria após indução; canaletas 4 e 13: lisado bacteriano; canaletas 5 e 14: sobrenadante do lisado bacteriano após centrifugação; canaletas 6 e 15: precipitado do lisado bacteriano após solubilização em 4 M de ureia; canaletas 7, 16 e 17: coletas da purificação em coluna HisTrap; canaletas 8 e 18: amostras dialisadas (diluições mais concentradas); canaletas 9 e 19: amostras dialisadas (diluições menos concentradas). SDS-PAGE com géis de poliacrilamida de 15%, corados com Coomassie Blue.

Durante a purificação e diálises da proteína recombinante LepA962_Nt, uma banda menos intensa e pouco maior em massa molecular (~30 kDa) foi observada nas amostras analisadas (Figura 6A). Como tentativa de separá-la da banda inferior principal, foram realizadas duas modalidades de cromatografia: troca aniônica e interação hidrofóbica, em amostras purificadas em 4 M de ureia. Mesmo após as duas purificações, as bandas continuaram a ser eluídas nas mesmas frações (Figuras 7A e 7B). A fim de identificar a origem das bandas e confirmar qual das duas correspondiam à proteína recombinante LepA962_Nt, cada banda foi analisada por espectrometria de massas com ionização por *electrospray* (ESI), seguida pela combinação de analisadores de massas por quadrupolo e tempo de voo (QUAD-TOF). O sequenciamento dos fragmentos gerados de ambas as bandas resultou em trechos da sequência referente à proteína recombinante LepA962_Nt e, portanto, a banda adicional não se tratava de um contaminante e as duas bandas representam esta proteína.

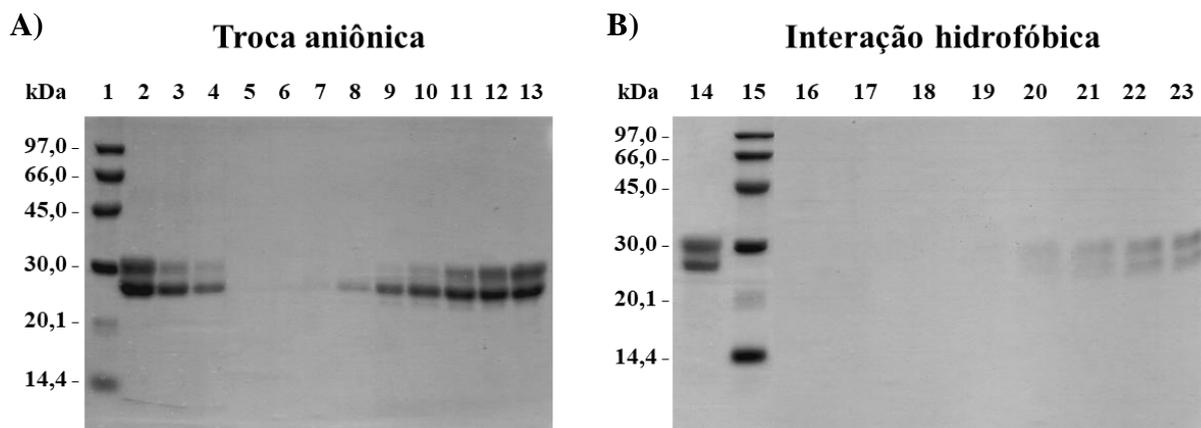


Figura 7. SDS-PAGE com amostras de purificação por troca aniônica (**A**), seguida de purificação por interação hidrofóbica (**B**) da proteína LepA962_Nt previamente purificada por IMAC, após diálise para remoção de NaCl da solução (concentração final de aproximadamente 0,3 mM). Em **A**), estão representadas as seguintes amostras: proteína LepA962_Nt em 4 M ureia a ser purificada por troca aniônica em resina *Q Sepharose Fast Flow* (canaleta 2), coleta durante aplicação da amostra (*flow-through*; canaleta 3), coleta após lavagem da coluna (canaleta 4), e coletas de eluição com concentrações de NaCl variando de 80 mM a 160 mM (canaletas 5 a 13). Em **B**), estão representadas as seguintes amostras: proteína LepA962_Nt em 4 M ureia e 200 mM NaCl a ser purificada por interação hidrofóbica em resina *Phenyl Sepharose 6 Fast Flow High Sub* (canaleta 14), e coletas da purificação *in batch* com concentrações de NaCl decrescentes (2500 mM → 2000 mM → 1500 mM → 1000 mM → 750 mM → 500 mM → 250 mM → 0 mM; canaletas 16 a 23). Canaletas 1 e 15: marcador de massa molecular. SDS-PAGE com géis de poli(acrilamida) de 15%, corados com Coomassie Blue.

4.2.2. Análise da estrutura secundária das proteínas recombinantes LepA962_Nt e

LepA962_Phg e obtenção de antissoros

A recuperação das proteínas recombinantes LepA962_Nt e LepA962_Phg após a remoção da ureia foi bem sucedida, sendo possível sua análise estrutural e funcional. Inicialmente, a estrutura secundária de cada proteína foi avaliada por dicróismo circular, e as elipticidades molares resultantes entre os comprimentos de onda de 188 nm a 260 nm revelaram curvas compatíveis com proteínas estruturadas (Figuras 8A e 8B). A deconvolução dos dados obtidos mostrou que a proteína LepA962_Nt possui 4% de α -hélices, 35% de fitas β , 20% de voltas e 40% de segmentos desordenados (Figura 8A), enquanto que a proteína LepA962_Phg possui 5% de α -hélices, 30% de fitas β , 18% de voltas e 47% de segmentos desordenados (Figura 8B).

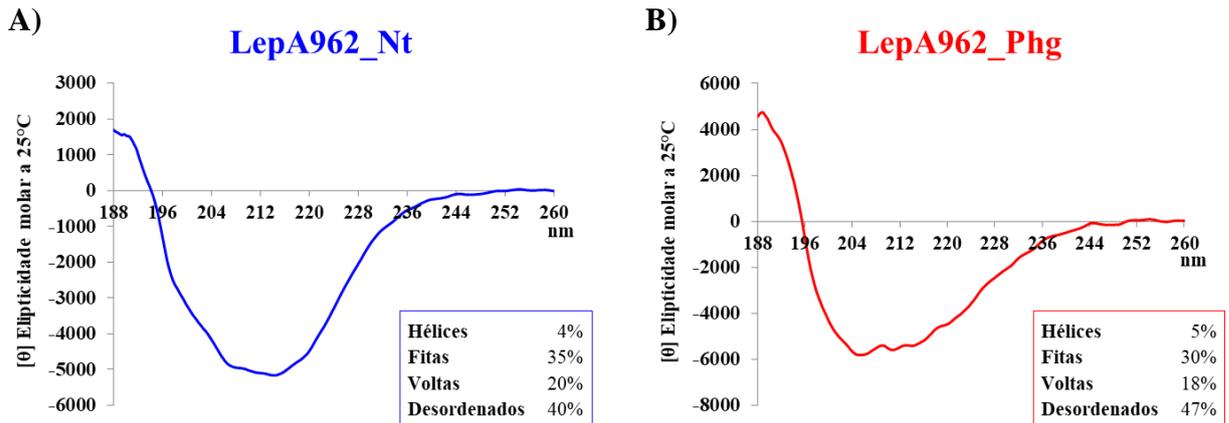


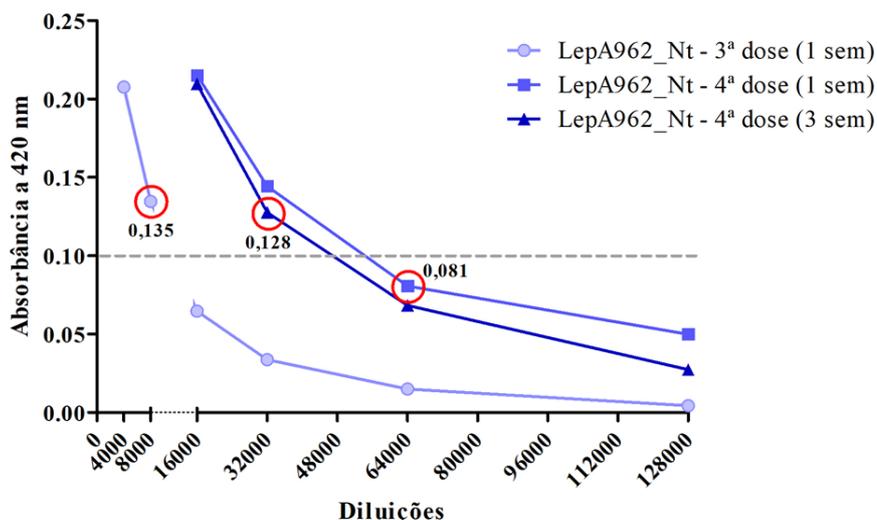
Figura 8. Espectros de elipticidade molar obtidos pela análise de dicroísmo circular das proteínas recombinantes LepA962_Nt (A) e LepA962_Phg (B). Dados gerados a partir de dez leituras da polarização de luz medida a 25 °C e no intervalo de 188 a 260 nm, com incrementos de 0,1 nm, e posteriormente convertidos em valores de elipticidade molar segundo Errington e colaboradores (ERRINGTON; IQBALSZYAH; DOIG, 2006). A interpretação das curvas para revelar a proporção das principais estruturas secundárias em cada proteína analisada foi realizada pelo algoritmo CDSSTR disponível no servidor Dichroweb (quadros inferiores e à direita de cada gráfico).

Em paralelo, os antissoros das proteínas recombinantes foram obtidos a partir da imunização de camundongos com doses administradas a cada 7 dias. A avaliação dos títulos dos antissoros foi iniciada após uma semana da 3ª dose de cada proteína, resultando nos títulos de 1:8.000 para o antissoros de LepA962_Nt (Figura 9A) e apenas 1:160 para o antissoros de LepA962_Phg (Figura 9B).

O baixo título do antissoros de LepA962_Phg foi resultado da degradação da proteína recombinante durante seu armazenamento a 4 °C. A obtenção de mais proteína recombinante LepA962_Phg purificada e dialisada em paralelo permitiu que o esquema de imunização continuasse sem interrupções, sendo necessárias mais 3 doses além das 4 doses planejadas para o esquema de imunização (total de 7 doses). Além disso, alíquotas da proteína LepA962_Phg passaram a ser armazenadas a -20 °C, garantindo uma maior estabilidade de sua estrutura. Ao final de cada esquema de imunização de 4 doses (LepA962_Nt) ou 7 doses (LepA962_Phg), os antissoros obtidos em 1, 2 ou 3 semanas após a última dose atingiram o título satisfatório de 1:64.000 (Figuras 9A e 9B).

A)

Avaliação dos títulos dos antissoros de
LepA962_Nt



B)

Avaliação dos títulos dos antissoros de
LepA962_Phg

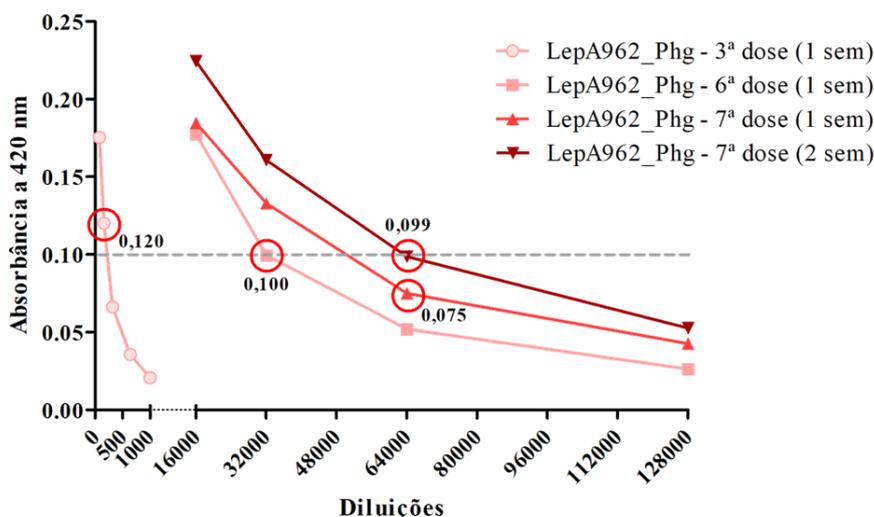


Figura 9. Gráficos dos títulos de antissoros obtidos em camundongos imunizados com as proteínas recombinantes LepA962_Nt (A) e LepA962_Phg (B), avaliados em ensaios de ELISA. Soros obtidos de amostras de sangue coletadas após 1, 2 ou 3 semanas das doses indicadas (1 sem, 2 sem e 3 sem, respectivamente). Valores plotados após subtração dos valores obtidos nos poços sem adição de antissoros. Títulos definidos pelos valores de absorbância a 492 nm obtidos mais próximos a 0,1 (rotulados e destacados por círculos vermelhos),

Após os ajustes de armazenamento ideal das proteínas recombinantes LepA962_Nt e LepA962_Phg e da obtenção de seus antissoros com alto título, foi realizada a avaliação das proteínas e de seus antissoros em SDS-PAGE e *Western Blotting*. Ambas as proteínas se

apresentaram íntegras e com tamanho esperado de 23 kDa (LepA962_Nt) e 15 kDa (LepA962_Phg), enquanto que os antissoros foram bem sucedidos em reconhecer sua proteína recombinante correspondente, utilizando uma concentração de apenas 1 μ M de cada proteína e a diluição de 1:5.000 dos antissoros (Figura 10).

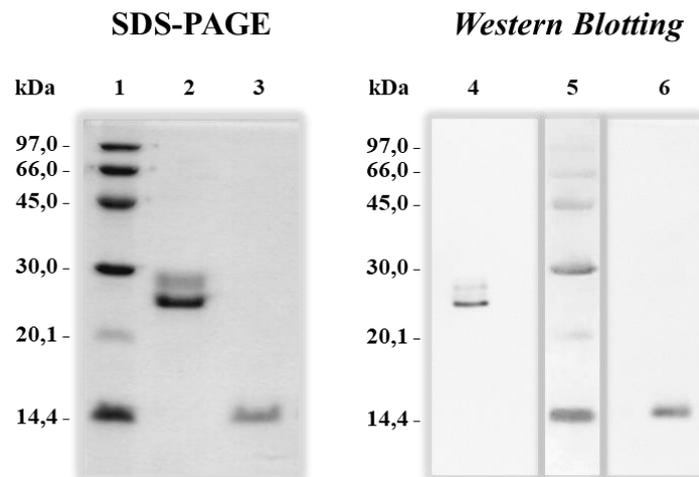


Figura 10. Análise das proteínas recombinantes LepA962_Nt e LepA962_Phg e de seus antissoros obtidos em camundongos. À esquerda, está representado o SDS-PAGE com as proteínas recombinantes LepA962_Nt (canaleta 2) e LepA962_Phg (canaleta 3) purificadas e dialisadas, com concentração de 2 μ M cada. À direita, está representado o *Western Blotting* com as proteínas recombinantes LepA962_Nt (canaleta 4) e LepA962_Phg (canaleta 6) purificadas e dialisadas, com concentração de 1 μ M cada, e detectadas com seu antissoro correspondente diluído 1:5.000. Canaletas 1 e 5: marcador de massa molecular. Géis de poliacrilamida de 15%, e SDS-PAGE corado com Coomassie Blue.

4.2.3. Análise da interação entre as proteínas recombinantes (LepA962_Nt e LepA962_Phg) e componentes da MEC

A caracterização das proteínas recombinantes LepA962_Nt e LepA962_Phg seguiu para o estudo de sua capacidade de interagir com alguns dos componentes presentes na MEC. O gráfico na Figura 11 mostra que a proteína recombinante LepA962_Nt apresentou uma discreta interação com as proteínas de MEC escolhidas, em conformidade com os dados obtidos na literatura. Já a proteína recombinante LepA962_Phg apresentou uma marcante interação com as mesmas proteínas de MEC, principalmente com a fibronectina plasmática, colágeno I,

fibrinogênio e laminina, com as quais obteve níveis de absorvância superiores à proteína controle recombinante LigBC, correspondente a uma porção recombinante da região C-terminal da proteína LigB, um importante fator de virulência de leptospiros e de sua interação com o hospedeiro (Figura 11). Os dados representam a absorvância de cada interação, subtraindo-se a interação das proteínas recombinantes com o controle negativo BSA; nenhuma das proteínas recombinantes apresentou interação com decorina superior à interação com BSA (dados não apresentados).

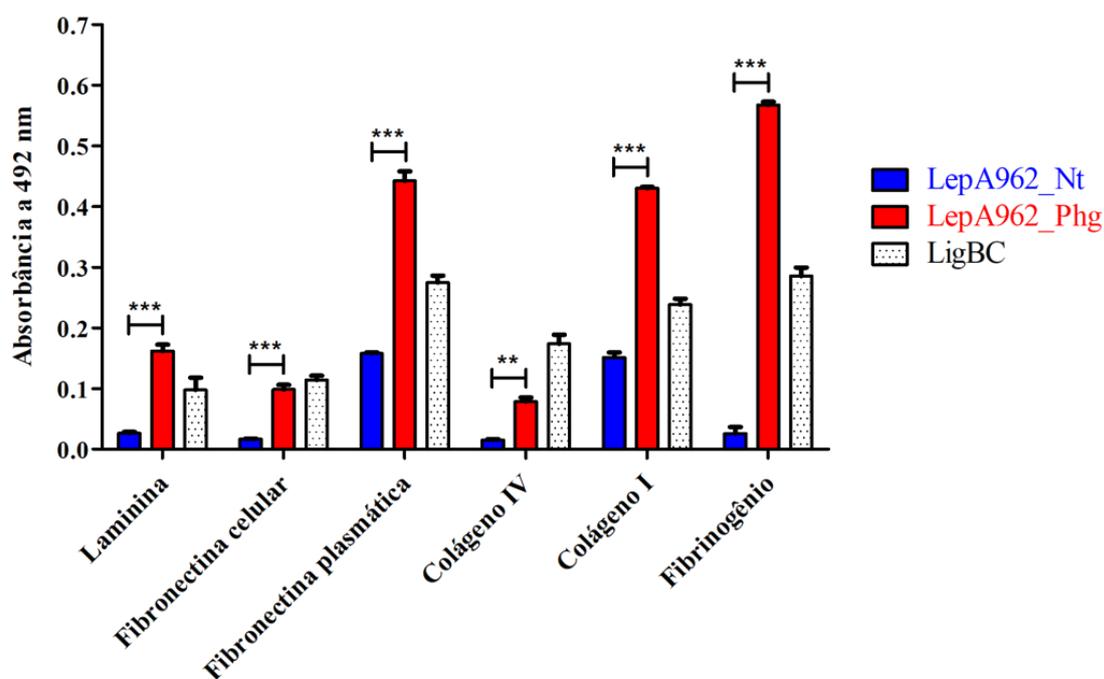


Figura 11. Análise da ligação das proteínas recombinantes LepA962_Nt e LepA962_Phg a diferentes componentes da MEC e plasma por ELISA. As proteínas laminina, fibronectinas plasmática e celular, colágenos IV e I e fibrinogênio foram utilizadas para a sensibilização das placas, na concentração de 1 µg/poço. Após bloqueio, 1 µg de cada proteína recombinante foi adicionado aos poços em triplicatas. A interação entre as proteínas de matriz/plasma e as proteínas recombinantes foi detectada com a adição do respectivo antissoro diluído a 1:5.000. Proteína recombinante LigBC adicionada como controle positivo. Dados de absorvância a 492 nm calculados, descontando os valores de controle (poços sem adição de antissoros) e da interação de cada proteína recombinante ao BSA (dados não representados). Estatística representada apenas entre os grupos LepA962_Nt e LepA962_Phg, com análise de variância *One-way* ANOVA e comparação múltipla de Tukey; $P < 0,05$ (**) ou $P < 0,001$ (***)

As proteínas de MEC com as quais a proteína recombinante LepA962_Phg apresentou as maiores interações (laminina, fibronectina plasmática, colágeno I e fibrinogênio) foram utilizadas para a determinação de curvas dose-resposta (Figura 12).

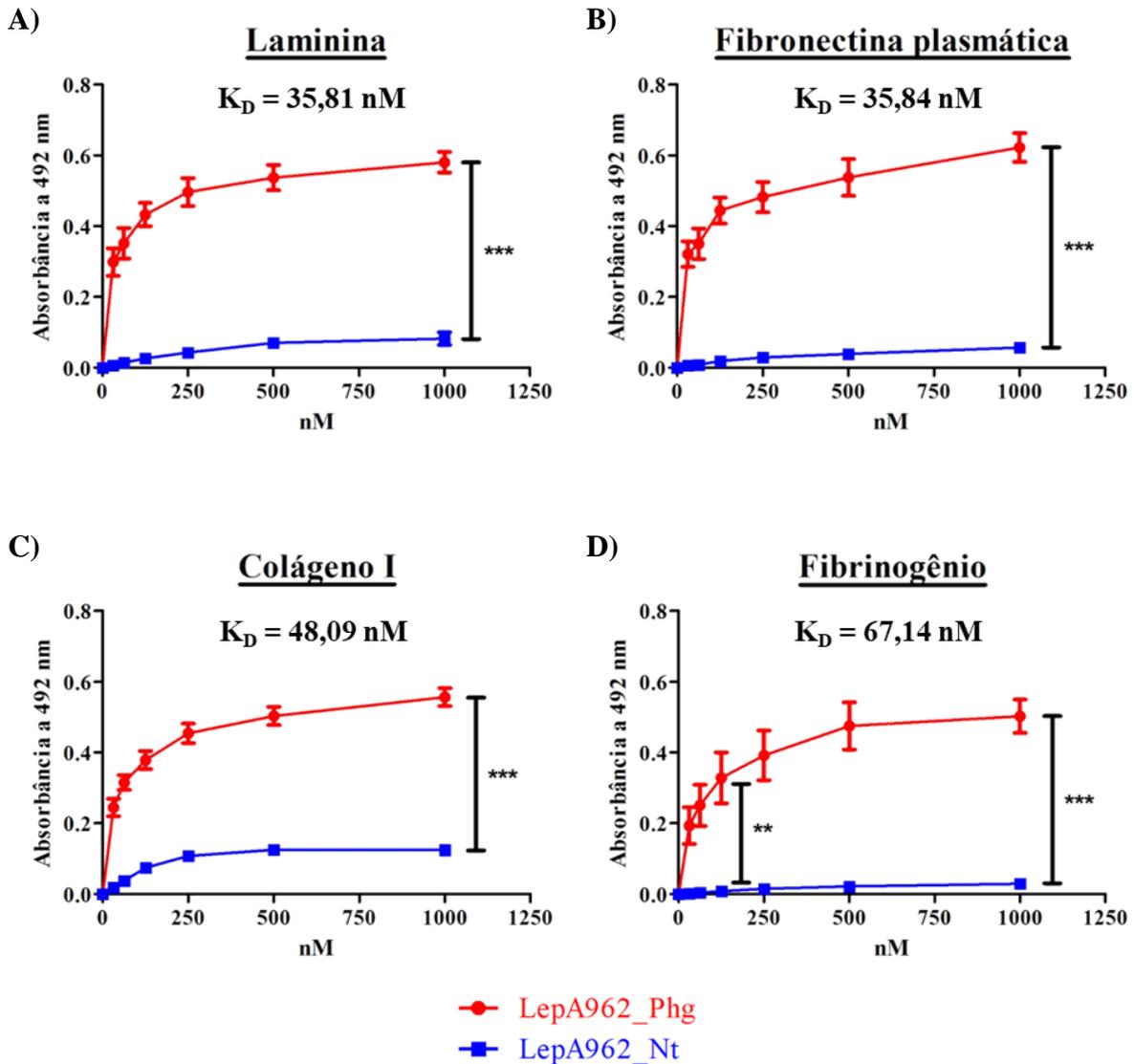


Figura 12. ELISA das curvas dose-resposta das proteínas recombinantes LepA962_Nt e LepA962_Phg com componentes da MEC e plasma selecionadas. As proteínas laminina (A), fibronectina plasmática (B), colágeno I (C) e fibrinogênio (D) foram utilizadas para a sensibilização das placas, na concentração de 1 $\mu\text{g}/\text{poço}$. Após bloqueio, concentrações crescentes de cada proteína recombinante, variando de 0 a 1000 nM foram adicionadas aos poços em duplicata. A interação entre as proteínas de matriz/plasma e as proteínas recombinantes foi detectada com a adição do respectivo antissoro diluído a 1:5.000. Dados de absorbância a 492 nm obtidos a partir de três ensaios independentes. Estatísticas: teste t não pareado (*two-tailed*); $P < 0,005$ (**) e $P < 0,001$ (***); Coeficientes de dissociação (K_D) determinados pelo software GraphPad Prism (Regressão não linear > Ligação - Saturação > 1 sítio - Ligação específica).

Foram utilizadas concentrações entre 0 e 1.000 nM das proteínas recombinantes LepA962_Phg e LepA962_Nt, além do controle negativo BSA. Os resultados apresentados nos gráficos da Figura 12 representam os dados descontando as interações apresentadas entre as proteínas da MEC e BSA. As diferenças de interação das proteínas LepA962_Phg e LepA962_Nt foram significativas e os coeficientes de dissociação (K_D) calculados para a proteína LepA962_Phg foram de 35,81 nM com laminina, 35,84 nM com fibronectina plasmática, 48,09 nM com colágeno I e 67,14 nM com fibrinogênio (Figura 12).

4.2.4. Análise da interação entre as proteínas recombinantes (LepA962_Nt e LepA962_Phg) a células VERO e sua MEC secretada

A capacidade de ligação das proteínas recombinantes LepA962_Nt e LepA962_Phg também foi testada em células VERO em cultura e sua MEC secretada em ensaio baseado na técnica de ELISA (Figura 13).

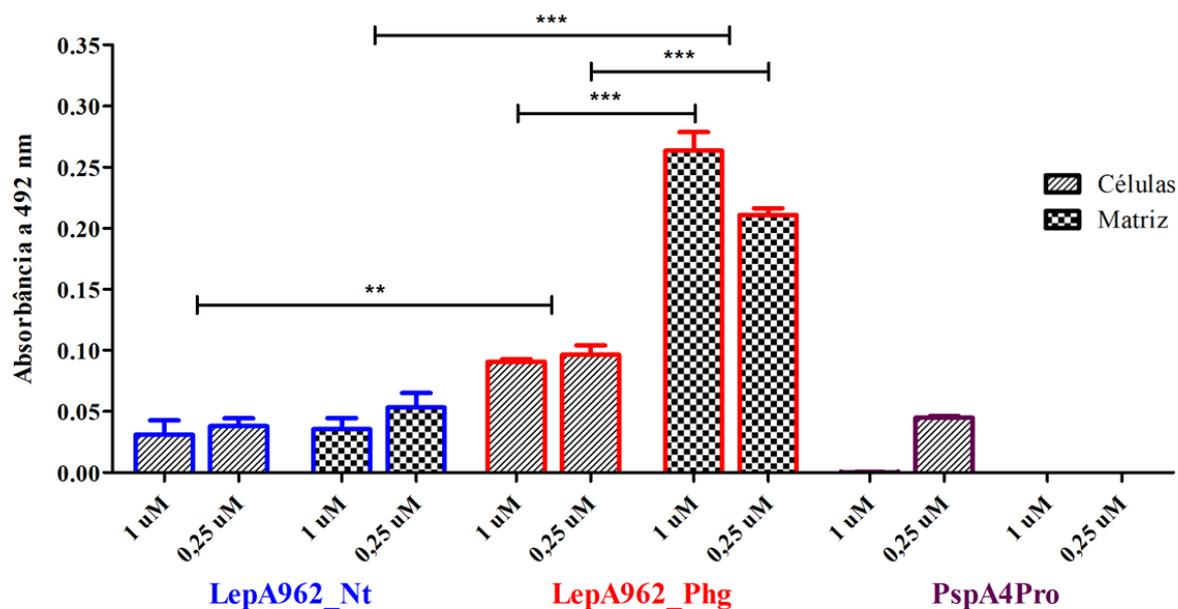


Figura 13. Ligação das proteínas recombinantes LepA962_Nt e LepA962_Phg a células VERO e a sua MEC secretada. Após adição das células a uma placa de 96 poços tratada (5×10^4 células/poço), seguida

por incubação por 24 h a 37 °C e 5% CO₂, parte das células foi removida por lavagem com 10 mM EDTA, a fim de expor a matriz extracelular secretada. As proteínas recombinantes foram adicionadas em duas concentrações (1 µM e 0,25 µM), em triplicatas. A proteína recombinante PspA4Pro também foi adicionada nas mesmas concentrações como controle negativo de proteína produzida em bactérias BL21-SI. Após fixação, a interação entre as células e sua matriz secretada com as proteínas recombinantes foi detectada em ensaio baseado em ELISA, com a adição dos respectivos antissoros diluídos a 1:5.000. Dados de absorvância a 492 nm. Estatística: teste t não pareado (*two-tailed*) e comparação múltipla de Tukey; P < 0,05 (**) ou P < 0,001 (***).

As interações das proteínas foram avaliadas com duas concentrações diferentes: 1 µM e 0,25 µM, também sendo adicionada uma porção recombinante da proteína A de superfície pneumocócica (PspA4Pro) como um controle de proteína marcada com 6x-His e produzida em bactérias BL21-SI. A proteína LepA962_Phg apresentou maior interação às células VERO e sua MEC secretada do que as proteínas LepA962_Nt e PspA4Pro, principalmente com a MEC secretada (Figura 13), reforçando sua forte interação com proteínas de MEC.

4.2.5. Análise da interação dos antissoros das proteínas recombinantes LepA962_Nt e LepA962_Phg à superfície de leptospiros

Além da avaliação da capacidade de ligação das proteínas recombinantes LepA962_Nt e LepA962_Phg com elementos do hospedeiro, seus antissoros também foram avaliados quanto à sua ligação à superfície de leptospiros. A partir de um ensaio de ELISA sanduíche, a fase de captura consistiu em soros obtidos em coelhos pré-ímmunes ou desafiados com *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130, seguida pela adição de leptospiros virulentos da mesma cepa, dos antissoros correspondentes e do anticorpo secundário para detecção. O antissoro da proteína recombinante LipL32 também foi adicionado como controle positivo por representar uma lipoproteína encontrada em abundância na superfície de leptospiros patogênicas.

Como esperado, a interação do antissoro de LipL32 à superfície das leptospiros foi superior aos dos antissoros das proteínas recombinantes de LepA962, entretanto, a diferença

dos valores obtidos para os antissoros de LepA962_Phg e LepA962_Nt foram considerados significativos, indicando uma maior exposição do trecho correspondente à proteína LepA962_Phg na superfície de leptospiros (Figura 14).

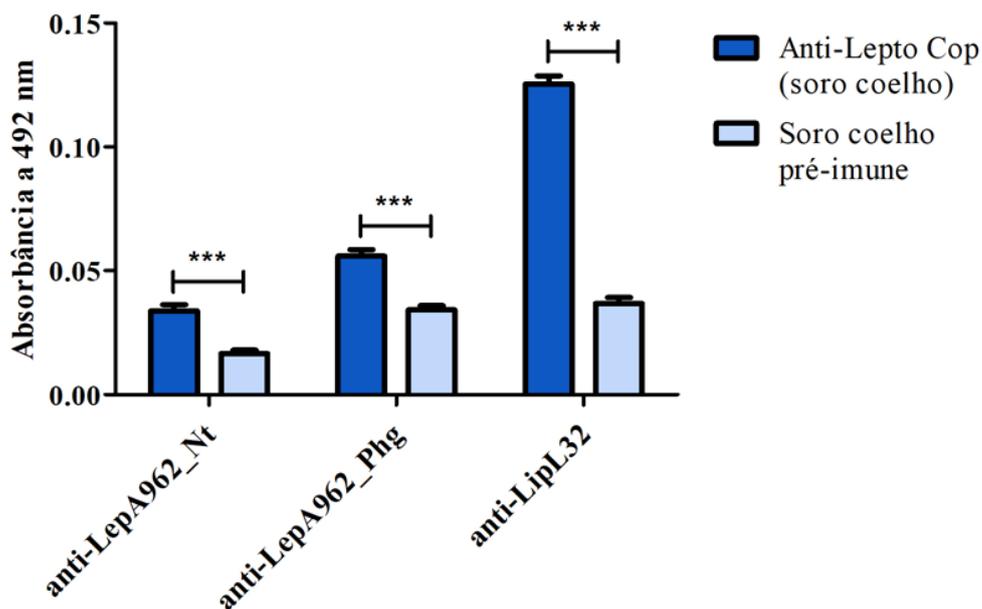


Figura 14. Ligação dos antissoros das proteínas recombinantes LepA962_Nt e LepA962_Phg à superfície de leptospiros. Ensaio de ELISA sanduíche, onde os anticorpos de captura aplicados foram: soro de coelho pré-imune e soro de coelho contra o extrato proteico de *L. interrogans* sorovar Copenhageni (diluição 1:2.000). Leptospiros virulentas da cepa Fiocruz L1-130 obtidas de culturas em meio EMJH foram adicionadas na concentração de 10^7 células/poço para captura, sendo detectadas pela posterior adição dos antissoros das proteínas recombinantes de LepA962 diluídos 1:200. Como controle positivo, o antissoro da proteína LipL32 diluído a 1:1.000 também foi adicionado. Dados de absorbância a 492 nm. Estatística representada para comparar a interação de cada antissoro às leptospiros utilizando os dois soros de coelho (captura): teste t não pareado (*two-tailed*); $P < 0,001$ (***)

4.3. Análise da lipoproteína hipotética LepA388

4.3.1. Expressão gênica em camundongos infectados com *L. interrogans*

LepA388 também foi escolhida como proteína de interesse por se tratar de uma proteína hipotética possivelmente secretada por leptospiros. A análise desta proteína de 631 resíduos de aminoácidos por programas de predição de peptídeos sinais, de regiões transmembrana e de localização subcelular revelou a presença de um peptídeo sinal N-terminal típico de proteínas secretadas pela via Sec/SPI, sendo uma proteína provavelmente secretada para o meio extracelular das leptospiros. Entretanto, a característica mais marcante desta proteína é que ela faz parte de uma família de proteínas parálogas em leptospiros com sequências semelhantes entre si, reunidas pelo domínio de função desconhecida referido ao longo da tese como DUF_61 (Pfam: PF07598), o qual ocupa praticamente toda a extensão das proteínas. Segundo a ferramenta InterPro, o domínio DUF_61 está presente em 818 proteínas apenas no reino Eubacteria: nos clados Actinobacteria e Firmicutes (< 0,4%); no clado Proteobacteria (15,4%), incluindo unidades nos gêneros *Helicobacter*, *Campylobacter*, *Bartonella*, *Piscirickettsia* e na família *Vibrionaceae*. No gênero *Leptospira* (> 84%), o domínio está presente apenas em espécies patogênicas, além de ter sofrido expansão nas espécies consideradas mais virulentas (LEHMANN et al., 2014; FOUTS et al., 2016; PICARDEAU, 2017).

A presença da família DUF_61 em espécies patogênicas de leptospiros e a relação de seus membros com a interação patógeno-hospedeiro e virulência foram demonstradas em trabalhos anteriores (MATSUNAGA et al., 2007; LEHMANN et al., 2013; MARCSISIN et al., 2013; CHOU et al., 2014; PUTZ et al., 2021). Em especial, o trabalho de Lehmann e colaboradores apresentou o estudo de mutações de nucleotídeo único em bactérias *L. interrogans* sorovar Lai 56601 atenuadas por 18 passagens em subculturas *in vitro*, encontrando mutações em duas proteínas da família DUF_61, uma das quais é ortóloga à proteína LepA388; além disso, análises de RT-qPCR mostraram que todos os membros desta família foram

regulados positivamente na circulação e fígado de hamsters infectados (LEHMANN et al., 2013). A partir dos resultados obtidos na literatura, foi realizada uma reação de qPCR com oligonucleotídeos específicos para a proteína LepA388 e cDNA obtido a partir do RNA extraído do fígado de camundongos após três dias do desafio com *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130. O aumento da expressão do gene de LepA388 foi confirmado, resultando em um aumento médio de 40 vezes, em comparação à expressão do gene constitutivo ribossomal 16S (Figura 15).

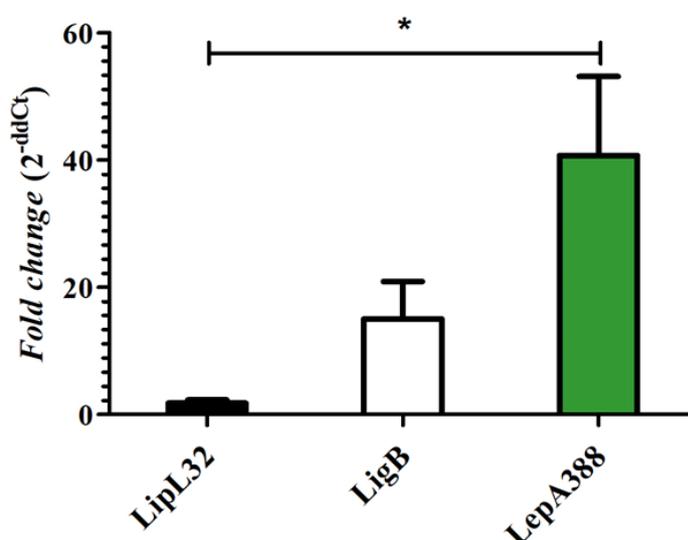


Figura 15. Análise da expressão dos genes codificadores de LipL32, LigB e LepA388, apresentada em *fold change* e pelo método delta-delta (2^{-ddCt}), a partir de RT-qPCR utilizando amostras de fígado de camundongos susceptíveis C3H/HeJ após três dias de infecção por *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130. As amostras foram analisadas em triplicata para cada par de oligonucleotídeo, e os dados foram normalizados a partir da expressão do gene ribossomal 16S e comparadas à expressão dos mesmos genes em leptospiras cultivadas em meio EMJH a 30 °C. Análise estatística por *One-way* ANOVA, com $P < 0,05$ (*).

4.3.2. Predições da estrutura tridimensional da proteína LepA388

4.3.2.1. Modelos a partir do servidor Phyre2

As dificuldades prévias encontradas para a produção de porções recombinantes da proteína LepA388 incentivou o estudo mais detalhado da possível estrutura tridimensional desta

proteína. Inicialmente, foi realizada a predição de sua estrutura com o auxílio do servidor Phyre2, o qual utiliza uma combinação de modelagem com múltiplos moldes proteicos já conhecidos intercalados com regiões simuladas de forma *ab initio*, ou seja, apenas pelo fornecimento de estrutura primária da proteína de interesse (Figura 16).

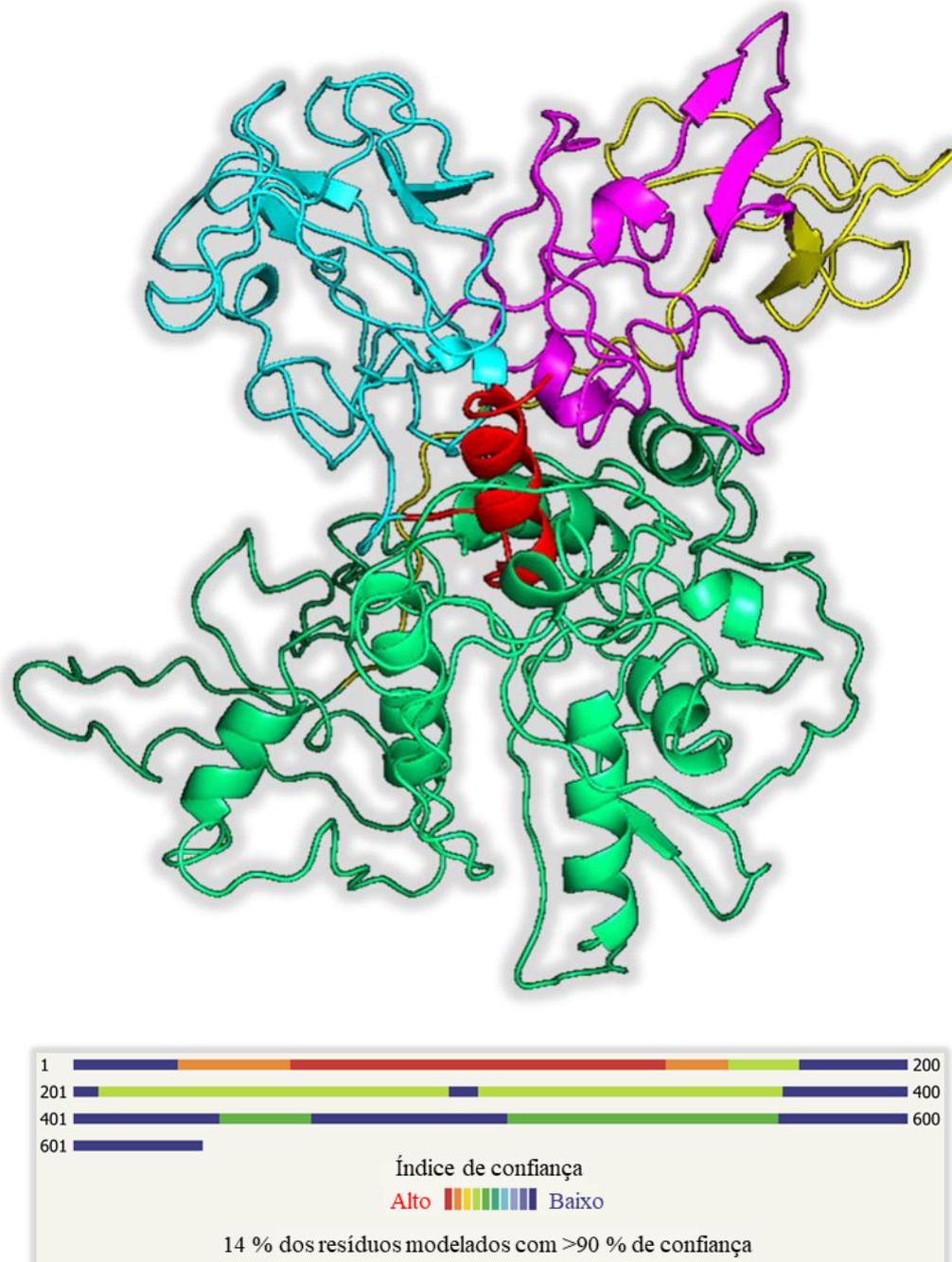


Figura 16. Modelo tridimensional da proteína LepA388 gerado pelo servidor Phyre2. Representação modificada no software PyMol para a delimitação de possíveis domínios da proteína: peptídeo sinal (em

vermelho), domínio de lectina B (em azul ciano), sequência identificada por *phage display* (em rosa), região C-terminal do possível “domínio 1” (porção N-terminal da proteína, em amarelo) e a porção C-terminal da proteína (em verde). O quadro inferior mostra o índice de confiança ao longo da sequência de LepA388, onde 14% dos resíduos foram modelados com > 90% de confiança, representados pela região em vermelho coincidente com o possível domínio de lectina B.

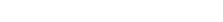
O modelo gerado baseou-se principalmente na semelhança estrutural de 14% dos resíduos de LepA388 com modelos do banco de dados de proteínas tridimensionais (PDB) referentes a domínios de ligação a carboidratos de diversas proteínas, que possuem o dobramento semelhante ao domínio B da toxina ricina (Tabela 4). A porção N-terminal da proteína LepA388 foi modelada com mais de 90% de confiança, sendo o domínio de lectina B da toxina abrina (PDB 1abr) utilizado como referência de maior confiança para o modelo (Figura 16).

Curiosamente, a única proteína que contém a anotação desse domínio em *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130 é uma proteína pertencente à família DUF_61 (LepA400). O alinhamento múltiplo entre as proteínas DUF_61 na cepa Fiocruz L1-130 mostra a semelhança em toda a sua extensão (Figura S4), mas principalmente na região N-terminal, onde se encontra o domínio semelhante à lectina B da ricina em LepA400 (Figura 17A). A análise detalhada desta sequência destaca as semelhanças, como a posição de quatro resíduos de cisteínas e as possíveis repetições QxW, as quais são características marcantes do domínio semelhante à lectina B da ricina (RUTENBER et al., 1987; HAZES, 1996), e que estão estruturalmente alinhadas no modelo gerado no servidor Phyre2 com a estrutura correspondente na toxina abrina (Figura 17B). Esta sobreposição também foi analisada pelo software UCSF Chimera para o cálculo do desvio médio quadrático (RMSD) das posições atômicas após o alinhamento das estruturas com a ferramenta *MatchMaker*: o RMSD obtido foi de 0,864 Å, confirmando a grande semelhança entre as estruturas analisadas. Todas as análises indicam que o domínio semelhante à lectina B da ricina está presente não só na proteína LepA400, mas também na proteína LepA388 e possivelmente em todas as proteínas da família DUF_61.

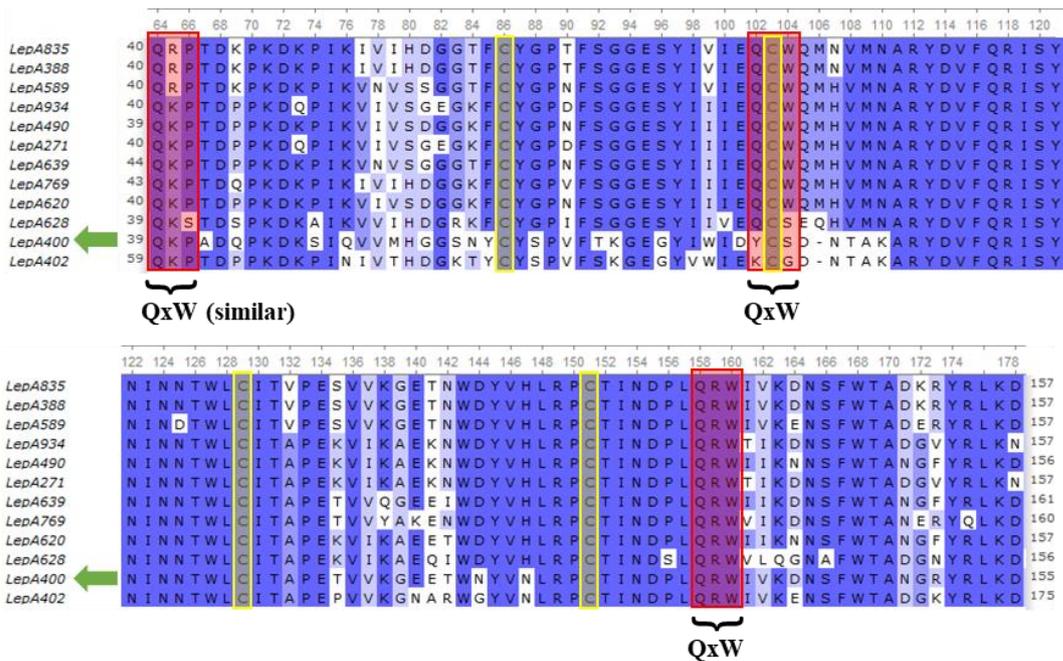
Devido ao baixo nível de confiança no restante da estrutura gerada pelo servidor Phyre2, novas análises resultaram em diferentes modelos, representados na Figura 18. Apesar da diferença entre si, é possível ver a esperada conservação do domínio semelhante à lectina B da ricina (em azul ciano) e uma ligeira separação entre os domínios N- e C-terminais (em verde).

Tabela 4

Relação dos principais modelos estruturais alinhados à proteína LepA388 pelo servidor Phyre2.

Nº	PDB	Descrição	Espécie	Cobertura do alinhamento	Confiança	Identidade (%)
1	1ABR_B	Abrina-A	<i>Abrus precatorius</i>		93,2	19
2	1ONK_B	Lectina I do visco	<i>Viscum album</i>		91,8	19
3	1GGP_B	Proteína lectina 1, cadeia B	<i>Trichosanthes kirilowii</i>		91,2	14
4	1OQL_B	Lectina I do visco	<i>Viscum album</i>		89,9	13
5	2Q3N_B	Aglutinina	<i>Abrus precatorius</i>		89,4	15
6	1XHB_A	Polipeptídeo N-acetilgalactosaminiltransferase-T1	<i>Mus musculus</i>		89,4	16
7	2AAI_B	Ricina, cadeia B	<i>Ricinus communis</i>		89,3	19
8	1DQG_A	Receptor de manose rico em cisteínas	<i>Mus musculus</i>		88,5	16
9	1CE7_B	Lectina I do visco	<i>Viscum album</i>		88,1	13
10	1HWM_B	Ebulina	<i>Sambucus ebulus</i>		84,9	23
11	1XHB_A	Polipeptídeo N-acetilgalactosaminiltransferase-T1	<i>Mus musculus</i>		84,6	14
12	3PG0_A	Proteína projetada de 3 dobras simétricas (<i>Three Foil</i>)	Construção sintética		83,9	11
13	2FFU_A	Polipeptídeo N-acetilgalactosaminiltransferase 2	<i>Homo sapiens</i>		81,9	18
14	2AAI_B	Ricina, cadeia B	<i>Ricinus communis</i>		81,9	15
15	1RZO_B	Aglutinina	<i>Ricinus communis</i>		81,5	14
16	4HR6_C	Lectina da semente de cabaça de cobra	<i>Trichosanthes anguina</i>		81	17
17	1M2T_B	Lectina I do visco, cadeia B	<i>Viscum album</i>		80,9	13
18	2RG9_B	Receptor de manose de macrófago 1	<i>Homo sapiens</i>		79,9	16
19	1RZO_B	Lectina 1 específica de beta-galactosídeo, cadeia B	<i>Viscum album</i>		78,7	16
20	1OQL_B	Lectina I do visco	<i>Viscum album</i>		76,4	16

A)



B)

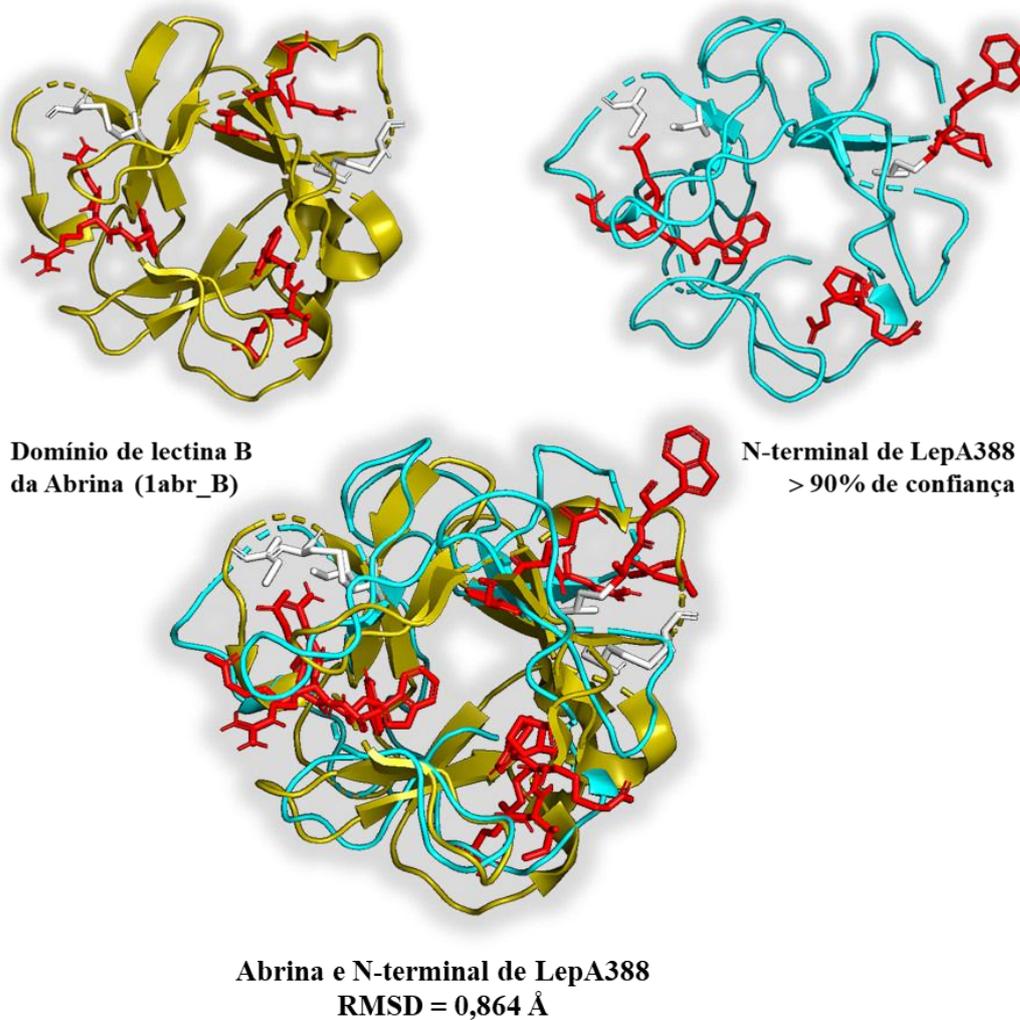


Figura 17. A) Alinhamento múltiplo das 12 principais proteínas da família DUF_61 encontradas na bactéria *L. interrogans* sorovar Copenhageni cepa Fiocruz LI-130, considerando a região

correspondente ao domínio de lectina B anotado somente na proteína LepA400 (setas verdes), e com código de cores baseado na porcentagem de identidade entre as sequências. Os quadros delimitam as repetições QxW características deste domínio (em vermelho) e os 4 resíduos de cisteína conservados em todas as sequências proteicas (em amarelo). Alinhamento múltiplo obtido pela ferramenta Clustal Omega e analisado pelo software UGene. **B)** Representação do domínio de lectina B da Abrina (PBD 1abr_B; à esquerda, em amarelo ouro), do modelo da região N-terminal da proteína LepA388 gerado pelo servidor Phyre2 e correspondente ao domínio de lectina B (à direita, em azul ciano) e do alinhamento dos dois domínios (ao centro). Representações modificadas no software PyMol para destacar as repetições QxW (em vermelho) e os resíduos de cisteína/pontes dissulfeto (em branco). Cálculo do RMSD em angstroms realizado pelo software UCSF Chimera.

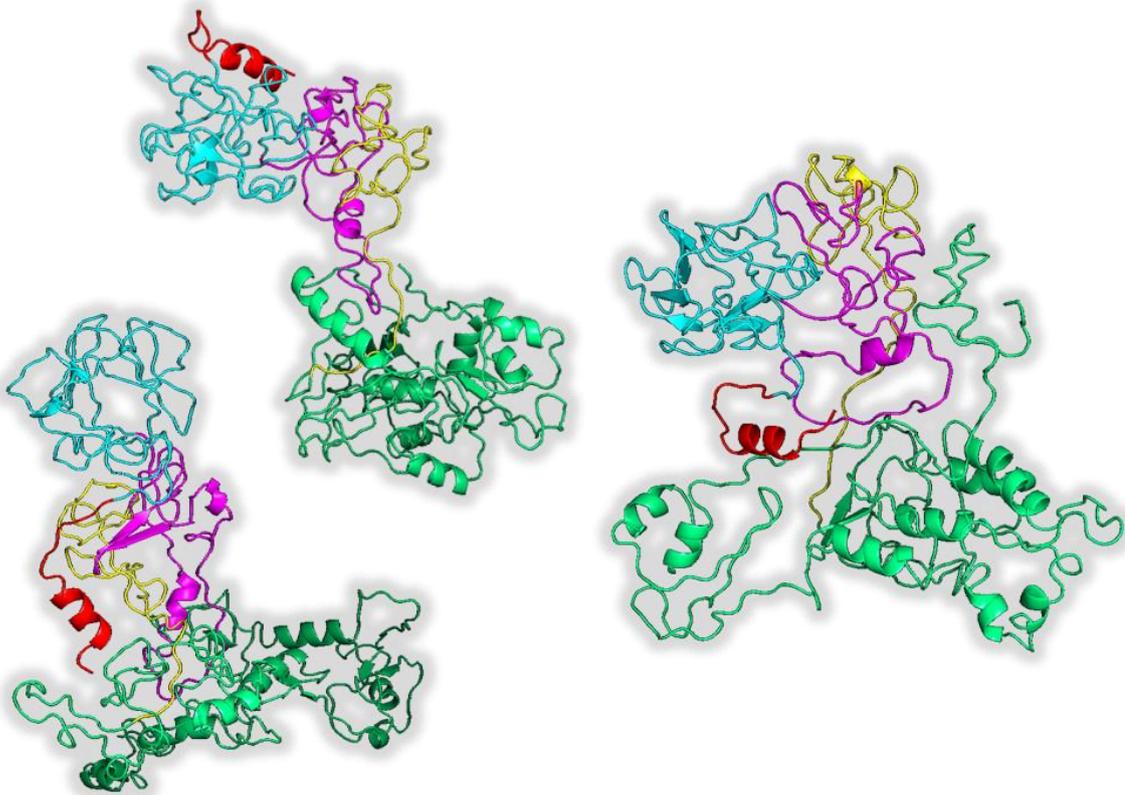


Figura 18. Modelos tridimensionais da proteína LepA388 gerados em análises independentes pelo servidor Phyre2. Representações modificadas no software PyMol para a delimitação dos possíveis domínios da proteína: peptídeo sinal (em vermelho), domínio de lectina B (em azul ciano), sequência identificada por *phage display* (em rosa), região C-terminal do possível “domínio 1” (porção N-terminal da proteína, em amarelo) e a porção C-terminal da proteína (em verde). O índice de confiança para cada modelo é semelhante ao apresentado na Figura 16. Apesar da diferença destes modelos com o primeiro modelo gerado, é possível observar a separação entre os domínios N- e C-terminais.

4.3.2.2. Modelo a partir dos programas SWISS-Model, HHpred e I-TASSER

A proteína LepA388 também foi analisada por outros programas que detectam homologies entre proteínas para a predição da estrutura de proteínas de interesse: SWISS-Model e HHpred. Ambos os programas concordaram no alinhamento estrutural de aproximadamente 40% da proteína LepA388 com os modelos PDB 4tlw e 4tlv, correspondentes à toxina CARDS (*Community-Acquired Respiratory Distress Syndrome*) de *Mycoplasma pneumoniae* (Figura 19 e Tabelas 5 e 6).

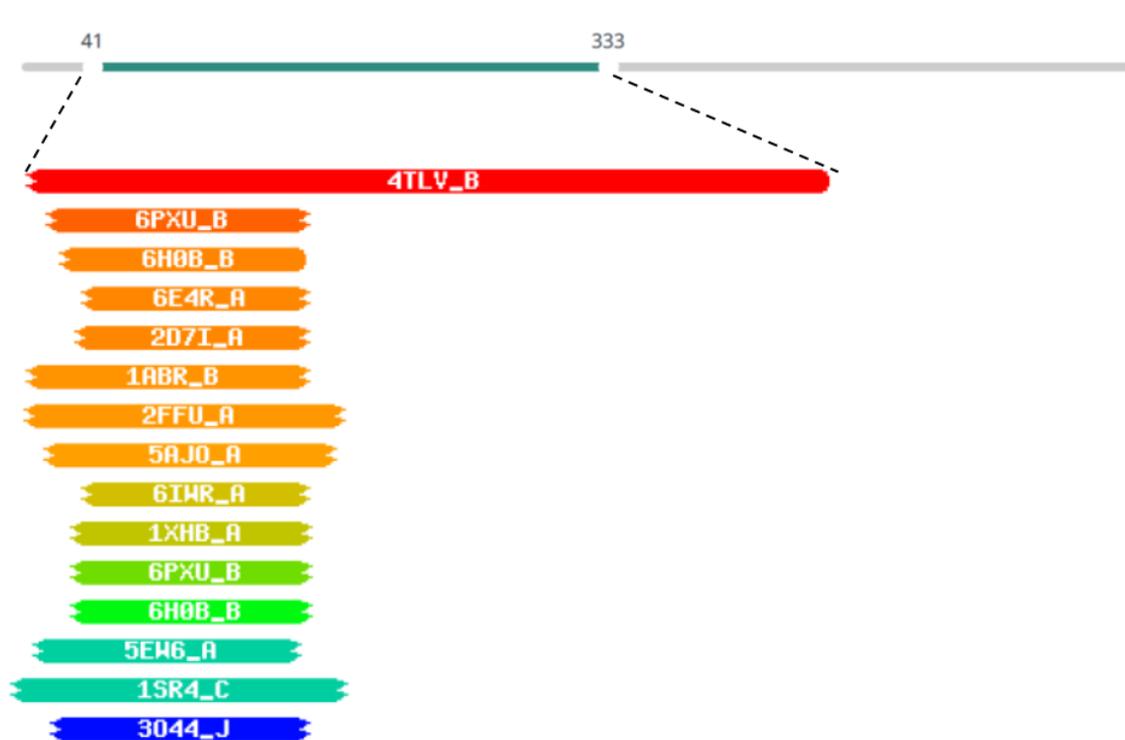


Figura 19. Representação gráfica da predição estrutural e detecção de homologia realizada pelo servidor HHpred (<https://toolkit.tuebingen.mpg.de/tools/hhpred>), a partir da sequência proteica da proteína LepA388. Na figura, o modelo PDB 4tlv_B da toxina CARDS aparece em destaque, e é alinhado à porção N-terminal da proteína de leptospira. Detalhes do alinhamento de modelos à proteína LepA388 se encontram na Tabela 5.

O alinhamento estrutural entre as proteínas LepA388 e CARDS ocorre entre a porção N-terminal de LepA388 e toda a extensão dos domínios D2 e D3 da toxina CARDS (Figura 20A). Os domínios D2 e D3 da toxina CARDS são estruturados em três folhas β , similares a

domínios da toxina ricina (RUTENBER & ROBERTUS, 1991). Além da ligação e internalização da toxina em células, os domínios também são responsáveis pela atividade de vacuolização mediado pela toxina. Dados na literatura mostram que o domínio D3 é especialmente necessário para a plena atividade da toxina, possivelmente mediada por um “caminho aromático” formado dentro do domínio: dos 152 resíduos de aminoácidos, 35 são aromáticos (23%) e 15 deles fazem parte do “caminho aromático” (KANNAN et al., 2014; BECKER et al., 2015). A sequência de LepA388 alinhada ao domínio D3 da toxina CARDS apresenta 25 resíduos aromáticos (17% dos 146 aminoácidos totais), dos quais 14 deles se alinham a aminoácidos aromáticos da toxina CARDS e 7 deles coincidem com os aminoácidos pertencentes ao “caminho aromático” (Figura 20B).

Tabela 5

Descrição dos modelos alinhados à proteína LepA388, a partir da análise de homologia e predição de estruturas por comparação HMM-HMM, realizada pelo programa HHpred.

Nº	PDB	Descrição	Espécie	Probabilidade	Valor E
1	4TLV_B	Toxina de ADP-ribosilação CARDS	<i>Mycoplasma pneumoniae</i>	100	6,70E-66
2	6PXU_B	Polipeptídeo N-acetilgalactosaminiltransferase 12	<i>Homo sapiens</i>	91,77	0,42
3	6H0B_B	Polipeptídeo N-acetilgalactosaminiltransferase 4	<i>Homo sapiens</i>	88,82	1,6
4	6E4R_A	Polipeptídeo N-acetilgalactosaminiltransferase 9	<i>Drosophila melanogaster</i>	88,65	1,2
5	2D7I_A	Polipeptídeo N-acetilgalactosaminiltransferase 10	<i>Homo sapiens</i>	88,59	1,2
6	1ABR_B	Abrina	<i>Abrus precatorius</i>	87,59	2,4
7	2FFU_A	Polipeptídeo N-acetilgalactosaminiltransferase 2	<i>Homo sapiens</i>	87,15	3,9
8	5AJO_A	Polipeptídeo N-acetilgalactosaminiltransferase 2	<i>Homo sapiens</i>	86,54	2,9
9	6IWR_A	N-acetilgalactosaminiltransferase 7	<i>Homo sapiens</i>	82,39	3,2
10	1XHB_A	Polipeptídeo N-acetilgalactosaminiltransferase 10	<i>Mus musculus</i>	81,25	5,5
11	6PXU_B	Polipeptídeo N-acetilgalactosaminiltransferase 12	<i>Homo sapiens</i>	76,58	5,1
12	6H0B_B	Polipeptídeo N-acetilgalactosaminiltransferase 4	<i>Homo sapiens</i>	69,01	19
13	5EW6_A	Receptor de manose tipo C 2	<i>Homo sapiens</i>	60,62	28

Tabela 5 - Continuação

Nº	PDB	Descrição	Espécie	Probabilidade	Valor E
14	1SRA_C	Toxina distensiva citoletal, subunidade A	<i>Haemophilus ducreyi</i>	60,54	120
15	3O44_J	Hemolisina	<i>Vibrio cholerae</i>	40,85	89
16	5EHA_A	Proteína de dobra semelhante à lectina	<i>Agaricus bisporus</i>	38,88	61
17	6INU_A	Receptor de manose de macrófago 1	<i>Homo sapiens</i>	37,77	140
18	5XTS_A	Receptor de manose de macrófago 1	<i>Homo sapiens</i>	36,3	270
19	4TLV_B	Toxina de ADP-ribosilação CARDS	<i>Mycoplasma pneumoniae</i>	32,72	220
20	1ABR_B	Abrina	<i>Abrus precatorius</i>	28,89	460

Tabela 6

Relação dos principais modelos estruturais alinhados à proteína LepA388 pelo programa SWISS-Model

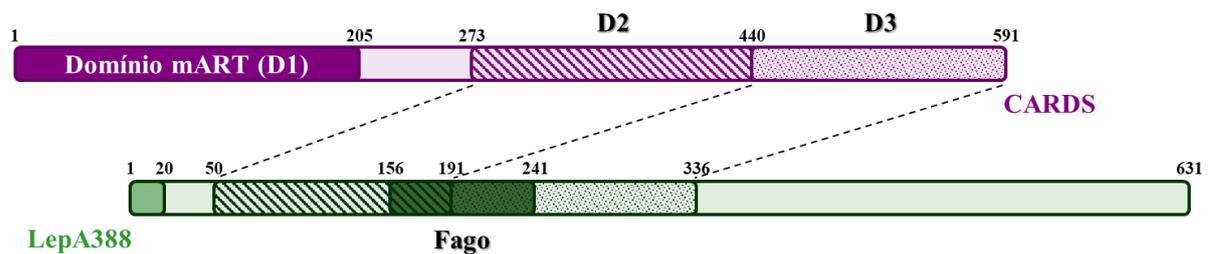
Nº	PDB	Descrição	Espécie	Cobertura	GMQE*	Identidade (%)
1	4TLW.1.A	Toxina de ADP-ribosilação CARDS	<i>Mycoplasma pneumoniae</i>		0,14	19,58
2	4TLV.1.A	Toxina de ADP-ribosilação CARDS	<i>Mycoplasma pneumoniae</i>		0,13	19,17
3	4TLV.2.A	Toxina de ADP-ribosilação CARDS	<i>Mycoplasma pneumoniae</i>		0,13	19,17
4	1DQO.1.A	Receptor de manose	<i>Mus musculus</i>		0,04	19,75
5	1DQG.1.A	Receptor de manose	<i>Mus musculus</i>		0,04	19,75
6	1FWU.1.A	Domínio rico em cisteínas do receptor de manose	<i>Mus musculus</i>		0,04	18,52
7	1HWM.1.B	Ebulina	<i>Sambucus ebulus</i>		0,03	20,27
8	2RG9.1.B	Lectina 1 específica de beta-galactosídeo, cadeia B	<i>Viscum album</i>		0,03	17,07
9	2VLC.1.A	Proteína cinamonina III de inativação de ribossomo tipo 2	<i>Cinnamomum camphora</i>		0,03	20,27
10	4ZA3.1.B	N-glicosidase de rRNA	<i>Momordica charantia</i>		0,03	17,33
11	1ABR.1.B	Abrina	<i>Abrus precatorius</i>		0,03	17,57
12	2AAI.1.B	Ricina, cadeia B	<i>Ricinus communis</i>		0,02	18,92
13	4JKX.1.B	Lectina 1 específica de beta-galactosídeo, cadeia B	<i>Viscum album</i>		0,02	16,22
14	6ELY.1.B	Lectina I do visco	<i>Viscum album</i>		0,02	16,22
15	3D7W.1.B	Lectina 1 específica de beta-galactosídeo	<i>Viscum album</i>		0,02	16,22
16	1PUM.1.B	Lectina I, cadeia B	<i>Viscum album</i>		0,02	16,44
17	1ONK.1.B	Lectina 1 específica de galactose, cadeia B	<i>Viscum album</i>		0,02	16,22
18	1M2T.1.B	Lectina I do visco, cadeia B	<i>Viscum album</i>		0,02	17,81

Tabela 6 - Continuação

N°	PDB	Descrição	Espécie	Cobertura	GMQE*	Identidade (%)
19	1GGP.1.B	Proteína lectina 1, cadeia B	<i>Trichosanthes kirilowii</i>		0,02	13,7
20	2MLL.1.B	Proteína de inativação de ribossomo tipo II	<i>Viscum album</i>		0,02	15,49

*GMQE = Estimativa de Qualidade do Modelo Global. A pontuação GMQE resultante é expressa como um número entre 0 e 1, refletindo a precisão esperada de um modelo construído com esse alinhamento e modelo, normalizado pela cobertura da sequência alvo. Números mais altos indicam maior confiabilidade.

A)



B)

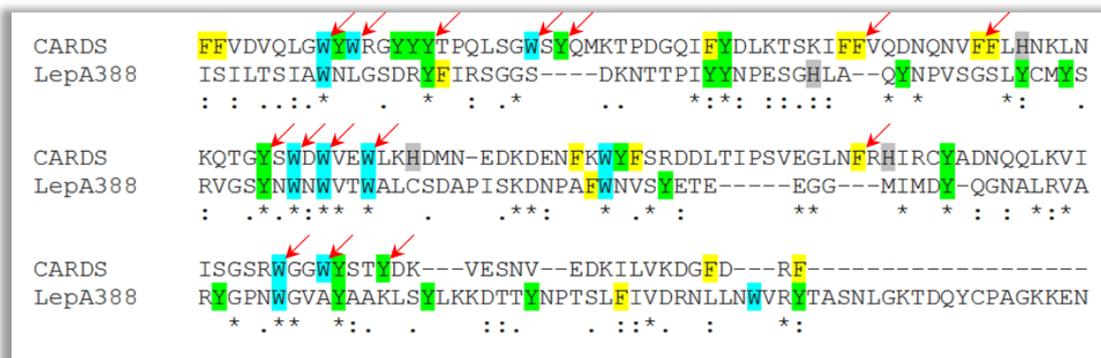


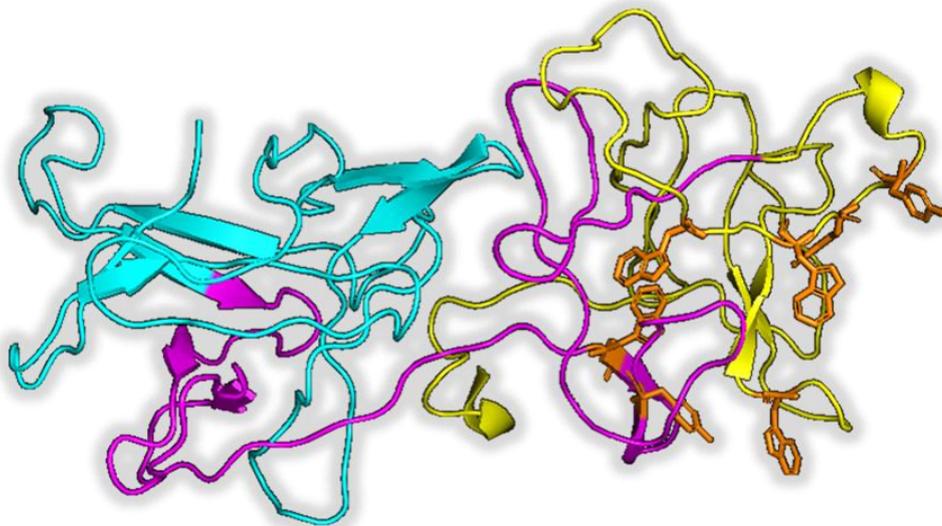
Figura 20. A) Esquema da toxina CARDS de *Mycoplasma pneumoniae* (em roxo) e seus domínios D2 e D3 alinhados ao esquema da proteína LepA388 (em verde). B) Detalhe do alinhamento do domínio D3 da toxina CARDS com a sequência de 191-336 aa de LepA388, onde estão destacados os resíduos aromáticos. As setas vermelhas indicam os resíduos que formam o “caminho aromático”, essencial para a função de ligação e internalização da toxina CARDS em células.

A partir do alinhamento obtido entre LepA388 e a estrutura PDB 4tlw correspondente à toxina CARDS no programa SWISS-Model, foi possível a geração de um modelo estrutural da proteína de LepA388 pelo servidor de predição de estruturas proteicas I-TASSER. Na Figura 21A, está representado o modelo da porção de LepA388 alinhada à toxina CARDS, onde estão destacados os resíduos coincidentes ao “caminho aromático”. A sobreposição deste modelo

com a toxina CARDS demonstra a semelhança entre as duas estruturas (Figura 21B). Esta sobreposição também foi analisada pelo software UCSF Chimera com a ferramenta *MatchMaker*, sendo encontrado o RMSD de 0,772 Å, confirmando a grande semelhança entre as estruturas analisadas.

O modelo completo da proteína LepA388 gerado pelo servidor I-TASSER possui a porção C-terminal da proteína representado como um único domínio mais compacto (Figura 21C). Esta representação se assemelha aos modelos gerados pelo servidor Phyre2 (Figuras 16 e 18) e reforça uma possível separação entre os domínios N- e C-terminais. Adicionalmente, o alinhamento da extremidade N-terminal da proteína LepA388 a domínios de ligação a carboidratos de diversas proteínas visto no servidor Phyre2 também é representado nos programas SWISS-Model e HHpred, incluindo o modelo PDB 1abr referente à toxina abrina e ao seu domínio semelhante à lectina B da ricina.

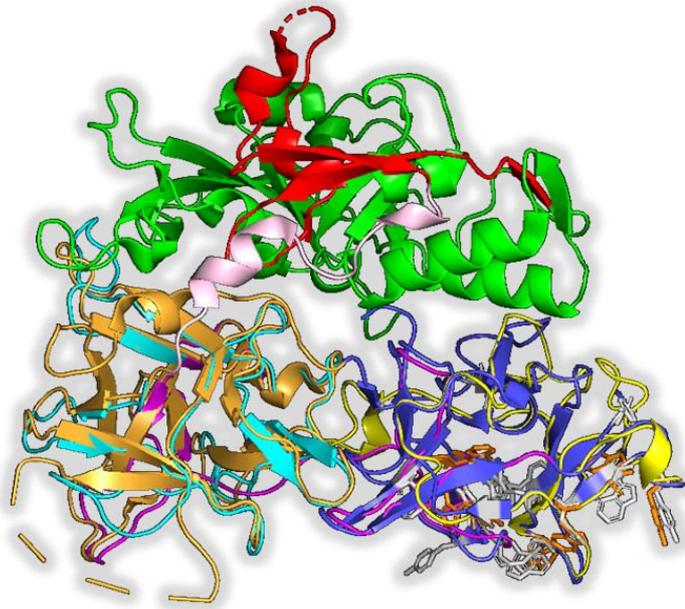
A)



LepA388

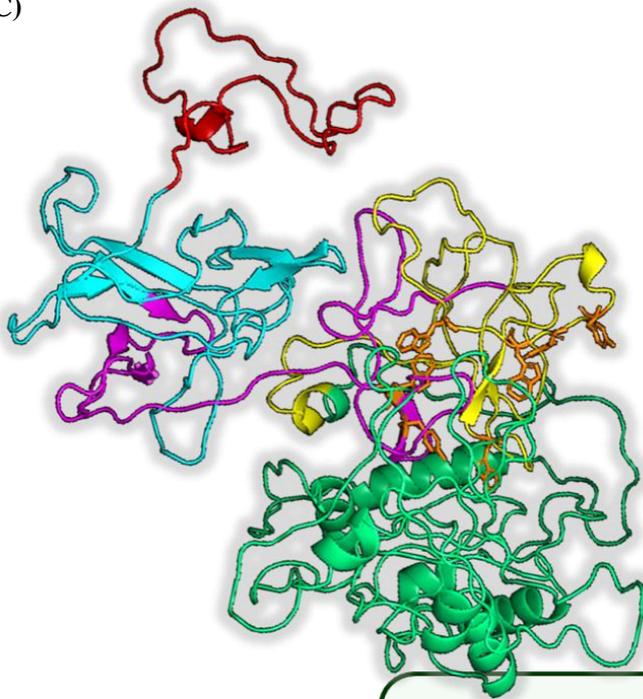
- | | |
|---------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|
|  Alinhamento com D2 (CARDS) |  Alinhamento com D3 (CARDS) |
|  Sequência apresentada pelo fago |  Resíduos do “caminho aromático” |

B)



■ Domínio mART (D1)	■ Domínio 2	Toxina CARDS
■ Sítio de ligação ao NAD ⁺	■ Domínio 3	
■ Conexão entre D1 a D2+D3	■ Resíduos do “caminho aromático”	

C)



LepA388	
■ Alinhamento com D2 (CARDS)	■ Alinhamento com D3 (CARDS)
■ Sequência apresentada pelo fago	■ Resíduos do “caminho aromático”
■ Peptídeo sinal	■ C-terminal

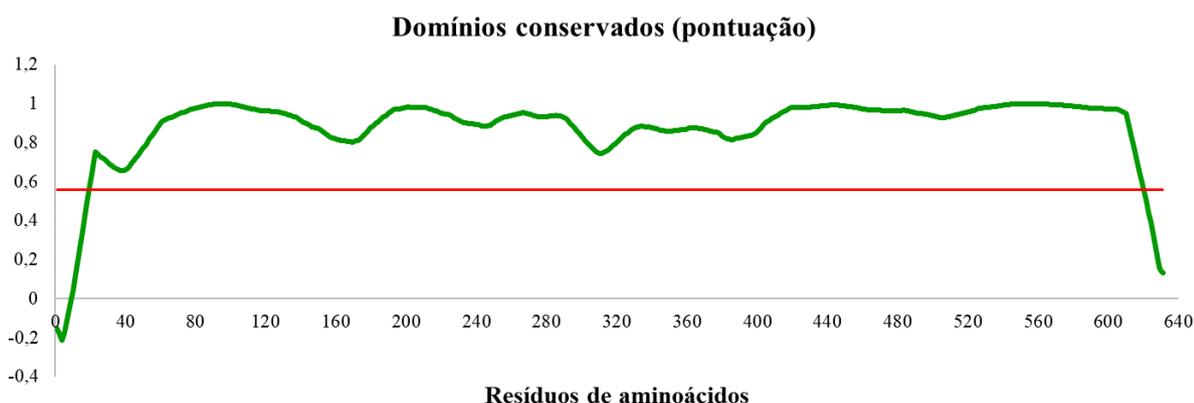
Figura 21. A) Modelo com a sequência de 50-336 aa da proteína LepA388 alinhados aos domínios D2 (em azul ciano) e D3 (em amarelo) da toxina CARDS de *Mycoplasma pneumoniae*. A sequência

identificada por *phage display* está representada em rosa, e os resíduos do “caminho aromático” estão representados em laranja. **B)** Modelo da toxina CARDS de *M. pneumoniae*, alinhado ao modelo da proteína LepA388 representado em A. **C)** Modelo da proteína LepA388 completa gerado pelo servidor I-TASSER a partir do alinhamento com a toxina CARDS (gerado pelo programa SWISS-MODEL). As regiões do peptídeo sinal e C-terminal da proteína LepA388 estão representadas em vermelho e verde, respectivamente.

4.3.2.3. Predição de domínios conservados na estrutura tridimensional de LepA388

Adicionalmente, foi realizada uma nova predição da estrutura tridimensional da proteína LepA388 com o programa ThreaDom, o qual compara a sequência primária de proteínas com modelos tridimensionais publicados e estabelece uma pontuação de conservação de domínio (DCS = *domain conservation score*) para cada resíduo. Dessa forma, o programa pode prever as extremidades de possíveis domínios contínuos e descontínuos na proteína de interesse (XUE et al., 2013). A predição resultante revelou que a proteína LepA388 seria uma proteína de domínio único, segundo o limiar adotado pelo programa (Figura 22A). Apesar deste resultado, as oscilações apresentadas no gráfico com o DCS de cada resíduo curiosamente coincidem com as características marcantes apresentadas pelos modelos apresentados nos itens 4.4.2.1 e 4.4.2.2, principalmente com os modelos gerados pelo servidor Phyre2 (Figura 22B), indicando e reforçando a existência de possíveis divisões espaciais na estrutura tridimensional de LepA388.

A)



B)

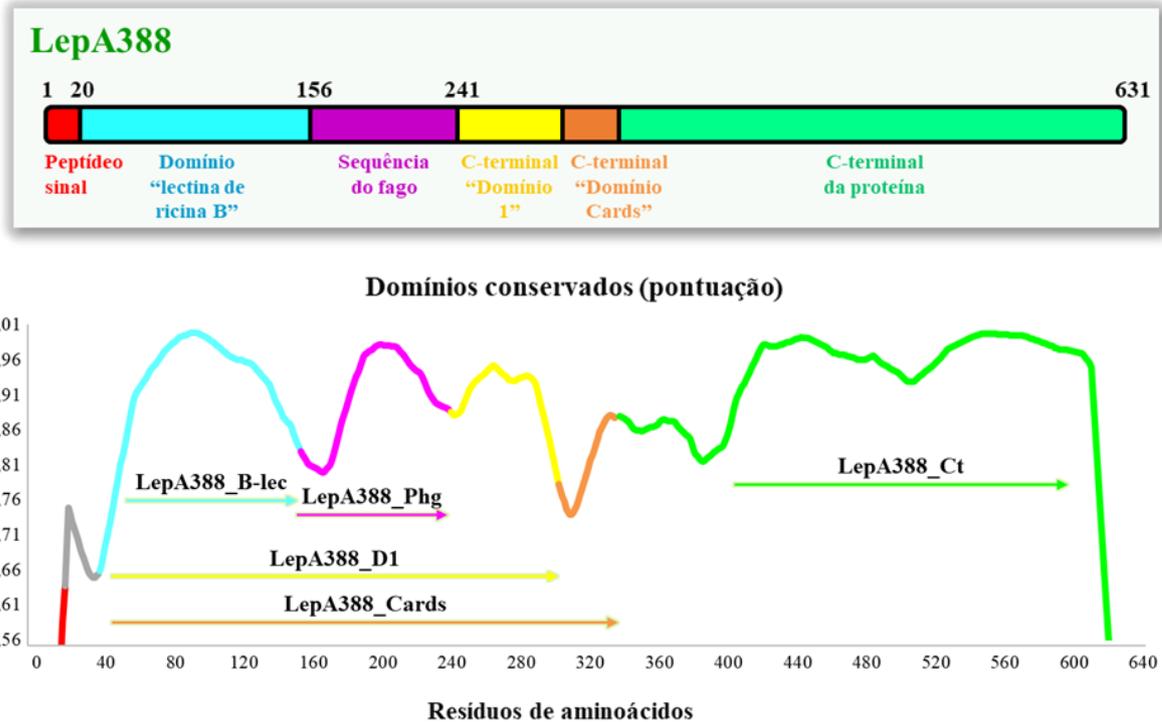


Figura 22. A) Gráfico de predição de domínios conservados gerado pelo programa ThreaDom. A linha vermelha representa o limiar do programa para delimitação de possíveis domínios; no caso da proteína LepA388, apenas um único domínio foi predito pelo programa. B) Gráfico apresentado em A modificado para ampliar a curva acima do limiar definido pelo programa ThreaDom. A curva resultante coincide com porções previstas pelos servidores de predição Phyre2, SWISS-MODEL e HHpred, principalmente na região N-terminal da proteína LepA388. No gráfico também estão representados os esquemas das proteínas recombinantes desenhadas a partir de LepA388.

4.3.3. Desenho, produção e purificação de proteínas recombinantes da proteína

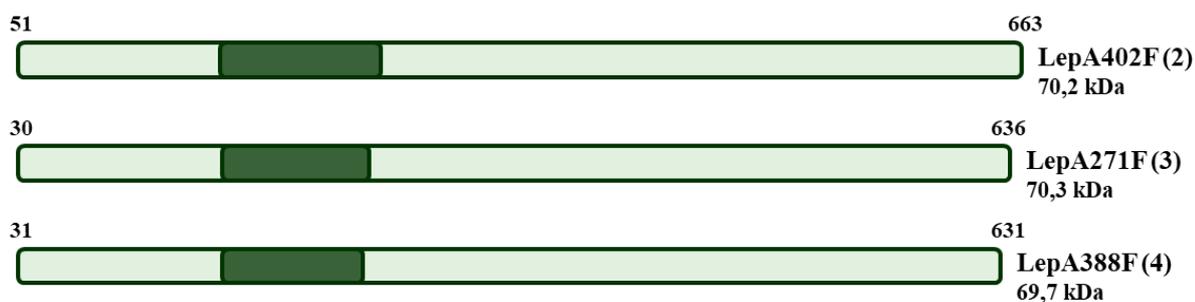
LepA388

A proteína recombinante correspondente à sequência completa da proteína LepA388 sem a região contendo o peptídeo sinal (LepA388F) já havia sido desenhada, amplificada e subclonada no vetor de expressão pAE durante o período do Mestrado, assim como as proteínas recombinantes LepA271F e LepA400F, as quais representam proteínas de maior e menor identidade com a sequência consenso resultante do alinhamento múltiplo das 12 proteínas da família DUF_61, respectivamente (FERREIRA, 2015) (Figura 23A). Além dessas construções, a partir dos modelos gerados com o auxílio de programas de predição, foi definido o desenho

de novas proteínas recombinantes da proteína LepA388, a fim de confirmar possíveis domínios e suas funções associadas: LepA388_Cards, correspondente à sequência alinhada aos domínios D2 e D3 da toxina CARDS (predição dos programas SWISS-Model e HHpred); LepA388_D1, correspondente ao possível domínio N-terminal da proteína LepA388 (predição do servidor Phyre2); LepA388_LecB, correspondente ao domínio semelhante à lectina B de ricina; LepA388_Phg, correspondente à sequência apresentada pelo fago selecionado após seleção em células (técnicas *phage display* e BRASIL); LepA388_Ct, correspondente a uma sequência C-terminal da proteína LepA388 (Figura 23B).

As proteínas recombinantes LepA388F, LepA271F e LepA402F foram produzidas em bactérias *E. coli* BL21 Star (DE3) pLysS, enquanto que as proteínas LepA388_Cards, LepA388_D1, LepA388_LecB, LepA388_Phg e LepA388_Ct foram produzidas em bactérias *E. coli* BL21-SI. Todas as proteínas foram obtidas de forma insolúvel, sendo necessária a dissolução das proteínas em corpúsculos de inclusão com solução contendo 8 M de ureia ou 6 M de hidrocloreto de guanidina. Em seguida, cada proteína foi purificada em coluna cromatográfica de afinidade a níquel (coluna HisTrap), sendo eluída em tampão contendo 8 M de ureia. A produção e purificação das proteínas recombinantes foram avaliadas por SDS-PAGE, em que cada proteína recombinante apresentou massa molecular próximo ao esperado pelas predições e que sua purificação foi bem sucedida (Figura 23C).

A)



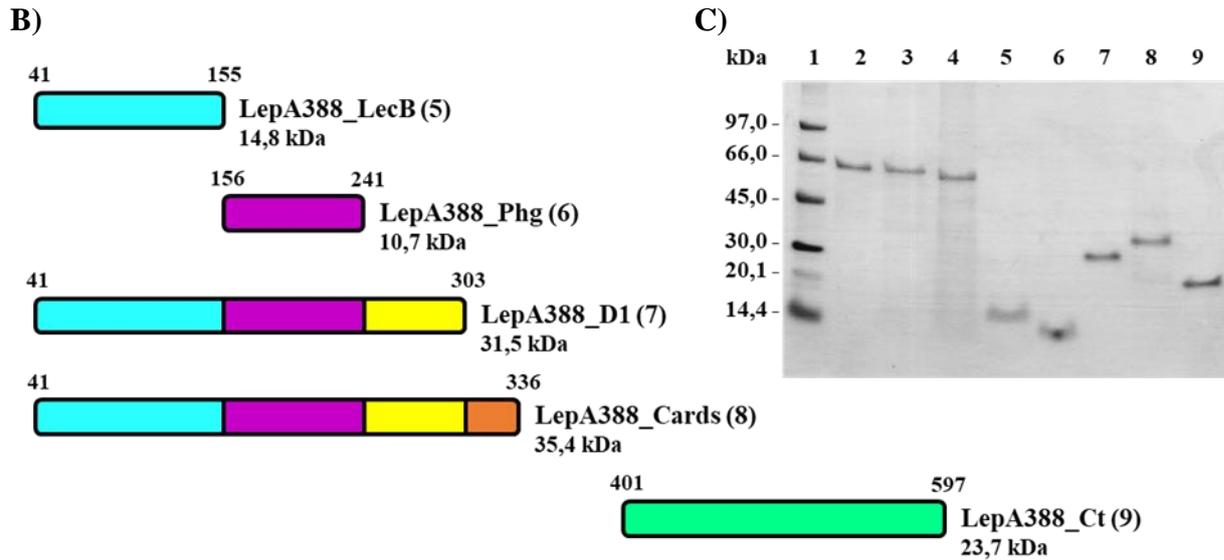


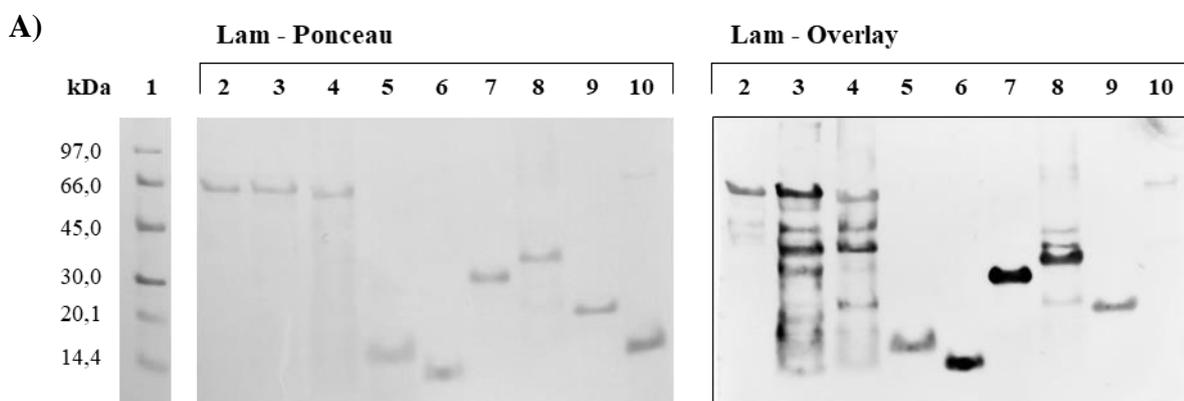
Figura 23. A) Esquema das três proteínas recombinantes correspondentes às sequências das proteínas LepA402, LepA271 e LepA388 (LepA402F, LepA271F e LepA388F, respectivamente). As três proteínas foram desenhadas e clonadas em pAE previamente a este estudo (FERREIRA, 2015). B) Esquema das cinco proteínas recombinantes de LepA388, desenhadas a partir dos modelos tridimensionais gerados pelos programas de predição e identificadas pelos seguintes sufixos: *_LecB*, referente ao domínio de lectina B; *_Phg*, referente à sequência identificada por *phage display*; *_D1*, referente ao domínio N-terminal delimitado pelo servidor Phyre2; *_Cards*, referente à sequência alinhada à toxina CARDS de *M. pneumoniae*; *_Ct*, referente à parte da região C-terminal da proteína. C) SDS-PAGE com as proteínas recombinantes purificadas por IMAC em 8 M ureia. A ordem das proteínas das canaletas 2 a 9 é a mesma da sequência dos esquemas em A e B (números em parêntesis). Canaleta 1: marcador de massa molecular. SDS-PAGE com gel de poliacrilamida de 9% e 16% corado com Coomassie Blue.

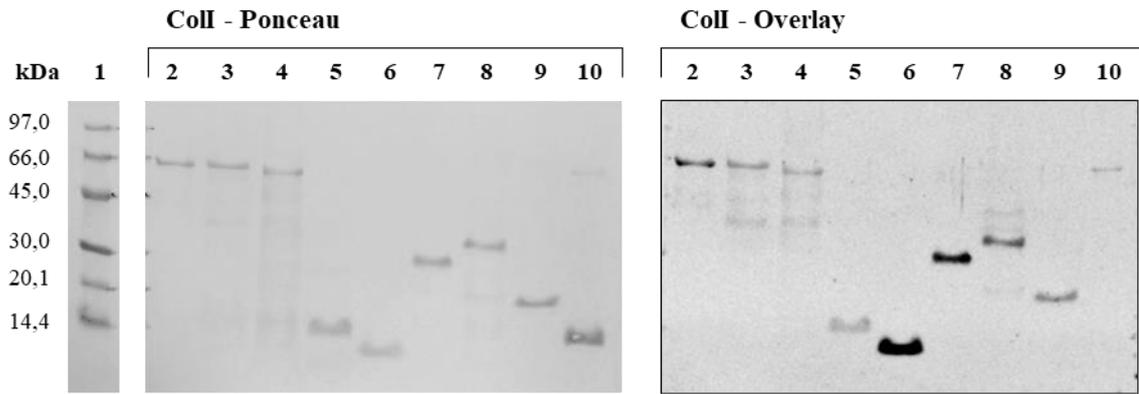
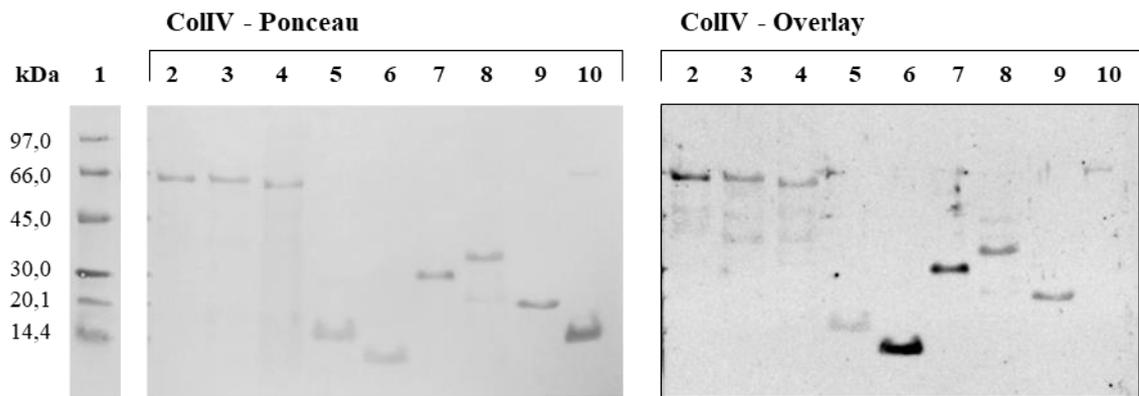
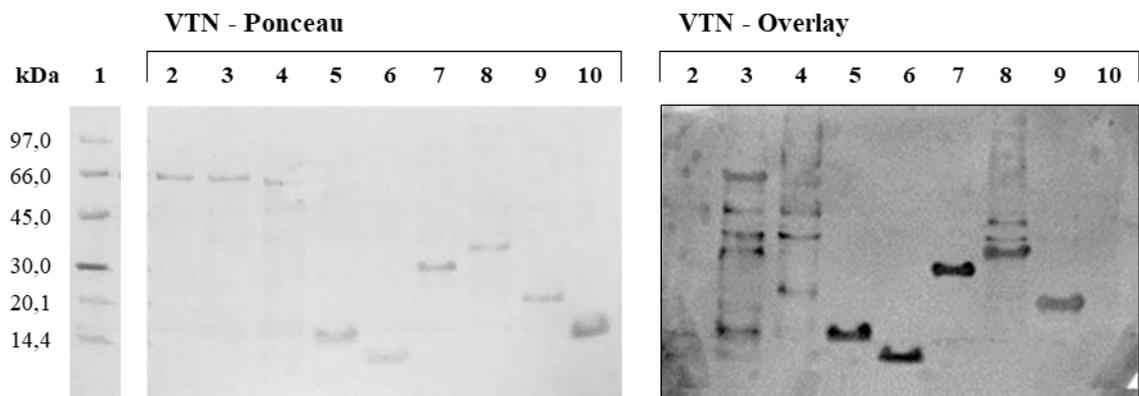
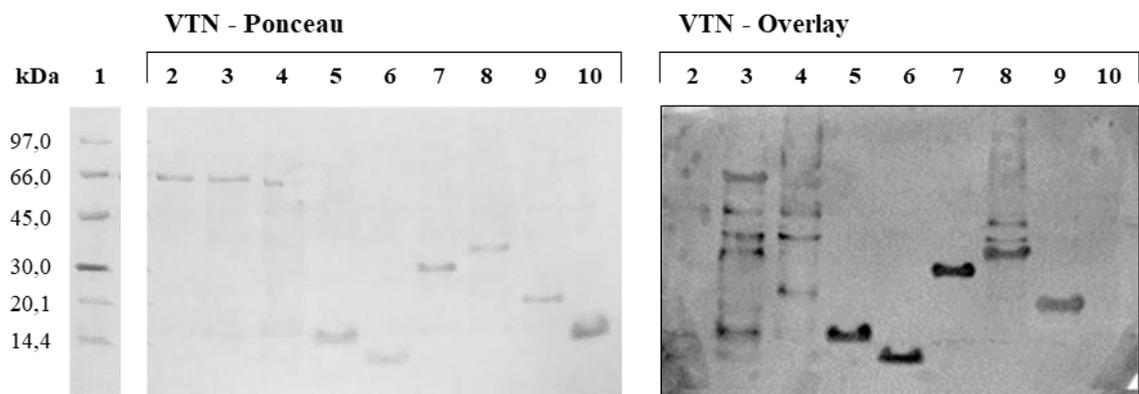
4.3.4. Análise da interação entre as proteínas recombinantes de LepA388 com proteínas de MEC

A fim de avaliar a interação entre as proteínas recombinantes da proteína LepA388 ou da família DUF_61 com proteínas de MEC, foram realizados ensaios de *Far Western Blotting*, também conhecidos como *Overlay*. Amostras das proteínas recombinantes foram aplicadas em géis contendo as concentrações 9% e 16% de poliacrilamida e, posteriormente, as membranas resultantes foram incubadas com as seguintes proteínas de MEC: laminina (Lam), colágeno I (ColI), colágeno IV (ColIV), vitronectina (VTN), fibronectina (FN) celular, fibronectina (FN) plasmática (450 kDa) e os fragmentos de 45 kDa e 30 kDa da fibronectina (FN) plasmática. As

proteínas recombinantes de LigA e de LIC10301 também foram aplicadas em todos os géis como controles positivo e negativo, respectivamente.

Com exceção do controle negativo LIC10301, todas as proteínas recombinantes apresentaram interação em níveis variados com uma ou mais proteínas de MEC utilizadas (Figura 24), mas é possível observar maior interação com as proteínas recombinantes localizadas na região N-terminal de LepA388. Em quase todas as membranas avaliadas, a proteína LepA388_Phg aparece em destaque, indicando sua interação com várias proteínas da MEC. A proteína LepA388_D1 também apresentou interação marcante com laminina, colágeno I, vitronectina e com as fibronectinas celular e plasmáticas. A proteína LepA388_Cards interagiu de forma ligeiramente menos marcante com laminina e as fibronectinas celulares e plasmáticas, enquanto que o fragmento LepA388_LecB interagiu da mesma forma com laminina, vitronectina e fibronectina celular. Quanto às proteínas recombinantes correspondentes às sequências completas de membros da família DUF_61, sua interação com as proteínas de MEC foi mais discreta em comparação aos fragmentos de LepA388, com exceção da marcante interação de LepA271F e LepA388F com laminina. Estes resultados sugerem que a sequência apresentada pelo fago pode ser a principal região onde ocorre a interação desta proteína com o hospedeiro.



B)**C)****D)****E)**

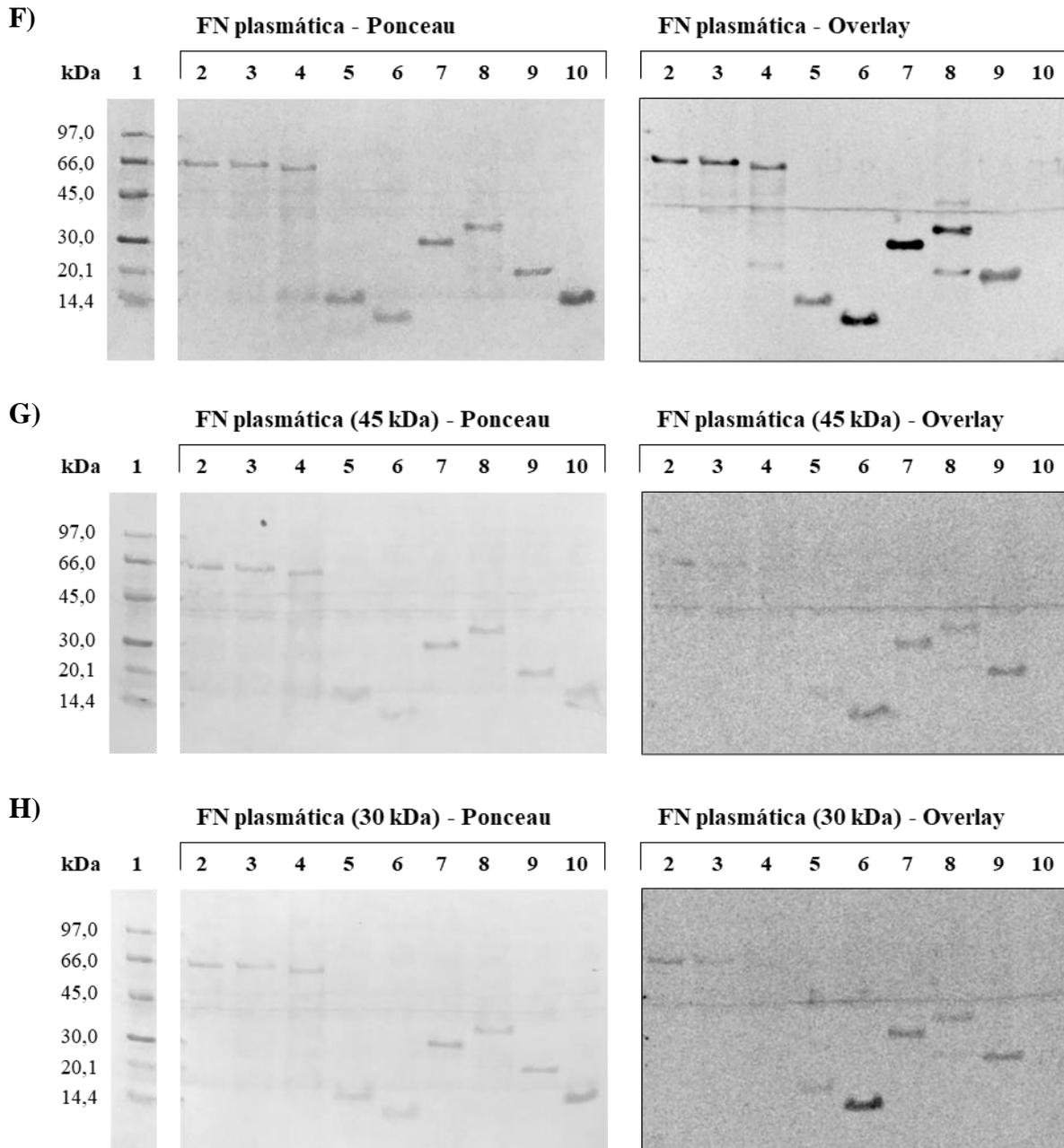


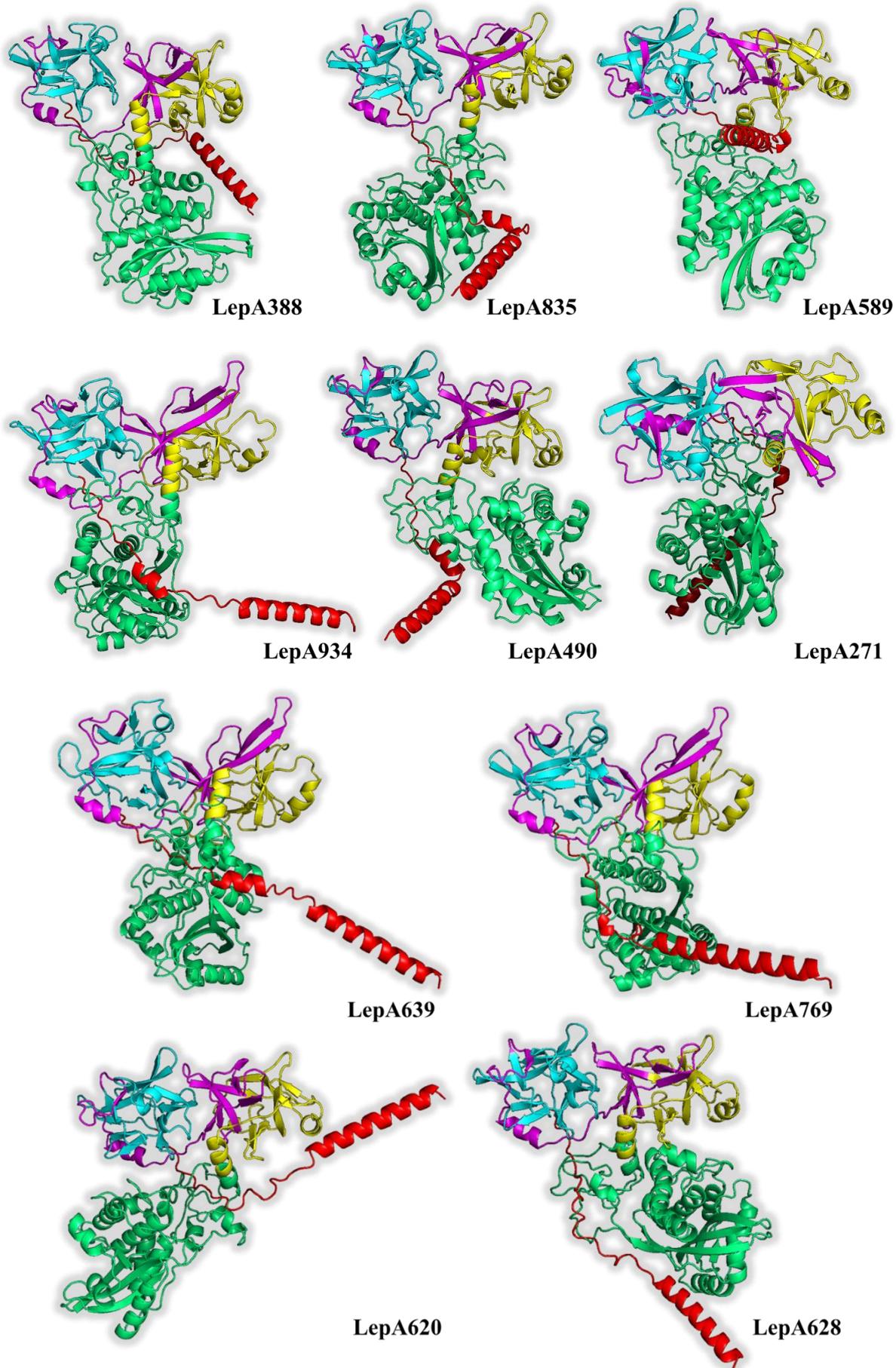
Figura 24. Ensaio de *Far Western Blotting* (*Overlay*) com as proteínas recombinantes de LepA388 e da família DUF_61, e sua interação com proteínas de MEC. À esquerda, estão os registros das membranas de nitrocelulose coradas com Ponceau S, obtidas a partir da transferência de SDS-PAGE com géis de poliácridamida de 9% e 16%, com as proteínas na seguinte ordem: LepA402F (canaleta 2), LepA271F (canaleta 3), LepA388F (canaleta 4), LepA388_LecB (canaleta 5), LepA388_Phg (canaleta 6), LepA388_D1 (canaleta 7), LepA388_Cards (canaleta 8), LepA388_Ct (canaleta 9), e proteínas recombinantes de controle positivo LigA e negativo LIC10301 (canaleta 10). À direita, estão as revelações por quimioluminescência após incubação das membranas de nitrocelulose correspondentes com 10 µg/mL das proteínas: laminina (Lam, **A**), colágeno I (ColI, **B**), colágeno IV (ColIV, **C**), vitronectina (VTN, **D**), fibronectina celular (FN celular, **E**), fibronectina plasmática (FN plasmática, **F**) ou fragmentos proteolíticos de 30 kDa (**G**) e 45 kDa (**H**) da fibronectina plasmática, seguida pela detecção com anticorpos específicos. Canaleta 1: marcador de massa molecular.

4.3.5. Novas modelagens de LepA388 e de todos os membros da família DUF_61

Uma das mais recentes metodologias de predição de estruturas proteicas baseadas em *deep learning*, RoseTTAFold, permite a geração de modelos de alta acurácia a partir de uma rede neural de três vias capaz de processar simultaneamente a sequência de aminoácidos, distância entre resíduos e informações de suas coordenadas (BAEK et al., 2021). Esta metodologia foi utilizada para validar os modelos gerados e descritos nos itens acima para a proteína LepA388, assim como para todos os outros 12 membros da família DUF_61 presentes em *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130.

A partir da sequência aminoacídica de cada proteína, foram gerados modelos com grande parte dos erros menores do que 5 angstroms e confiança média de 0,74 ($\pm 0,02$) (HIRANUMA et al., 2021), exceto pelos modelos das proteínas LepA402 e LepA591 que apresentaram confiança ligeiramente menor (0,70 e 0,65, respectivamente), as quais possuem sequências de aminoácidos mais distantes do restante dos membros da família DUF_61 nesta cepa de *Leptospira* (Figura 25).

Apesar de algumas pequenas diferenças estruturais, é possível observar a grande semelhança dos 13 modelos entre si, os quais são concordantes com os modelos gerados pelos servidores Phyre2 (Figura 16) e SWISS-Model em combinação com I-TASSER (Figura 21C) para a proteína LepA388: em todos os modelos existe a tendência de separação entre os domínios N- e C-terminais e a semelhança com os domínios de ligação e internalização celulares de toxinas, como CARDS, abrina e ricina (RMSD de 1,1-1,2 Å entre o modelo de LepA388 e os modelos PDB 4tlw, 1abr e 2aai), sugerindo que esta família de proteínas pode desempenhar um papel semelhante durante o ciclo infeccioso das leptospiras.



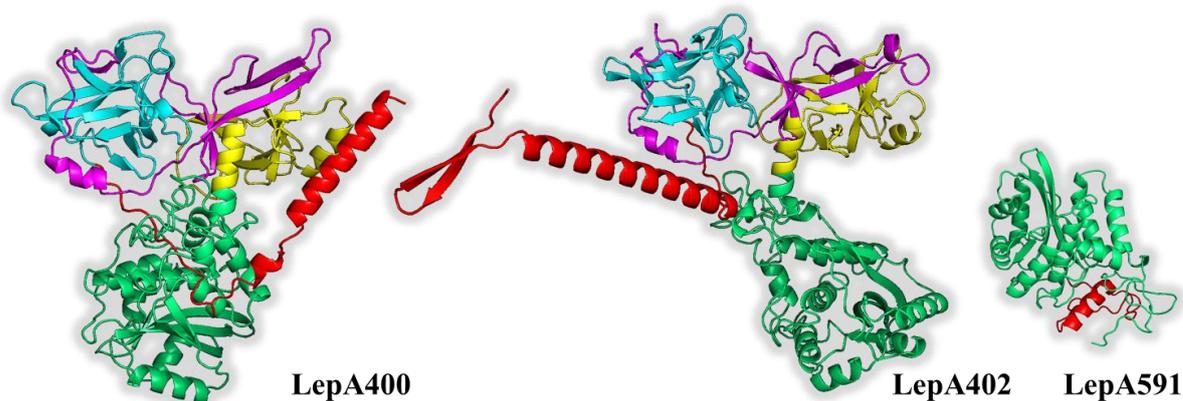


Figura 25. Modelos tridimensionais das proteínas da família DUF_61 gerados pelo servidor Robetta (metodologia RoseTTAFold). Representações modificadas no software PyMol para a delimitação de possíveis domínios da proteína: peptídeo sinal (em vermelho), alinhamento com o domínio D2 da toxina CARDS (em azul ciano), sequência identificada por *phage display* (em rosa), alinhamento com o domínio D3 da toxina CARDS (em amarelo) e a porção C-terminal da proteína (em verde). Modelos com erros menores do que 5 angstroms e índices de confiança variando entre 0,65 a 0,76.

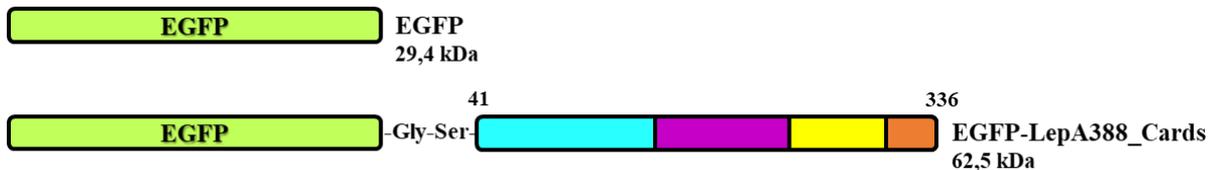
4.3.6. Produção e purificação de proteínas fluorescentes com EGFP

Novas proteínas recombinantes tiveram suas sequências nucleotídicas clonadas para posterior expressão em bactérias: a proteína EGFP e a proteína fusionada EGFP-LepA388_Cards (Figura 26A), com o objetivo de analisar a interação desta porção da proteína LepA388 a células mamíferas e seus receptores. As proteínas recombinantes foram produzidas em bactérias *E. coli* ArcticExpress (DE3) de forma solúvel a partir do vetor de expressão pAZ e, em seguida, cada proteína foi purificada em coluna HisTrap, com posterior diálise para a remoção de imidazol. Além delas, ajustes no cultivo para a produção da proteína LepA388F em bactérias *E. coli* BL21 Star (DE3) pLysS permitiram sua expressão de forma solúvel, a qual também foi purificada em coluna HisTrap.

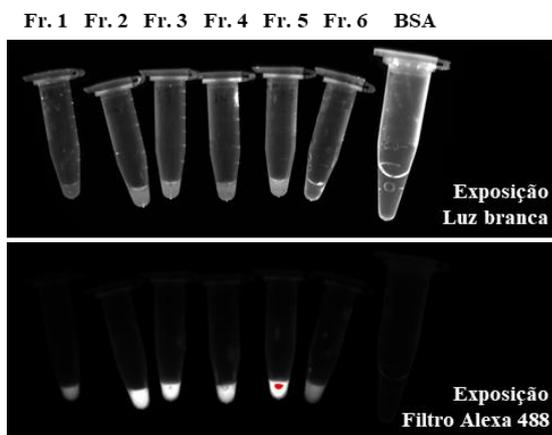
Após a purificação e diálise, amostras da proteína recombinante EGFP estavam claramente esverdeadas, indicando o êxito em sua expressão. Amostras da proteína fusionada EGFP-LepA388_Cards não apresentavam cor verde por avaliação visual, mas grande parte delas apresentou fluorescência sob exposição ao filtro Alexa 488 no fotodocumentador

Chemidoc (Bio-Rad) (Figura 26B). Todas as proteínas recombinantes solúveis foram posteriormente avaliadas por SDS-PAGE (Figura 26C).

A)



B)



C)

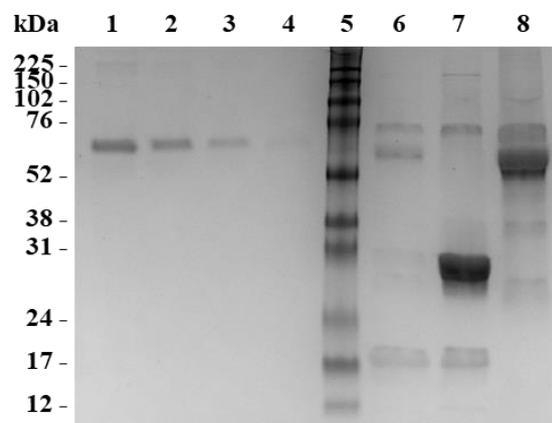


Figura 26. A) Esquema das duas proteínas recombinantes clonadas no vetor de expressão pAZ correspondentes às sequências da proteína EGFP e da proteína fusionada EGFP-LepA388_Cards. A porção da proteína LepA388 nesta fusão apresenta a mesma sequência aminoacídica que a apresentada na Figura 23B. B) Avaliação da fluorescência das amostras purificadas da proteína fusionada EGFP-LepA388_Cards sob exposição ao filtro Alexa 488 no fotodocumentador Chemidoc. As frações 1 a 6 (Fr. 1-6) representam as diferentes coletas de purificação da proteína em coluna HisTrap. Uma amostra contendo BSA também foi avaliada como controle negativo. C) SDS-PAGE com as proteínas recombinantes purificadas por IMAC. À esquerda, foram adicionadas amostras de BSA em concentrações decrescentes para comparação: 2 µg (canaleta 1), 1 µg (canaleta 2), 0,5 µg (canaleta 3) e 0,25 µg (canaleta 4). Amostras da Fr. 5 da proteína fusionada EGFP-LepA388_Cards (com maior fluorescência, ver figura B), de EGFP e LepA388F foram adicionadas nas canaletas 6, 7 e 8, respectivamente. Canaleta 5: marcador de massa molecular. SDS-PAGE com gel de poliacrilamida de 12% corado com Coomassie Blue.

4.4. Análise da lipoproteína hipotética LepA069

Finalmente, LepA069 consiste em uma proteína hipotética que contém 1002 resíduos de aminoácidos e sua análise por programas de predição revelou a presença de uma região transmembrana e sua possível localização na membrana externa de leptospiros. Esta proteína se destacou ao ter duas sequências sobrepostas selecionadas pelo método de *phage display*, seguido da seleção pela técnica BRASIL em células VERO, indicando sua possível interação com o hospedeiro (FERREIRA, 2015). Entretanto, não há informações de possíveis domínios ou estudos na literatura com esta proteína.

Como uma análise inicial, foi realizada uma reação de qPCR com oligonucleotídeos específicos para a proteína LepA069 e cDNA obtido a partir do RNA extraído do fígado de camundongos C3H/HeJ após três dias do desafio com *L. interrogans* sorovar *Copenhageni* cepa Fiocruz L1-130. A expressão do gene de LepA069 obteve um aumento médio de aproximadamente 70 vezes, em comparação à expressão do gene ribossomal 16S (Figura 27).

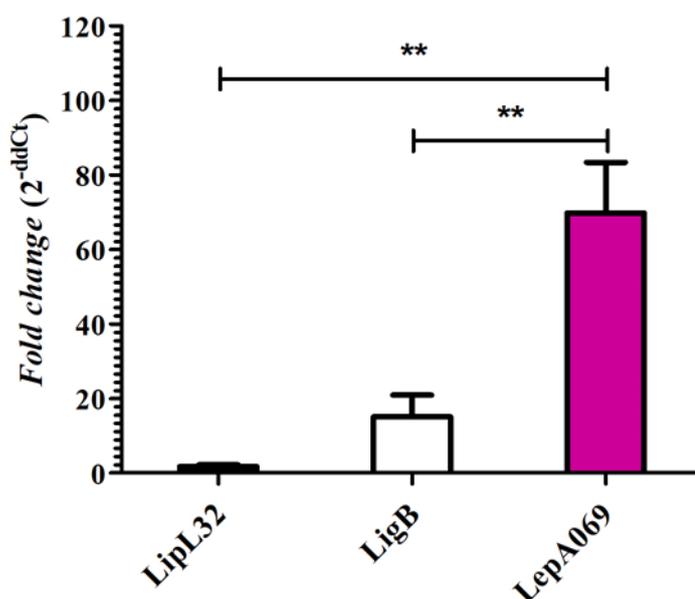


Figura 27. Análise da expressão dos genes codificadores de LipL32, LigB e LepA096, apresentada em *fold change* e pelo método delta-delta (2^{-ddCt}), a partir de RT-qPCR utilizando amostras de fígado de camundongos susceptíveis C3H/HeJ após três dias de infecção por *L. interrogans* sorovar *Copenhageni*

cepa Fiocruz L1-130. As amostras foram analisadas em triplicata para cada par de oligonucleotídeo, e os dados foram normalizados a partir da expressão do gene constitutivo ribossomal 16S e comparadas à expressão dos mesmos genes em leptospiros cultivadas em meio EMJH a 30 °C. Análise estatística por *One-way* ANOVA, com $P < 0,01$ (**).

A análise da proteína LepA069 através do desenho de proteínas recombinantes e produção dos fragmentos resultantes não foi bem sucedida (dados não apresentados). Entretanto, o forte aumento na expressão de seu gene em condições de infecção em animais susceptíveis, aliado aos resultados previamente obtidos através da técnica de *phage display*, sugerem sua participação na patogênese da leptospirose e uma intrínseca relação patógeno-hospedeiro a ser explorada.

5. DISCUSSÃO

A caracterização da bactéria virulenta *Leptospira interrogans* quanto à sua interação com diferentes hospedeiros permanece pouco conhecida, devido à limitação de ferramentas para a manipulação genética das leptospiras. Como alternativa, a técnica de *phage display* tem se mostrado uma ferramenta valiosa para identificar novas adesinas e outras proteínas que participam da interação patógeno-hospedeiro em leptospiras e bactérias espiroquetas (COBURN et al., 1999; ANTONARA et al., 2007; ROSANDER et al., 2011; CHING et al., 2012; LIMA et al., 2013; EVANGELISTA et al., 2014b). Através desta técnica, as proteínas hipotéticas LepA962, LepA069 e LepA388 foram previamente identificadas e sua caracterização foi continuada e aprofundada neste estudo.

Já havia evidências na literatura de que a lipoproteína hipotética LepA962, também conhecida como Lp95, poderia ser um componente candidato a contribuir com o processo infeccioso de leptospiras. No atual estudo, a interação da proteína recombinante LepA962_PhG com componentes de matriz extracelular e células VERO foi claramente demonstrada, corroborando o ensaio funcional que culminou na identificação desta sequência proteica pela técnica de *phage display*, após seleção na mesma linhagem de células epiteliais renais (FERREIRA, 2015). Os resultados obtidos com as proteínas recombinantes LepA962_PhG e LepA962_Nt também estão de acordo com os dados apresentados no trabalho de Atzingen e colaboradores (ATZINGEN et al., 2009), onde foram demonstradas diferenças claras de ligação entre as três porções recombinantes da proteína LepA962 (N-terminal, C-terminal e *Overlap*) com as proteínas de MEC testadas: todos os fragmentos e a proteína inteira mostraram ligação até certo ponto, mas o fragmento C-terminal, o qual possui o domínio DUF_54 e grande parte da sequência identificada por *phage display*, apresentou maior intensidade de ligação à laminina e fibronectinas plasmática e celular. Por outro lado, o fragmento “*Overlap*” que possui a

sequência completa encontrada no fago também apresentou marcante interação com laminina e fibronectina plasmática, e interação em menor intensidade à fibronectina celular. No atual estudo, também foi demonstrada a intensa ligação de LepA962_Phg à laminina e fibronectina plasmática, entretanto, ao contrário do trabalho de Atzingen, também foi observada interação marcante desta proteína recombinante ao colágeno tipo I, e adicionalmente ao fibrinogênio (Figura 11). As interações com as proteínas de MEC testadas neste estudo mostraram ser dependentes da dose de proteína LepA962_Phg, enquanto que a proteína LepA962_Nt apresentou interação significativamente menor com as mesmas proteínas de MEC, mesmo nas maiores concentrações testadas (Figura 12).

Adicionalmente, Atzingen e colaboradores demonstraram que a proteína recombinante inteira de LepA962 é capaz de ativar a molécula de adesão celular E-selectina em células endoteliais HUVEC. Seus antígenos também foram identificados tanto no lúmen tubular renal quanto em associação a infiltrados de células inflamatórias intersticiais, a partir de tecidos renais obtidos de cobaias após 14 dias de infecção por bactérias virulentas de *L. interrogans* cepa Fiocruz L1-130, indicando sua expressão durante a infecção animal e sua interação com um dos órgãos mais afetados na leptospirose (ATZINGEN et al., 2009). Em concordância com os dados na literatura, no atual estudo também foi demonstrada a ligação da proteína recombinante LepA962_Phg a células VERO, uma linhagem de células mamíferas provenientes do epitélio renal de macacos. Apesar desta interação ser claramente mais intensa com a matriz extracelular secretada pelas células, a ligação às células VERO com as duas concentrações testadas foi significativamente maior comparada à proteína LepA962_Nt e ao controle negativo de pneumococo PspA4Pro (Figura 13). Por outro lado, o ensaio de ELISA sanduíche realizado para avaliar a interação de antissoros à superfície de leptospiros virulentas da cepa Fiocruz L1-130 revelou uma discreta ligação dos antissoros das proteínas LepA962_Phg e LepA962_Nt em relação ao antissoro da lipoproteína LipL32, presente em

abundância na superfície de leptospiras (Figura 14). Entretanto, as bactérias utilizadas pertenciam à terceira passagem em cultura com meio EMJH após a coleta de hamsters infectados, condições que podem ser desfavoráveis para a expressão da proteína LepA962.

Curiosamente, fontes na literatura discordam quanto aos dados obtidos com a proteína LepA962 e com as 6 proteínas LenA-F pertencentes à mesma família do domínio DUF_54. Assim como a LepA962, trabalhos demonstraram a ligação de proteínas Len à laminina e à fibronectina plasmática (BARBOSA et al., 2006; STEVENSON et al., 2007), enquanto que no trabalho de Pinne e colaboradores, a interação de LepA962 e das proteínas Len com fibronectina não foi considerada significativa; os próprios autores consideraram as diferenças nas estratégias experimentais e nas origens das fibronectinas utilizadas entre os trabalhos como prováveis causas da incompatibilidade nos resultados obtidos (PINNE; MATSUNAGA; HAAKE, 2012). Adicionalmente, a expressão dos genes codificadores de LenA e LepA962 não foi consistentemente detectada por RT-qPCR em modelos de infecção de animais susceptíveis (hamsters) e resistentes (camundongos) com bactérias virulentas de *L. interrogans* sorovar Icterohaemorrhagiae cepa Verdun (MATSUI et al., 2012), entretanto, o desenho dos oligonucleotídeos baseado na sequência dos genes de *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130 pode ter contribuído para a falha na detecção destes genes.

O papel da proteína LepA962 na contribuição da virulência em leptospiras patogênicas ainda deve ser melhor elucidado, entretanto, o presente estudo confirmou a clara interação desta proteína e um de seus fragmentos (revelados por *phage display*) com células do hospedeiro e proteínas de MEC, sugerindo a participação da LepA962 na disseminação de leptospiras durante seu ciclo infeccioso nos hospedeiros.

Em relação à proteína LepA388, foi inicialmente demonstrado por RT-qPCR o aumento na expressão de seu gene em amostras de fígado de camundongos susceptíveis infectados por *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130. Similarmente, indícios da

participação da proteína LepA388 na virulência de leptospiras já havia sido demonstrada na literatura. A regulação positiva na expressão do seu gene ortólogo por RT-qPCR foi documentada em amostras de sangue e fígado, e em menor intensidade nas amostras de rins coletadas de hamsters em modelo de infecção aguda por bactérias *L. interrogans* sorovar Lai cepa 56601. O mesmo estudo demonstrou o aumento na expressão dos genes de outros membros também pertencentes à família DUF_61, utilizando as mesmas amostras de animais infectados (LEHMANN et al., 2013). Outro trabalho revelou uma forte indução de 6 a 11 vezes na expressão de dois genes da família DUF_61 (LepA402 e LepA400, respectivamente) em cultivos de *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130 com osmolaridade fisiológica atingida com a adição de 120 mM de NaCl ao meio EMJH (MATSUNAGA et al., 2007). Ainda considerando o gene da proteína LepA402, um modelo de infecção em células HK-2 por *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130 mostrou sua forte expressão após 4 h de infecção (CHOU et al., 2014) e, mais recentemente, análises de expressão diferencial utilizando RNA-seq revelaram o mesmo gene como um dos mais significativamente expressos quando as bactérias patogênicas de *L. borgpetersenii* sorovar Hardjo cepa JB197 foram cultivadas a 37 °C (PUTZ et al., 2021).

A análise da proteína LepA388 continuou com a predição de estruturas tridimensionais e possíveis domínios, sendo a porção N-terminal modelada de forma mais confiável, a partir do seu alinhamento com o domínio de β -lectina (ou lectina B) da toxina ricina. Este domínio está presente em diversas proteínas que interagem com carboidratos, como toxinas AB, glicosidases e proteases de plantas e bactérias. O domínio 1abr_B da proteína abrina foi um dos domínios mais utilizados como referência no alinhamento com a LepA388, e a abrina, assim como a ricina, trata-se de uma potente toxina do tipo AB, em que a cadeia A é responsável pela atividade tóxica da proteína, enquanto que a cadeia B consiste em um domínio de ligação à galactose que possibilita a ligação e internalização da toxina nas células (TAHIROV et al.,

1995; BAGARIA & KARANDE, 2014), sendo estruturalmente semelhante à cadeia B da toxina ricina. Atualmente, o domínio de lectina B se encontra anotado somente na proteína LepA400 em *L. interrogans* (cepa Fiocruz L1-130), entretanto, no presente estudo, foi observada a grande semelhança na sequência referente a este domínio na proteína LepA388 e em todas as outras proteínas pertencentes à família DUF_61, incluindo a presença das repetições QxW características deste domínio (RUTENBER; READY; ROBERTUS, 1987; HAZES, 1996), indicando o provável compartilhamento deste domínio entre todos os membros da família DUF_61 (Figura 17A). Curiosamente, no estudo de Lehmann e colaboradores (LEHMANN et al., 2013), a mutação G72E encontrada no gene ortólogo ao codificador da proteína LepA388 após atenuação de bactérias *L. interrogans* em subculturas *in vitro* ocorre no possível domínio de lectina B na proteína, sugerindo a importância desta estrutura na contribuição da virulência de leptospiros patogênicos.

Outros domínios marcantes presentes nas análises de predição da estrutura de LepA388 foram 4tlv e 4tlw, correspondentes à toxina CARDS de *Mycoplasma pneumoniae*. A toxina CARDS é um fator de virulência expresso pela bactéria com atividades de ADP-ribosilação e vacuolização em células das vias respiratórias, levando à morte celular. A estrutura da toxina é triangular, composta pelo domínio N-terminal (D1) com o sítio ativo mART (mono-ADP-ribosiltransferase), seguido por dois domínios (D2 e D3) responsáveis pela ligação e internalização da toxina em células (BECKER et al., 2015). Os domínios D2 e D3 também são estruturados de maneira semelhante ao domínio de lectina B da ricina, e a semelhança entre estes domínios com a porção N-terminal da proteína LepA388, incluindo a conservação parcial do caminho aromático característico do domínio D3 da toxina CARDS, sugerem que a proteína de leptospira pode desempenhar um papel semelhante de ligação e internalização celulares.

A predição de estruturas tridimensionais e domínios auxiliou no desenho de fragmentos da proteína LepA388 para análises funcionais subsequentes. Através da técnica de *Far Western*

Blotting, foi observada interações marcantes principalmente entre os fragmentos correspondentes à sequência identificada por *phage display* (LepA388_Phg) e à região N-terminal completa (LepA388_D1 e LepA388_Cards) com as proteínas de MEC como laminina, colágenos do tipo I e IV, vitronectina e fibronectinas plasmáticas e celular. Estes resultados somados às informações na literatura, às predições estruturais da região N-terminal da proteína LepA388 e ao prévio ensaio de *phage display* seguido pela seleção dos fagos em células VERO, sugerem fortemente que a proteína pode contribuir para a interação patógeno-hospedeiro durante a infecção por leptospiros. É importante ressaltar que este estudo foi o primeiro a explorar de forma inicial as possíveis atividades funcionais de diferentes regiões da proteína LepA388, contribuindo assim para a elucidação desta singular família de genes parálogos DUF_61 em leptospiros virulentas.

A família de proteínas parálogas DUF_61 é uma das milhares de famílias proteicas que carecem de homólogos com estrutura conhecida para comparação de modelagens de forma mais acurada. Entretanto, métodos computacionais utilizados para a predição de estruturas proteicas a partir da sequência de aminoácidos tornaram-se cada vez mais numerosos e acessíveis no final do século XX, facilitando e guiando o estudo funcional de proteínas com funções e estruturas ainda desconhecidas. O crescente aumento no sequenciamento de genomas também possibilita a previsão cada vez mais precisa de contatos resíduo-resíduo utilizando dados evolutivos (ANFINSEN, 1973; OVCHINNIKOV et al., 2017).

Os renomados servidores para modelagem de proteínas utilizados no presente estudo (Phyre2, I-TASSER, SWISS-Model e HHpred) utilizam métodos de predição de estruturas baseadas em modelos conhecidos por homologia e/ou reconhecimento de dobras, gerando modelos com características semelhantes entre si, mas com regiões de baixa confiança estrutural. A análise adicional das proteínas da família DUF_61 utilizando uma recente metodologia baseada na formação de uma rede neural de três vias denominada RoseTTAFold

(BAEK et al., 2021) permitiu a geração de modelos com mais acurácia (Figura 25) e, juntos aos modelos gerados em outros servidores, reforçam os indícios de que as proteínas desta família em leptospiras patogênicas podem se tratar de possíveis toxinas.

A hipótese de que as proteínas da família DUF_61 podem atuar como toxinas durante a infecção de leptospiras patogênicas também foi explorada e confirmada na recente pré-publicação de Chaurasia e colaboradores. O grupo explorou a atividade da proteína recombinante rLA3490, referente à proteína da família DUF_61 com maior regulação gênica positiva em amostras de sangue, fígado e rins de hamsters em modelo de infecção aguda demonstrada anteriormente pelo mesmo grupo (LEHMANN et al., 2013). A proteína recombinante mostrou compartilhar a mesma especificidade de ligação a glicoconjugados que a cadeia B da ricina, além da demonstração de efeitos citopáticos causados quando a proteína foi adicionada a culturas de células HeLa, incluindo arredondamento e destacamento celular, despolimerização de actina e fragmentação nuclear, levando à lise e morte celulares (CHAURASIA et al., 2020). Estes dados são os primeiros indícios da citotoxicidade desta proteína, função que pode ser exercida pelos outros membros da família DUF_61.

De maneira independente, seguimos um racional semelhante ao trabalho acima com as predições tridimensionais da proteína LepA388, além do desenho e produção da proteína fusionada EGFP-LepA388_Cards para avaliar a interação do domínio N-terminal de LepA388 com diferentes linhagens de células mamíferas e possíveis receptores específicos (experimentos em andamento). Outras construções também estão sendo avaliadas para análises adicionais, incluindo o provável efeito citotóxico da proteína LepA388. É possível que haja diferentes focos de atuação das proteínas da família DUF_61 dependendo das condições de infecção, além da cepa de *Leptospira* e hospedeiro envolvidos, abrindo um campo diverso de possibilidades de estudo, parte dos quais foram iniciados na presente tese, identificados inicialmente pela técnica de *phage display* pelo nosso grupo.

Finalmente, assim como demonstrado para a proteína LepA388, o estudo da proteína LepA069 também identificada por *phage display* teve início com a avaliação da expressão do seu gene codificador por RT-qPCR, resultando no aumento expressivo de sua expressão em amostras de fígado de camundongos susceptíveis após três dias de infecção por *L. interrogans* sorovar Copenhageni cepa Fiocruz L1-130. A técnica de RT-qPCR é eficientemente empregada para a quantificação e comparação acuradas da expressão gênica de bactérias virulentas em diversas amostras biológicas, e tem contribuído para o melhor entendimento de proteínas candidatas na participação durante o ciclo de infecção das leptospiros nos hospedeiros. Alguns exemplos incluem a demonstração da expressão gênica de colagenase (*colA*) nos pulmões de hamsters em modelo de infecção aguda (JANWITTHAYANAN et al., 2013), além de modelos de cultivo *in vitro* mimetizando as condições encontradas nos hospedeiros, como aumento da osmolaridade (de 67 mOsm para 300 mOsm) e/ou temperatura (de 30 °C para 37 °C), parâmetros que influenciam na regulação positiva de genes associados à virulência como *ligA*, *ligB*, *lenA*, *lipL32*, *lipL41*, *loa22*, *sph2*, entre outros (MATSUNAGA et al., 2005; PAPPAS; PICARDEAU, 2015; FRASER; BROWN, 2017). Considerando a possível proteína de membrana externa LepA069, o aumento de sua expressão gênica em 70 vezes nas amostras de animais infectados comparada à expressão em leptospiros cultivadas em meio EMJH a 30 °C (Figura 27), junto à sua identificação por *phage display* e seleção em células VERO (FERREIRA, 2015), são evidências inéditas de sua relação na interação patógeno-hospedeiro.

6. CONCLUSÕES E PERSPECTIVAS

Em resumo, o presente trabalho contribuiu para a caracterização de três proteínas hipotéticas de *L. interrogans*, as quais estão possivelmente envolvidas no arsenal de fatores de virulência utilizadas pelas bactérias patogênicas durante seu ciclo infeccioso. A caracterização da proteína recombinante LepA962_Phg identificada previamente por *phage display* enriqueceu o conhecimento já adquirido sobre a proteína LepA962 (ou Lp95), destacando um possível novo domínio de interação desta proteína com o hospedeiro durante a infecção. A proteína LepA069 ainda permanece com função desconhecida, mas demonstramos aqui os primeiros dados de sua possível participação na patogenicidade da leptospirose.

Dentre as proteínas estudadas, as proteínas da família DUF_61 se destacam. Os dados aqui apresentados com a proteína LepA388 são evidências de sua interação com o hospedeiro. Entretanto, devido à intrigante redundância desta família em *L. interrogans*, ainda há um imenso e empolgante campo de possibilidades no estudo desta família em leptospirosas patogênicas, o qual está sendo atualmente explorado pelo nosso grupo com a proteína em fusão EGFP-LepA388_Cards e outros domínios da proteína LepA388.

Como perspectivas, a avaliação do potencial imunogênico destas três proteínas hipotéticas como antígenos vacinais ou para uso em diagnóstico de leptospirose que constituía um dos objetivos iniciais deste projeto de Doutorado, também será explorado com a continuação dos atuais estudos. Adicionalmente, a elaboração de um manuscrito sobre os dados obtidos com proteínas recombinantes de LepA962 está sendo finalizada, assim como as presentes e futuras descobertas relacionadas à proteína LepA388 e outros membros da família DUF_61 também serão incluídas em outro manuscrito a ser elaborado.

7. REFERÊNCIAS

ADLER, B. Vaccines Against Leptospirosis. In: ADLER, B. (Ed.). *Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015a. 387p. 251–272.

ADLER, B. History of Leptospirosis and Leptospira. In: ADLER, B. (Ed.). *Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015b. 387p. 1–9.

ADLER, B.; DE LA PEÑA MOCTEZUMA, A. *Leptospira* and leptospirosis. *Veterinary Microbiology*, v. 140, n. 3–4, p. 287–296, 2010.

AHMED, S. A.; SANDAI, D. A.; MUSA, S.; HOCK, H. C.; RIADZI, M.; LEONG, L. K.; HOCK, T. T. Rapid Diagnosis of Leptospirosis by Multiplex PCR. *Malaysian Journal of Medical Sciences*, v. 19, n. 3, 2012.

ANFENSEN, C. B. Principles that Govern the Folding of Protein Chains. *Science*, v. 181, n. 4096, p. 223–230, 1973.

ANFENSEN, CHRISTIAN B., HABER, E. Studies on the Reduction and Re-formation of Protein Disulfide Bonds. *The Journal of Biological Chemistry*, v. 236, n. 5, p. 1361–1363, 1961.

ANTONARA, S.; CHAFEL, R. M.; LAFRANCE, M.; COBURN, J. *Borrelia burgdorferi* adhesins identified using in vivo phage display. *Molecular Microbiology*, v. 66, n. 1, 2007.

ATZINGEN, M. v.; BARBOSA, A. S.; DE BRITO, T.; VASCONCELLOS, S. A.; DE MORAIS, Z. M.; LIMA, D. M.; ABREU, P. A.; NASCIMENTO, A. L. Lsa21, a novel leptospiral protein binding adhesive matrix molecules and present during human infection. *BMC microbiology*, v. 8, p. 70, 2008.

ATZINGEN, M. v.; GÓMEZ, R. M.; SCHATTNER, M.; PRETTE, G.; GONÇALES, A. P.; DE MORAIS, Z. M.; VASCONCELLOS, S. A.; NASCIMENTO, A. L. T. O. Lp95, a novel leptospiral protein that binds extracellular matrix components and activates e-selectin on endothelial cells. *Journal of Infection*, v. 59, n. 4, p. 264–276, 2009.

BAEK, M.; DIMAIO, F.; ANISHCHENKO, I.; DAUPARAS, J.; OVCHINNIKOV, S.; LEE, G. R.; WANG, J.; CONG, Q.; KINCH, L. N.; DUSTIN SCHAEFFER, R.; MILLÁN, C.; PARK, H.; ADAMS, C.; GLASSMAN, C. R.; DEGIOVANNI, A.; PEREIRA, J. H.; RODRIGUES, A. v.; VAN DIJK, A. A.; EBRECHT, A. C.; OPPERMAN, D. J.; SAGMEISTER, T.; BUHLHELLER, C.; PAVKOV-KELLER, T.; RATHINASWAMY, M. K.; DALWADI, U.; YIP, C. K.; BURKE, J. E.; CHRISTOPHER GARCIA, K.; GRISHIN, N. v.; ADAMS, P. D.; READ, R. J.; BAKER, D. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, v. 373, n. 6557, 2021.

BAGARIA, S.; KARANDE, A. A. Abrin and Immunoneutralization: A Review. In: *Toxinology*. [s.l.: s.n.]p. 1–21.

BAL, A. E.; GRAVEKAMP, C.; HARTSKEERL, R. A.; DE MEZA-BREWSTER, J.; KORVER, H.; TERPSTRA, W. J. Detection of leptospires in urine by PCR for early diagnosis of leptospirosis. *Journal of Clinical Microbiology*, v. 32, n. 8, 1994.

- BARBOSA, A. S.; ABREU, P. A. E.; NEVES, F. O.; ATZINGEN, M. v.; WATANABE, M. M.; VIEIRA, M. L.; MORAIS, Z. M.; VASCONCELLOS, S. A.; NASCIMENTO, A. L. T. O. A newly identified leptospiral adhesin mediates attachment to laminin. *Infection and Immunity*, v. 74, n. 11, p. 6356–6364, 2006.
- BECKER, A.; KANNAN, T. R.; TAYLOR, A. B.; PAKHOMOVA, O. N.; ZHANG, Y. Structure of CARDS toxin , a unique ADP-ribosylating and vacuolating cytotoxin from *Mycoplasma pneumoniae*. v. 112, n. 16, p. 5165–5170, 2015.
- BERTONI, M.; KIEFER, F.; BIASINI, M.; BORDOLI, L.; SCHWEDE, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Scientific Reports*, v. 7, n. 1, p. 1–15, 2017.
- BIENERT, S.; WATERHOUSE, A.; DE BEER, T. A. P.; TAURIELLO, G.; STUDER, G.; BORDOLI, L.; SCHWEDE, T. The SWISS-MODEL Repository - new features and functionality. *Nucleic Acids Research*, v. 45, n. D1, p. D313–D319, 2017.
- BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, v. 72, n. 1–2, 1976.
- BRASIL. Ministério da Saúde. Leptospirose. <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/l/leptospirose>. . 2020.
- BREINER, D. D.; FAHEY, M.; SALVADOR, R.; NOVAKOVA, J.; COBURN, J. *Leptospira interrogans* binds to human cell surface receptors including proteoglycans. *Infection and Immunity*, v. 77, n. 12, 2009.
- BUCHANAN, R. E. Studies in the nomenclature and classification of the bacteria: II. The primary subdivisions of the schizomycetes. *Journal of Bacteriology*, v. 2, n. 2, p. 155–64, 1917.
- BULACH, D. M.; ZUERNER, R. L.; WILSON, P.; SEEMANN, T.; MCGRATH, A.; CULLEN, P. A.; DAVIS, J.; JOHNSON, M.; KUCZEK, E.; ALT, D. P.; PETERSON-BURCH, B.; COPPEL, R. L.; ROOD, J. I.; DAVIES, J. K.; ADLER, B. Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential. *Proceedings of the National Academy of Sciences of the United States of America*, v. 103, n. 39, p. 14560–14565, 2006.
- BUTERA, D.; FONTES PIAZZA, R. M.; MCLANE, M. A.; CHAMMAS, R.; DA SILVA, A. M. M. Molecular engineering of an EGFP/disintegrin-based integrin marker. *Toxicon*, v. 46, n. 2, p. 178–184, 2005.
- CAMERON, C. E. Leptospiral Structure, Physiology, and Metabolism. In: ADLER, B. (Ed.). *Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015. 387p. 21–41.
- CARVALHO, E.; BARBOSA, A. S.; GÓMEZ, R. M.; CIANCIARULLO, A. M.; HAUK, P.; ABREU, P. A.; FIORINI, L. C.; OLIVEIRA, M. L. S.; ROMERO, E. C.; GONÇALES, A. P.; MORAIS, Z. M.; VASCONCELLOS, S. A.; HO, P. L. Leptospiral TlyC is an extracellular matrix-binding protein and does not present hemolysin activity. *FEBS Letters*, v. 583, n. 8, p. 1381–1385, 2009.

- CASTIBLANCO-VALENCIA, M. M.; FRAGA, T. R.; SILVA, L. B. da; MONARIS, D.; ABREU, P. A. E.; STROBEL, S.; JÓZSI, M.; ISAAC, L.; BARBOSA, A. S. Leptospiral immunoglobulin-like proteins interact with human complement regulators factor H, FHL-1, FHR-1, and C4BP. *The Journal of infectious diseases*, v. 205, n. 6, p. 995–1004, 2012.
- CHAEMCHUEN, S.; RUNGPRAGAYPHAN, S.; POOVORAWAN, Y.; PATARAKUL, K. Identification of candidate host proteins that interact with LipL32, the major outer membrane protein of pathogenic *Leptospira*, by random phage display peptide library. *Veterinary Microbiology*, v. 153, n. 1–2, p. 178–185, 2011.
- CHAURASIA, R.; MARROQUIN, A.; MATTHIAS, M. A.; VINETZ, J. M. Novel R-type Lectin Domain-Containing Cytotoxins Comprise a Family of Virulence-Modifying Proteins in Pathogenic *Leptospira*. *bioRxiv*, p. 2020.05.13.094169, 1 jan. 2020.
- CHING, A. T. C.; FÁVARO, R. D.; LIMA, S. S.; CHAVES, A. D. A. M.; DE LIMA, M. A.; NADER, H. B.; ABREU, P. A. E.; HO, P. L. *Leptospira interrogans* shotgun phage display identified LigB as a heparin-binding protein. *Biochemical and Biophysical Research Communications*, v. 427, n. 4, p. 774–779, 2012.
- CHOU, L.-F.; CHEN, T.-W.; KO, Y.-C.; PAN, M.-J.; TIAN, Y.-C.; CHIU, C.-H.; TANG, P.; HUNG, C.-C.; YANG, C.-W. Potential impact on kidney infection: a whole-genome analysis of *Leptospira santarosai* serovar Shermani. *Emerging Microbes & Infections*, v. 3, n. 11, p. e82, 2014.
- CHOY, H. A.; KELLEY, M. M.; CHEN, T. L.; MØLLER, A. K.; MATSUNAGA, J.; HAAKE, D. A. Physiological osmotic induction of *Leptospira interrogans* adhesion: LigA and LigB bind extracellular matrix proteins and fibrinogen. *Infection and Immunity*, v. 75, n. 5, p. 2441–2450, 2007.
- CILIA, G.; BERTELLONI, F.; ALBINI, S.; FRATINI, F. Insight into the epidemiology of leptospirosis: A review of leptospira isolations from “unconventional” hosts. *Animals*, v. 11, n. 1, p. 1–16, 1 jan. 2021.
- COBURN, J.; CHEGE, W.; MAGOUN, L.; BODARY, S. C.; LEONG, J. M. Characterization of a candidate *Borrelia burgdorferi* β 3-chain integrin ligand identified using a phage display library. *Molecular Microbiology*, v. 34, n. 5, 1999.
- CONNOR, D. O.; ZANTOW, J.; HUST, M.; BIER, F. F.; VON NICKISCH-ROSENEGK, M. Identification of Novel Immunogenic Proteins of *Neisseria gonorrhoeae* by Phage Display. *Plos One*, v. 11, n. 2, p. e0148986, 2016.
- COSTA, F.; HAGAN, J. E.; CALCAGNO, J.; KANE, M.; TORGERSON, P.; MARTINEZ-SILVEIRA, M. S.; STEIN, C.; ABELA-RIDDER, B.; KO, A. I. Global Morbidity and Mortality of Leptospirosis: A Systematic Review. *PLoS neglected tropical diseases*, v. 9, n. 9, p. e0003898, 17 set. 2015.
- CROWTHER, J. R. Indirect ELISA. In: *ELISA. Methods in Molecular Biology*. [s.l.] Humana Press, 1995. p. 131–160.
- DA SILVA, J. B.; CARVALHO, E.; COVARRUBIAS, A. E.; CHING, A. T. C.; MATTARAIA, V. G. M.; PAIVA, D.; SE FRANCO, M.; FÁVARO, R. D.; PEREIRA, M. M.; VASCONCELLOS, S.; ZORN, T. T. M.; HO, P. L.; MARTINS, E. A. L. Induction of TNF-

alfa and CXCL-2 mRNAs in different organs of mice infected with pathogenic *Leptospira*. *Microbial Pathogenesis*, v. 52, n. 4, p. 206–216, 2012.

DE SOUZA, V. M. M.; ARSKY, M. de L. N. S.; DE CASTRO, A. P. B.; DE ARAUJO, W. N. Years of potential life lost and hospitalization costs associated with leptospirosis in Brazil. *Revista de Saude Publica*, v. 45, n. 6, 2011.

DELANO, W. L. The PyMOL Molecular Graphics System. *CCP4 Newsletter on protein crystallography*, v. 40, n. 1, 2002.

DELLAGOSTIN, O. A.; GRASSMANN, A. A.; RIZZI, C.; SCHUCH, R. A.; JORGE, S.; OLIVEIRA, T. L.; MCBRIDE, A. J. A.; HARTWIG, D. D. Reverse vaccinology: An approach for identifying leptospiral vaccine candidates. *International Journal of Molecular Sciences*, v. 18, n. 1, p. 158, 2017.

DOMINGOS, R. F.; VIEIRA, M. L.; ROMERO, E. C.; GONÇALES, A.; DE MORAIS, Z. M.; VASCONCELLOS, S. A.; NASCIMENTO, A. L. T. O. Features of two proteins of *Leptospira interrogans* with potential role in host-pathogen interactions. *BMC Microbiology*, v. 12, n. 1, p. 50, 2012.

DONAHUE, R. A.; BEBEE, R. L. BL21-SI Competent Cells for Protein Expression in *E. coli*. *Focus*, v. 21, n. 2, p. 49–51, 1999.

DONG, W. H.; CHEN, Z. Leptospirosis with pulmonary haemorrhage and multiple organ failure: a case report and literature review. *Journal of International Medical Research*, v. 49, n. 5, 2021.

EL RAYES, J.; RODRÍGUEZ-ALONSO, R.; COLLET, J. F. Lipoproteins in Gram-negative bacteria: new insights into their biogenesis, subcellular targeting and functional roles. *Current Opinion in Microbiology*, v. 61, 2021.

ELLINGHAUSEN, H. C.; MCCULLOUGH, W. G. Nutrition of *Leptospira pomona* and growth of 13 other serotypes: a serum-free medium employing oleic albumin complex. *American journal of veterinary research*, v. 26, p. 39–44, 1965.

ERRINGTON, N.; IQBALSAYAH, T.; DOIG, A. J. Structure and Stability of the α -Helix. In: GUEROIS, R.; DE LA PAZ, M. L. (Ed.). *Protein Design: Methods and Applications*. Totowa, NJ: Humana Press, 2006. p. 3–26.

EVANGELISTA, K.; FRANCO, R.; SCHWAB, A.; COBURN, J. *Leptospira interrogans* Binds to Cadherins. *PLoS Neglected Tropical Diseases*, v. 8, n. 1, p. 15, 2014a.

EVANGELISTA, K. v; COBURN, J. *Leptospira* as an emerging pathogen: a review of its biology, pathogenesis and host immune responses. *Future Microbiology*, v. 5, n. 9, p. 1413–1425, 2010.

EVANGELISTA, K. v.; HAHN, B.; WUNDER, E. A.; KO, A. I.; HAAKE, D. A.; COBURN, J. Identification of Cell-Binding Adhesins of *Leptospira interrogans*. *PLoS Neglected Tropical Diseases*, v. 8, n. 10, p. e3215, 2014b.

FAINE, S.; ADLER, B.; BOLIN, C.; PEROLAT, P. *Leptospira and leptospirosis*. 2nd. ed. Melbourne: MediSci, 1999. 296 p

FAISAL, S. M.; YAN, W. W.; MCDONOUGH, S. P.; CHANG, C. F.; PAN, M. J.; CHANG, Y. F. Leptosome-entrapped leptospiral antigens conferred significant higher levels of protection than those entrapped with PC-liposomes in a hamster model. *Vaccine*, v. 27, n. 47, 2009a.

FAISAL, S. M.; YAN, W. W.; MCDONOUGH, S. P.; MOHAMMED, H. O.; DIVERS, T. J.; CHANG, Y. F. Immune response and prophylactic efficacy of smegmosomes in a hamster model of leptospirosis. *Vaccine*, v. 27, n. 44, 2009b.

FELIX, C. R.; SIEDLER, B. S.; BARBOSA, L. N.; TIMM, G. R.; MCFADDEN, J.; MCBRIDE, A. J. A. An overview of human leptospirosis vaccine design and future perspectives. *Expert Opinion on Drug Discovery*, v. 15, n. 2, 2020.

FERNANDES, L. G. v; VIEIRA, M. L.; KIRCHGATTER, K.; ALVES, I. J.; DE MORAIS, Z. M.; VASCONCELLOS, S. A.; ROMERO, E. C.; NASCIMENTO, A. L. T. O. Omp11 is an extracellular matrix- and plasminogen-interacting protein of *Leptospira* spp. *Infection and Immunity*, v. 80, n. 10, p. 3679–3692, 2012.

FERREIRA, F. L. Identificação e avaliação de novas adesinas em *Leptospira interrogans* por shotgun phage display. *Dissertação de Mestrado - Instituto de Química, Universidade de São Paulo, São Paulo*, p. 180, 2015.

FERRER, M.; CHERNIKOVA, T. N.; YAKIMOV, M. M.; GOLYSHIN, P. N.; TIMMIS, K. N. Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nature Biotechnology*, v. 21, n. 11, p. 1266–1267, 2003.

FEVRE, C.; SCHEEPMAKER, L.; HAAS, P. J. Identifying bacterial immune evasion proteins using phage display. In: *Methods in Molecular Biology*. [s.l.] Humana Press Inc., 2017. 1535p. 43–61.

FOUTS, D. E.; MATTHIAS, M. A.; ADHIKARLA, H.; ADLER, B.; AMORIM-SANTOS, L.; BERG, D. E.; BULACH, D.; BUSCHIAZZO, A.; CHANG, Y. F.; GALLOWAY, R. L.; HAAKE, D. A.; HAFT, D. H.; HARTSKEERL, R.; KO, A. I.; LEVETT, P. N.; MATSUNAGA, J.; MECHALY, A. E.; MONK, J. M.; NASCIMENTO, A. L. T.; NELSON, K. E.; PALSSON, B.; PEACOCK, S. J.; PICARDEAU, M.; RICARDI, J. N.; THAIPANDUNGPANIT, J.; WUNDER, E. A.; YANG, X. F.; ZHANG, J. J.; VINETZ, J. M. What Makes a Bacterial Species Pathogenic?: Comparative Genomic Analysis of the Genus *Leptospira*. *PLoS Neglected Tropical Diseases*, v. 10, n. 2, p. 1–57, 2016.

FRASER, T.; BROWN, P. D. Temperature and oxidative stress as triggers for virulence gene expression in pathogenic *Leptospira* spp. *Frontiers in Microbiology*, v. 8, n. MAY, 2017.

GOARANT, C.; DELLAGI, K.; PICARDEAU, M. Ending the neglect of treatable bacterial zoonoses responsible for non-Malaria fevers. *Yale Journal of Biology and Medicine*, v. 94, n. 2, 2021.

GOARANT, C.; TRUEBA, G.; BIERQUE, E.; THIBEAUX, R.; DAVIS, B.; DE LA PENA-MOCTEZUMA, A. *Leptospira* and Leptospirosis. In: *Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management (Global Water Pathogen Project)*. [s.l.] Michigan State University, 2019.

GRASSMANN, A. A.; SOUZA, J. D.; JOHN, A.; MCBRIDE, A.; JOHN, A.; MCBRIDE, A. A Universal Vaccine against Leptospirosis: Are We Going in the Right Direction? v. 8, n. March, p. 1–8, 2017.

GREENFIELD, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols*, v. 1, n. 6, p. 2876–2890, 2006.

GRUNBERG-MANAGO, M. Messenger RNA stability and its role in control of gene expression in bacteria and phages. *Annual Review of Genetics*, v. 33, p. 193–227, 1999.

GUEX, N.; PEITSCH, M. C.; SCHWEDE, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*, v. 30, n. S1, p. S162–S173, 2009.

GUO, J. Q.; LI, Q. M.; ZHOU, J. Y.; ZHANG, G. P.; YANG, Y. Y.; XING, G. X.; ZHAO, D.; YOU, S. Y.; ZHANG, C. Y. Efficient recovery of the functional IP10-scFv fusion protein from inclusion bodies with an on-column refolding system. *Protein Expression and Purification*, v. 45, n. 1, 2006.

GUPTA, R. S.; MAHMOOD, S.; ADEOLU, M. Erratum : A phylogenomic and molecular signature based approach for characterization of the Phylum Spirochaetes and its major clades: proposal for a taxonomic revision of the phylum. *Frontiers in Microbiology*, v. 4, n. October 31, p. 322, 2013.

GUPTA, R. S.; MAHMOOD, S.; ADEOLU, M.; LY, B. A phylogenomic and molecular signature based approach for characterization of the phylum Spirochaetes and its major clades: proposal for a taxonomic revision of the phylum. *Frontiers in Microbiology*, v. 4, n. July 30, p. 1–18, 2013.

HAAKE, D. A.; LEVETT, P. N. Leptospirosis in Humans. In: ADLER, B. (Ed.). *Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015. 387p. 65–97.

HAAKE, D. A.; MAZEL, M. K.; MCCOY, A. M.; MILWARD, F.; CHAO, G.; MATSUNAGA, J.; WAGAR, E. A. Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection. *Infection and Immunity*, v. 67, n. 12, 1999.

HAAKE, D. A.; ZÜCKERT, W. R. The Leptospiral Outer Membrane. In: ADLER, B. (Ed.). *Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015. 387p. 187–221.

HALL, R. A. Studying protein-protein interactions via blot overlay/far western blot. *Methods in Molecular Biology*, v. 1278, 2015.

HARTWIG, D. D.; FORSTER, K. M.; OLIVEIRA, T. L.; AMARAL, M.; MCBRIDE, A. J. A.; DELLAGOSTINA, O. A. A prime-boost strategy using the novel vaccine candidate, LemA, protects hamsters against leptospirosis. *Clinical and Vaccine Immunology*, v. 20, n. 5, 2013.

HATTA, M.; SMITS, H. L.; GUSSENHOVEN, G. C.; GOOSKENS, J. Introduction of a rapid dipstick assay for the detection of leptospira-specific immunoglobulin M antibodies in the laboratory diagnosis of leptospirosis in a hospital in Makassar, Indonesia. *Southeast Asian Journal of Tropical Medicine and Public Health*, v. 31, n. 3, 2000.

HAUK, P.; MACEDO, F.; ROMERO, E. C.; VASCONCELLOS, S. A.; DE MORAIS, Z. M.; BARBOSA, A. S.; HO, P. L. In LipL32, the major leptospiral lipoprotein, the C terminus is the primary immunogenic domain and mediates interaction with collagen IV and plasma fibronectin. *Infection and immunity*, v. 76, n. 6, p. 2642–50, 2008.

HAZES, B. The (QxW)₃ domain: A flexible lectin scaffold. *Protein Science*, v. 5, n. 8, p. 1490–1501, 1996.

HILDEBRAND, A.; REMMERT, M.; BIEGERT, A.; SÖDING, J. Fast and accurate automatic structure prediction with HHpred. *Proteins*, v. 77, n. SUPPL. 9, p. 128–132, 2009.

HIRANUMA, N.; PARK, H.; BAEK, M.; ANISHCHENKO, I.; DAUPARAS, J.; BAKER, D. Improved protein structure refinement guided by deep learning based accuracy estimation. *Nature Communications*, v. 12, n. 1, 2021.

HOKE, D. E.; EGAN, S.; CULLEN, P. A.; ADLER, B. LipL32 is an extracellular matrix-interacting protein of *Leptospira* spp. and *Pseudoalteromonas tunicata*. *Infection and Immunity*, v. 76, n. 5, p. 2063–2069, 2008.

HORNBECK, P.; WINSTON, S. E.; FULLER, S. A. Enzyme-Linked Immunosorbent Assays (ELISA). In: *Current Protocols in Molecular Biology*. [s.l.: s.n.]15p. 11.2.1-11.2.22.

JANWITTHAYANAN, W.; KEELAWAT, S.; PAYUNGPORN, S.; LOWANITCHAPAT, A.; SUWANCHAROEN, D.; POOVORAWAN, Y.; CHIRATHAWORN, C. *In vivo* gene expression and immunoreactivity of *Leptospira* collagenase. *Microbiological Research*, v. 168, n. 5, 2013.

JEONG, H.; BARBE, V.; LEE, C. H.; VALLENET, D.; YU, D. S.; CHOI, S. H.; COULOUX, A.; LEE, S. W.; YOON, S. H.; CATTOLICO, L.; HUR, C. G.; PARK, H. S.; SÉGURENS, B.; KIM, S. C.; OH, T. K.; LENSKI, R. E.; STUDIER, F. W.; DAEGELEN, P.; KIM, J. F. Genome Sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *Journal of Molecular Biology*, v. 394, n. 4, p. 644–652, 2009.

JOHNSON, D. I.; JOHNSON, D. I. Bacterial Virulence Factors. In: *Bacterial Pathogens and Their Virulence Factors*. [s.l.] Springer International Publishing, 2018. p. 1–38.

KANNAN, T. R.; KRISHNAN, M.; RAMASAMY, K.; BECKER, A.; PAKHOMOVA, O. N.; HART, P. J.; BASEMAN, J. B. Functional mapping of community-acquired respiratory distress syndrome (CARDS) toxin of *Mycoplasma pneumoniae* defines regions with ADP-ribosyltransferase, vacuolating and receptor-binding activities. *Molecular Microbiology*, v. 93, n. 3, p. 568–581, 2014.

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

KHAKI, P. Clinical Laboratory Diagnosis of Human Leptospirosis. *International Journal of Enteric Pathogens*, v. 4, n. 1, 2016.

KLINE, K. A.; FÄLKER, S.; DAHLBERG, S.; NORMARK, S.; HENRIQUES-NORMARK, B. Bacterial Adhesins in Host-Microbe Interactions. *Cell Host and Microbe*, v. 5, n. 6, p. 580–592, 2009.

- KO, A. I.; GOARANT, C.; PICARDEAU, M. *Leptospira*: the dawn of the molecular genetics era for an emerging zoonotic pathogen. *Nature reviews. Microbiology*, v. 7, n. 10, p. 736–747, 2009.
- KWON, S. K.; KIM, S. K.; LEE, D. H.; KIM, J. F. Comparative genomics and experimental evolution of *Escherichia coli* BL21(DE3) strains reveal the landscape of toxicity escape from membrane protein overproduction. *Nature Scientific reports*, v. 5, p. 16076, 2015.
- LAURETTI-FERREIRA, F.; SILVA, P. L. D.; ALCÂNTARA, N. M.; SILVA, B. F.; GRABHER, I.; SOUZA, G. O.; NAKAJIMA, E.; AKAMATSU, M. A.; VASCONCELLOS, S. A.; ABREU, P. A. E.; CARVALHO, E.; MARTINS, E. A. L.; HO, P. L.; DA SILVA, J. B. New strategies for *Leptospira* vaccine development based on LPS removal. *PLoS ONE*, v. 15, n. 3, 2020.
- LEHMANN, J.; MATTHIAS, M.; VINETZ, J.; FOUTS, D. Leptospiral Pathogenomics. *Pathogens*, v. 3, n. 2, p. 280–308, 2014.
- LEHMANN, J. S.; FOUTS, D. E.; HAFT, D. H.; CANNELLA, A. P.; RICARDI, J. N.; BRINKAC, L.; HARKINS, D.; DURKIN, S.; SANKA, R.; SUTTON, G.; MORENO, A.; VINETZ, J. M.; MATTHIAS, M. A. Pathogenomic Inference of Virulence-Associated Genes in *Leptospira interrogans*. *PLoS Neglected Tropical Diseases*, v. 7, n. 10, p. 1–11, 2013.
- LEMERCIER, G.; BAKALARA, N.; SANTARELLI, X. On-column refolding of an insoluble histidine tag recombinant exopolyphosphatase from *Trypanosoma brucei* overexpressed in *Escherichia coli*. *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, v. 786, n. 1–2, 2003.
- LEVETT, P. N. Leptospirosis. *Clinical Microbiology*, v. 14, n. 2, p. 296–326, 2001.
- LEVETT, P. N. Systematics of Leptospiraceae. In: ADLER, B. (Ed.). *Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015. 387p. 11–20.
- LEVETT, P. N.; WHITTINGTON, C. U. Evaluation of the indirect hemagglutination assay for diagnosis of acute leptospirosis. *Journal of Clinical Microbiology*, v. 36, n. 1, 1998.
- LIMA, S. S.; CHING, A. T. C.; FÁVARO, R. D.; DA SILVA, J. B.; OLIVEIRA, M. L. S.; CARVALHO, E.; ABREU, P. A. E.; VASCONCELLOS, S. A.; HO, P. L. Adhesin activity of *Leptospira interrogans* lipoprotein identified by in vivo and in vitro shotgun phage display. *Biochemical and Biophysical Research Communications*, v. 431, n. 2, p. 342–347, 2013.
- LIN, Y. P.; LEE, D. W.; MCDONOUGH, S. P.; NICHOLSON, L. K.; SHARMA, Y.; CHANG, Y. F. Repeated domains of *Leptospira* immunoglobulin-like proteins interact with elastin and tropoelastin. *Journal of Biological Chemistry*, v. 284, n. 29, p. 19380–19391, 2009.
- LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods*, v. 25, n. 4, p. 402–8, 2001.
- MARCHIORI, E.; LOURENÇO, S.; SETÚBAL, S.; ZANETTI, G.; GASPARETTO, T. D.; HOCHHEGGER, B. Clinical and imaging manifestations of hemorrhagic pulmonary leptospirosis: A state-of-the-art review. *Lung*, v. 189, n. 1, p. 1–9, fev. 2011.
- MARCSISIN, R. A.; BARTPHO, T.; BULACH, D. M.; SRIKRAM, A.; SERMSWAN, R. W.; ADLER, B.; MURRAY, G. L. Use of a high-throughput screen to identify *Leptospira* mutants

unable to colonize the carrier host or cause disease in the acute model of infection. *Journal of Medical Microbiology*, v. 62, n. 10, p. 1601–1608, 2013.

MATSUI, M.; SOUPÉ, M. E.; BECAM, J.; GOARANT, C. Differential *In Vivo* gene expression of major *Leptospira* proteins in resistant or susceptible animal models. *Applied and Environmental Microbiology*, v. 78, n. 17, p. 6372–6376, 2012.

MATSUNAGA, J.; LO, M.; BULACH, D. M.; ZUERNER, R. L.; ADLER, B.; HAAKE, D. A. Response of *Leptospira interrogans* to physiologic osmolarity: Relevance in signaling the environment-to-host transition. *Infection and Immunity*, v. 75, n. 6, p. 2864–2874, 2007.

MATSUNAGA, J.; SANCHEZ, Y.; XU, X.; HAAKE, D. A. Osmolarity, a key environmental signal controlling expression of leptospiral proteins LigA and LigB and the extracellular release of LigA. *Infection and Immunity*, v. 73, n. 1, p. 70–78, 2005.

MENDES, R. S.; VON ATZINGEN, M.; DE MORAIS, Z. M.; GONÇALES, A. P.; SERRANO, S. M. T.; ASEGA, A. F.; ROMERO, E. C.; VASCONCELLOS, S. A.; NASCIMENTO, A. L. T. O. The novel leptospiral surface adhesin Lsa20 binds laminin and human plasminogen and is probably expressed during infection. *Infection and Immunity*, v. 79, n. 11, p. 4657–4667, 2011.

MERIEN, F.; PORTNOI, D.; BOURHY, P.; CHARAVAY, F.; BERLIOZ-ARTHAUD, A.; BARANTON, G. A rapid and quantitative method for the detection of *Leptospira* species in human leptospirosis. *FEMS Microbiology Letters*, v. 249, n. 1, p. 139–147, 2005.

MERIEN, F.; TRUCCOLO, J.; BARANTON, G.; PEROLAT, P. Identification of a 36-kDa fibronectin-binding protein expressed by a virulent variant of *Leptospira interrogans* serovar icterohaemorrhagiae. *FEMS Microbiology Letters*, v. 185, n. 1, 2000.

MORENO, A. T.; OLIVEIRA, M. L. S.; FERREIRA, D. M.; HO, P. L.; DARRIEUX, M.; LEITE, L. C. C.; FERREIRA, J. M. C.; PIMENTA, F. C.; ANDRADE, A. L. S. S.; MIYAJI, E. N. Immunization of mice with single PspA fragments induces antibodies capable of mediating complement deposition on different pneumococcal strains and cross-protection. *Clinical and Vaccine Immunology*, v. 17, n. 3, 2010.

MULLEN, L. M.; NAIR, S. P.; WARD, J. M.; RYCROFT, A. N.; HENDERSON, B. Phage display in the study of infectious diseases. *Trends in Microbiology*, v. 14, n. 3, p. 141–147, 2006.

MURRAY, G. L. The lipoprotein LipL32, An enigma of leptospiral biology. *Veterinary Microbiology*, v. 162, n. 2–4, 2013.

MURRAY, G. L. The Molecular Basis of Leptospiral Pathogenesis. In: ADLER, B. (Ed.). *Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015. 387p. 139–185.

MURRAY, G. L.; SRIKRAM, A.; HOKE, D. E.; WUNDER, E. A.; HENRY, R.; LO, M.; ZHANG, K.; SERMSWAN, R. W.; KO, A. I.; ADLER, B. Major surface protein LipL32 is not required for either acute or chronic infection with *Leptospira interrogans*. *Infection and Immunity*, v. 77, n. 3, p. 952–958, 2009.

- NILSSON, M.; FRYKBERG, L.; FLOCK, J. I.; PEI, L.; LINDBERG, M.; GUSS, B. A fibrinogen-binding protein of *Staphylococcus epidermidis*. *Infection and immunity*, v. 66, n. 6, p. 2666–2673, 1998.
- NOGUEIRA, S. V.; BACKSTEDT, B. T.; SMITH, A. A.; QIN, J. H.; WUNDER, E. A.; KO, A.; PAL, U. *Leptospira interrogans* Enolase Is Secreted Extracellularly and Interacts with Plasminogen. *PLoS ONE*, v. 8, n. 10, 2013.
- ODIR, A. D.; GRASSMANN, A. A.; HARTWIG, D. D.; FÉLIX, S. R.; DA SILVA, É. F.; MCBRIDE, A. J. A. Recombinant vaccines against leptospirosis. *Human Vaccines*, v. 7, n. 11, 2011.
- OLIVEIRA, R.; DE MORAIS, Z. M.; GONÇALES, A. P.; ROMERO, E. C.; VASCONCELLOS, S. A.; NASCIMENTO, A. L. T. O. Characterization of novel OmpA-like protein of *Leptospira interrogans* that binds extracellular matrix molecules and plasminogen. *PLoS ONE*, v. 6, n. 7, 2011.
- OVCHINNIKOV, S.; PARK, H.; VARGHESE, N.; HUANG, P.-S.; PAVLOPOULOS, G. A.; KIM, D. E.; KAMISSETTY, H.; KYRPIDES, N. C.; BAKER, D. Protein structure determination using metagenome sequence data. *Science*, v. 355, n. 6322, p. 294–298, 2017.
- PAPPAS, C. J.; PICARDEAU, M. Control of gene expression in *Leptospira* spp. by transcription activator-like effectors demonstrates a potential role for LigA and LigB in *Leptospira interrogans* virulence. *Applied and Environmental Microbiology*, v. 81, n. 22, 2015.
- PASTER, B. J. Phylum XV. Spirochaetes Garrity and Holt 2001. In: KRIEG, N. R.; STALEY, J. T.; BROWN, D. R.; HEDLUND, B. P.; PASTER, B. J.; WARD, N. L.; LUDWIG, W.; WHITMAN, W. B. (Ed.). *Bergey's Manual of Systematic Bacteriology*. New York: Springer, 2010. p. 546–563.
- PETTERSEN, E. F.; GODDARD, T. D.; HUANG, C. C.; COUCH, G. S.; GREENBLATT, D. M.; MENG, E. C.; FERRIN, T. E. UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, v. 25, n. 13, 2004.
- PFAFFL, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, v. 29, n. 9, 2001.
- PICARDEAU, M. Genomics, Proteomics, and Genetics of *Leptospira*. In: ADLER, B. (Ed.). *Leptospira and Leptospirosis. Current Topics in Microbiology and Immunology*. [s.l.] Springer-Verlag Berlin Heidelberg, 2015. 387p. 43–63.
- PICARDEAU, M. Virulence of the zoonotic agent of leptospirosis: Still terra incognita? *Nature Reviews Microbiology*, v. 15, p. 297–307, 2017.
- PICARDEAU, M. Toolbox of molecular techniques for studying leptospira Spp. In: *Current Topics in Microbiology and Immunology*. [s.l.] Springer Verlag, 2018. 415p. 141–162.
- PICARDEAU, M. *Leptospira* and Leptospirosis. In: *Methods in Molecular Biology*. [s.l.] Humana Press Inc., 2020. 2134p. 271–275.
- PICARDEAU, M.; BULACH, D. M.; BOUCHIER, C.; ZUERNER, R. L.; ZIDANE, N.; WILSON, P. J.; CRENO, S.; KUCZEK, E. S.; BOMMEZZADRI, S.; DAVIS, J. C.; MCGRATH, A.; JOHNSON, M. J.; BOURSAUX-EUDE, C.; SEEMANN, T.; ROUY, Z.;

COPPEL, R. L.; ROOD, J. I.; LAJUS, A.; DAVIES, J. K.; MÉDIGUE, C.; ADLER, B. Genome sequence of the saprophyte *Leptospira biflexa* provides insights into the evolution of *Leptospira* and the pathogenesis of leptospirosis. *PLoS ONE*, v. 3, n. 2, p. 1–9, 2008.

PINNE, M.; CHOY, H. A.; HAAKE, D. A. The OmpL37 surface-exposed protein is expressed by pathogenic *Leptospira* during infection and binds skin and vascular elastin. *PLoS neglected tropical diseases*, v. 4, n. 9, p. e815, 2010.

PINNE, M.; HAAKE, D. A. LipL32 Is a Subsurface Lipoprotein of *Leptospira interrogans*: Presentation of New Data and Reevaluation of Previous Studies. *PLoS ONE*, v. 8, n. 1, 2013.

PINNE, M.; MATSUNAGA, J.; HAAKE, D. A. Leptospiral Outer Membrane Protein Microarray, a Novel Approach: To Identification of Host Ligand-Binding Proteins. *Journal of Bacteriology*, v. 194, n. 22, p. 6074–6087, 2012.

PIZARRO-CERDÁ, J.; COSSART, P. Bacterial adhesion and entry into host cells. *Cell*, v. 124, n. 4, p. 715–727, 2006.

PRASAD, S.; KHADATARE, P. B.; ROY, I. Effect of chemical chaperones in improving the solubility of recombinant proteins in *Escherichia coli*. *Applied and Environmental Microbiology*, v. 77, n. 13, p. 4603–4609, 2011.

PRINZ, W. A.; ÅSLUND, F.; HOLMGREN, A.; BECKWITH, J. The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the *Escherichia coli* cytoplasm. *Journal of Biological Chemistry*, v. 272, n. 25, p. 15661–15667, 1997.

PUTZ, E. J.; SIVASANKARAN, S. K.; FERNANDES, L. G. V.; BRUNELLE, B.; LIPPOLIS, J. D.; ALT, D. P.; BAYLES, D. O.; HORNSBY, R. L.; NALLY, J. E. Distinct transcriptional profiles of *Leptospira borgpetersenii* serovar hardjo strains jb197 and hb203 cultured at different temperatures. *PLoS Neglected Tropical Diseases*, v. 15, n. 4, 2021.

RAMOS, C. R. R.; ABREU, P. A. E.; NASCIMENTO, A. L. T. O.; HO, P. L. A high-copy T7 *Escherichia coli* expression vector for the production of recombinant proteins with a minimal N-terminal His-tagged fusion peptide. *Brazilian Journal of Medical and Biological Research*, v. 37, n. 8, p. 1103–1109, 2004.

REN, S.-X.; FU, G.; JIANG, X.-G.; ZENG, R.; MIAO, Y.-G.; XU, H.; ZHANG, Y.-X.; XIONG, H.; LU, G.; LU, L.-F.; JIANG, H.-Q.; JIA, J.; TU, Y.-F.; JIANG, J.-X.; GU, W.-Y.; ZHANG, Y.-Q.; CAI, Z.; SHENG, H.-H.; YIN, H.-F.; ZHANG, Y.; ZHU, G.-F.; WAN, M.; HUANG, H.-L.; QIAN, Z.; WANG, S.-Y.; MA, W.; YAO, Z.-J.; SHEN, Y.; QIANG, B.-Q.; XIA, Q.-C.; GUO, X.-K.; DANCHIN, A.; SAINT GIRONS, I.; SOMERVILLE, R. L.; WEN, Y.-M.; SHI, M.-H.; CHEN, Z.; XU, J.-G.; ZHAO, G.-P. Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature*, v. 422, n. 6934, p. 888–893, 2003.

RICALDI, J. N.; FOUTS, D. E.; SELENGUT, J. D.; HARKINS, D. M.; PATRA, K. P.; MORENO, A.; LEHMANN, J. S.; PURUSHE, J.; SANKA, R.; TORRES, M.; WEBSTER, N. J.; VINETZ, J. M.; MATTHIAS, M. A. Whole Genome Analysis of *Leptospira licerasiae* Provides Insight into Leptospiral Evolution and Pathogenicity. *PLoS Neglected Tropical Diseases*, v. 6, n. 10, p. e1853, 2012.

- ROBBINS, G. T.; HAHN, B. L.; EVANGELISTA, K. v.; PADMORE, L.; ARANDA, P. S.; COBURN, J. Evaluation of Cell Binding Activities of Leptospira ECM Adhesins. *PLOS Neglected Tropical Diseases*, v. 9, n. 4, p. 1–16, 2015.
- ROGL, H.; KOSEMUND, K.; KÜHLBRANDT, W.; COLLINSON, I. Refolding of *Escherichia coli* produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. *FEBS Letters*, v. 432, n. 1–2, p. 21–26, 1998.
- ROSANDER, A.; GUSS, B.; FRYKBERG, L.; BJÖRKMAN, C.; NÄSLUND, K.; PRINGLE, M. Identification of immunogenic proteins in *Treponema phagedenis*-like strain V1 from digital dermatitis lesions by phage display. *Veterinary Microbiology*, v. 153, n. 3–4, 2011.
- ROY, A.; KUCUKURAL, A.; ZHANG, Y. I-TASSER: A unified platform for automated protein structure and function prediction. *Nature Protocols*, v. 5, n. 4, p. 725–738, 2010.
- RUTENBER, E.; READY, M.; ROBERTUS, J. D. Structure and evolution of ricin B chain. *Nature*, v. 326, n. 6113, p. 624–626, 1987.
- RUTENBER, E.; ROBERTUS, J. D. Structure of ricin B-chain at 2.5 Å resolution. *Proteins: Structure, Function, and Bioinformatics*, v. 10, n. 3, 1991.
- SAMBROOK, J.; RUSSELL, D. W. *Molecular cloning: a laboratory manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, 2001. 999 p
- SAMROT, A. v.; SEAN, T. C.; BHAVYA, K. S.; SAHITHYA, C. S.; CHANDRASEKARAN, S.; PALANISAMY, R.; ROBINSON, E. R.; SUBBIAH, S. K.; MOK, P. L. Leptospiral infection, pathogenesis and its diagnosis—a review. *Pathogens*, v. 10, n. 2, 2021.
- SANHUEZA, J. M.; BAKER, M. G.; BENSCHOP, J.; COLLINS-EMERSON, J. M.; WILSON, P. R.; HEUER, C. Estimation of the burden of leptospirosis in New Zealand. *Zoonoses and Public Health*, v. 67, n. 2, 2020.
- SELA, M.; WHITE, F. H.; ANFINSEN, C. B. Reductive cleavage of disulfide bridges in ribonuclease. *Science*, v. 125, n. 3250, p. 691–692, 1957.
- SHIMOKAWA-FALCÃO, L. H. A. L.; CAPORRINO, M. C.; BARBARO, K. C.; DELLA-CASA, M. S.; MAGALHÃES, G. S. Toxin fused with SUMO tag: A new expression vector strategy to obtain recombinant venom toxins with easy tag removal inside the bacteria. *Toxins*, v. 9, n. 3, 2017.
- SILVEIRA, M. M.; OLIVEIRA, T. L.; SCHUCH, R. A.; MCBRIDE, A. J. A.; DELLAGOSTIN, O. A.; HARTWIG, D. D. DNA vaccines against leptospirosis: A literature review. *Vaccine*, v. 35, n. 42, 2017.
- SIMIZU, B.; RHIM, J. S.; WIEBENGA, N. H. Characterization of the Tacaribe Group of Arboviruses. 1. Propagation and Plaque Assay of Tacaribe Virus in a Line of African Green Monkey Kidney Cells (Vero). *Proceedings of the Society for Experimental Biology and Medicine*, v. 125, n. 1, p. 119–123, 1967.
- SINGH, A.; UPADHYAY, V.; UPADHYAY, A. K.; SINGH, S. M.; PANDA, A. K. Protein recovery from inclusion bodies of *Escherichia coli* using mild solubilization process. *Microbial Cell Factories*, v. 14, n. 1, 2015.

SMITH, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, v. 228, n. 4705, p. 1315–1317, 1985.

SREERAMA, N.; WOODY, R. W. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical Biochemistry*, v. 287, n. 2, 2000.

STEVENSON, B.; CHOY, H. A.; PINNE, M.; ROTONDI, M. L.; MILLER, M. C.; DEMOLL, E.; KRAICZY, P.; COOLEY, A. E.; CREAMER, T. P.; SUCHARD, M. A.; BRISSETTE, C. A.; VERMA, A.; HAAKE, D. A. *Leptospira interrogans* endostatin-like outer membrane proteins bind host fibronectin, laminin and regulators of complement. *PLoS ONE*, v. 2, n. 11, p. 1–11, 2007.

STUDER, G.; REMPFER, C.; WATERHOUSE, A. M.; GUMIENNY, R.; HAAS, J.; SCHWEDE, T. QMEANDisCo - distance constraints applied on model quality estimation. *Bioinformatics*, v. 36, n. 6, p. 1765–1771, 2020.

TAHIROV, T. H.; LU, T. H.; LIAW, Y. C.; CHEN, Y. L.; LIN, J. Y. Crystal structure of abrin-a at 2.14 Å. *Journal of Molecular Biology*, v. 250, n. 3, p. 354–367, 1995.

TAN, W. S.; HO, K. L. Phage display creates innovative applications to combat hepatitis B virus. *World Journal of Gastroenterology*, v. 20, n. 33, p. 11650–11670, 2014.

TEIXEIRA, A. I. P.; SILVA, D. M.; VITAL, T.; NITZ, N.; DE CARVALHO, B. C.; HECHT, M.; OLIVEIRA, D.; OLIVEIRA, E.; RABELLO, A.; ROMERO, G. A. S. Improving the reference standard for the diagnosis of canine visceral leishmaniasis: A challenge for current and future tests. *Memorias do Instituto Oswaldo Cruz*, v. 114, n. 1, 2019.

TERPSTRA, W. J.; LIGTHART, G. S.; SCHOONE, G. J. ELISA for the detection of specific IgM and IgG in human leptospirosis. *Journal of General Microbiology*, v. 131, n. 2, 1985.

THIBEAUX, R.; IRAOLA, G.; FERRÉS, I.; BIERQUE, E.; GIRAULT, D.; SOUPÉ-GILBERT, M. E.; PICARDEAU, M.; GOARANT, C. Deciphering the unexplored *Leptospira* diversity from soils uncovers genomic evolution to virulence. *Microbial genomics*, v. 4, n. 1, 2018.

VERMA, A.; BRISSETTE, C. A.; BOWMAN, A. A.; SHAH, S. T.; ZIPFEL, P. F.; STEVENSON, B. Leptospiral endostatin-like protein A is a bacterial cell surface receptor for human plasminogen. *Infection and Immunity*, v. 78, n. 5, p. 2053–2059, 2010.

VERMA, A.; HELLWAGE, J.; ARTIUSHIN, S.; ZIPFEL, P. F.; KRAICZY, P.; TIMONEY, J. F.; STEVENSON, B. LfhA, a novel factor H-binding protein of *Leptospira interrogans*. *Infection and immunity*, v. 74, n. 5, p. 2659–66, 1 maio 2006.

VIEIRA, M. L.; ATZINGEN, M. v.; OLIVEIRA, T. R.; OLIVEIRA, R.; ANDRADE, D. M.; VASCONCELLOS, S. A.; NASCIMENTO, A. L. T. O. *In Vitro* identification of novel plasminogen-binding receptors of the pathogen *Leptospira interrogans*. *PLoS ONE*, v. 5, n. 6, 2010a.

VIEIRA, M. L.; DE MORAIS, Z. M.; GONÇALES, A. P.; ROMERO, E. C.; VASCONCELLOS, S. A.; NASCIMENTO, A. L. T. O. Lsa63, a newly identified surface

protein of *Leptospira interrogans* binds laminin and collagen IV. *Journal of Infection*, v. 60, n. 1, p. 52–64, 2010b.

VINCENT, A. T.; SCHIETTEKATTE, O.; GOARANT, C.; NEELA, V. K.; BERNET, E.; THIBEAUX, R.; ISMAIL, N.; KHALID, M. K. N. M.; AMRAN, F.; MASUZAWA, T.; NAKAO, R.; KORBA, A. A.; BOURHY, P.; VEYRIER, F. J.; PICARDEAU, M. Revisiting the taxonomy and evolution of pathogenicity of the genus *Leptospira* through the prism of genomics. *PLoS Neglected Tropical Diseases*, v. 13, n. 5, 2019.

WATERHOUSE, A.; BERTONI, M.; BIENERT, S.; STUDER, G.; TAURIELLO, G.; GUMIENNY, R.; HEER, F. T.; DE BEER, T. A. P.; REMPFER, C.; BORDOLI, L.; LEPORE, R.; SCHWEDE, T. SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, v. 46, n. W1, p. W296–W303, 2018.

WHITMORE, L.; WALLACE, B. A. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, v. 32, n. WEB SERVER ISS., 2004.

WHITMORE, L.; WALLACE, B. A. Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers*, v. 89, n. 5, 2008.

WILLIAMS, R. J.; HENDERSON, B.; SHARP, L. J.; NAIR, S. P. Identification of a Fibronectin-Binding Protein from *Staphylococcus epidermidis*. *Infect Immun*, v. 70, n. 12, p. 6805–10, 2002.

WU, Y.; LI, Q.; CHEN, X. Z. Detecting protein-protein interactions by far western blotting. *Nature Protocols*, v. 2, n. 12, 2007.

WYNWOOD, S. J.; BURNS, M. A. A.; GRAHAM, G. C.; WEIER, S. L.; MCKAY, D. B.; CRAIG, S. B. Validation of a Microsphere Immunoassay for Serological Leptospirosis Diagnosis in Human Serum by Comparison to the Current Gold Standard. *PLoS Neglected Tropical Diseases*, v. 9, n. 3, 2015.

XU, Y.; ZHU, Y.; WANG, Y.; CHANG, Y. F.; ZHANG, Y.; JIANG, X.; ZHUANG, X.; ZHU, Y.; ZHANG, J.; ZENG, L.; YANG, M.; LI, S.; WANG, S.; YE, Q.; XIN, X.; ZHAO, G.; ZHENG, H.; GUO, X.; WANG, J. Whole genome sequencing revealed host adaptation-focused genomic plasticity of pathogenic *Leptospira*. *Scientific Reports*, v. 6, 2016.

XUE, Z.; XU, D.; WANG, Y.; ZHANG, Y. ThreaDom: extracting protein domain boundary information from multiple threading alignments. *Bioinformatics*, v. 29, n. 13, p. i247-256, 2013.

YANG, J.; YAN, R.; ROY, A.; XU, D.; POISSON, J.; ZHANG, Y. The I-TASSER suite: Protein structure and function prediction. *Nature Methods*, v. 12, n. 1, p. 7–8, 2015.

YANG, J.; ZHANG, Y. I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Research*, v. 43, n. W1, p. W174–W181, 2015.

ZIMMERMANN, L.; STEPHENS, A.; NAM, S. Z.; RAU, D.; KÜBLER, J.; LOZAJIC, M.; GABLER, F.; SÖDING, J.; LUPAS, A. N.; ALVA, V. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *Journal of Molecular Biology*, v. 430, n. 15, p. 2237–2243, 2018.

ZÜCKERT, W. R. Secretion of Bacterial Lipoproteins: Through the Cytoplasmic Membrane, the Periplasm and Beyond. *Biochimica et Biophysica Acta - Molecular Cell Research*, v. 1843, n. 8, 2014.

APÊNDICES

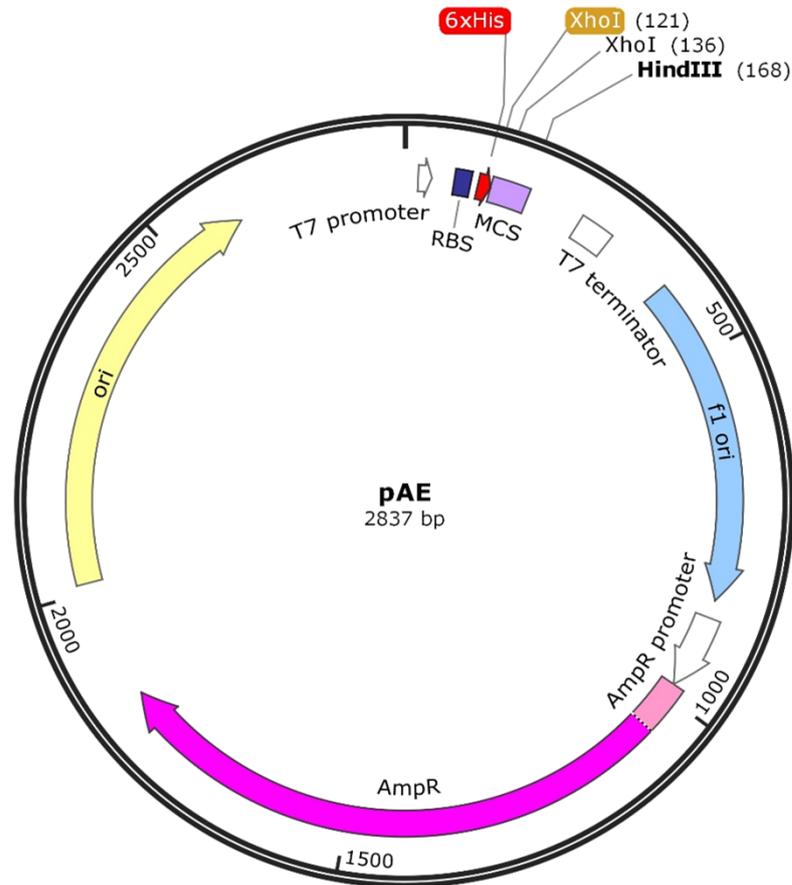


Figura S1. Mapa do vetor de expressão **pAE**, onde estão detalhados componentes como origens de replicação, promotores e o gene de resistência ao antibiótico ampicilina (**AmpR**, em rosa). A sequência codificando as seis histidinas (**6xHis**, em vermelho) se encontra antes do sítio de clonagem múltipla (**MCS**, em lilás), estando na extremidade N-terminal da proteína recombinante após sua expressão. Os sítios de restrição **XhoI** (destaque em laranja) e **HindIII** foram utilizados para a clonagem das sequências dos genes recombinantes. Mapa gerado pelo software SnapGene (GSL Biotech LLC).

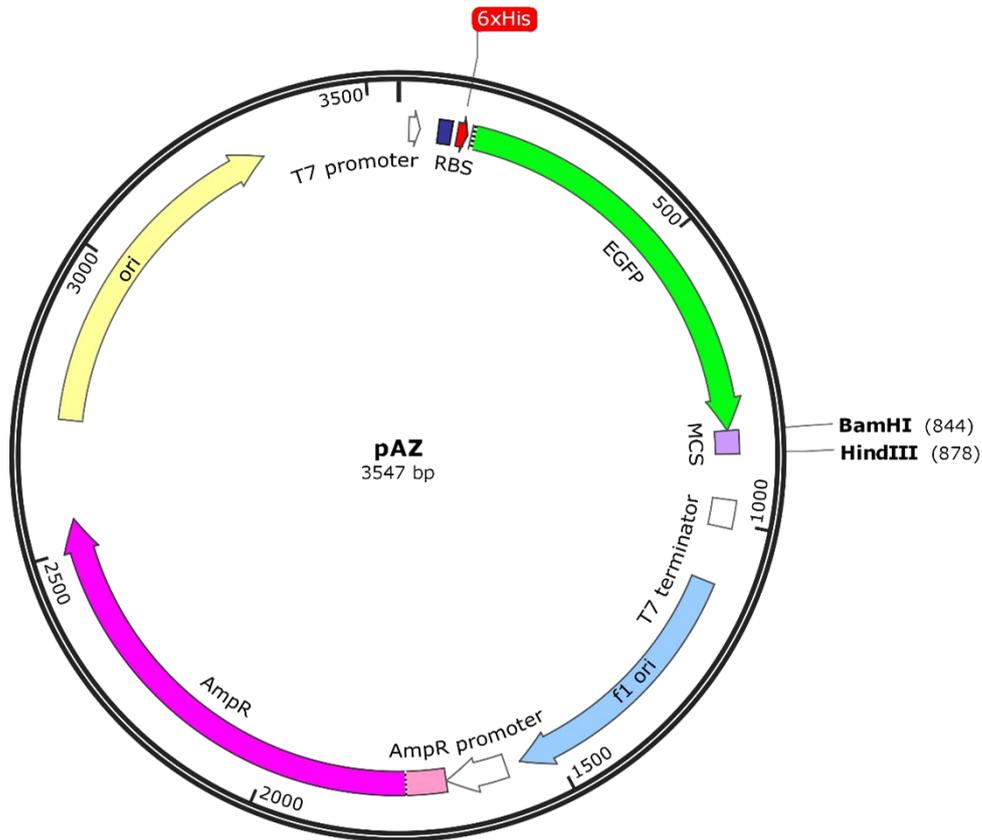


Figura S2. Mapa do vetor de expressão **pAZ**, onde estão detalhados componentes como origens de replicação, promotores e o gene de resistência ao antibiótico ampicilina (**AmpR**, em rosa). A sequência codificando as seis histidinas (**6xHis**, em vermelho) se encontra antes do sítio de clonagem múltipla (**MCS**, em lilás), estando na extremidade N-terminal da proteína recombinante após sua expressão. Os sítios de restrição **BamHI** e **HindIII** foram utilizados para a clonagem das sequências dos genes recombinantes. Mapa gerado pelo software SnapGene (GSL Biotech LLC).

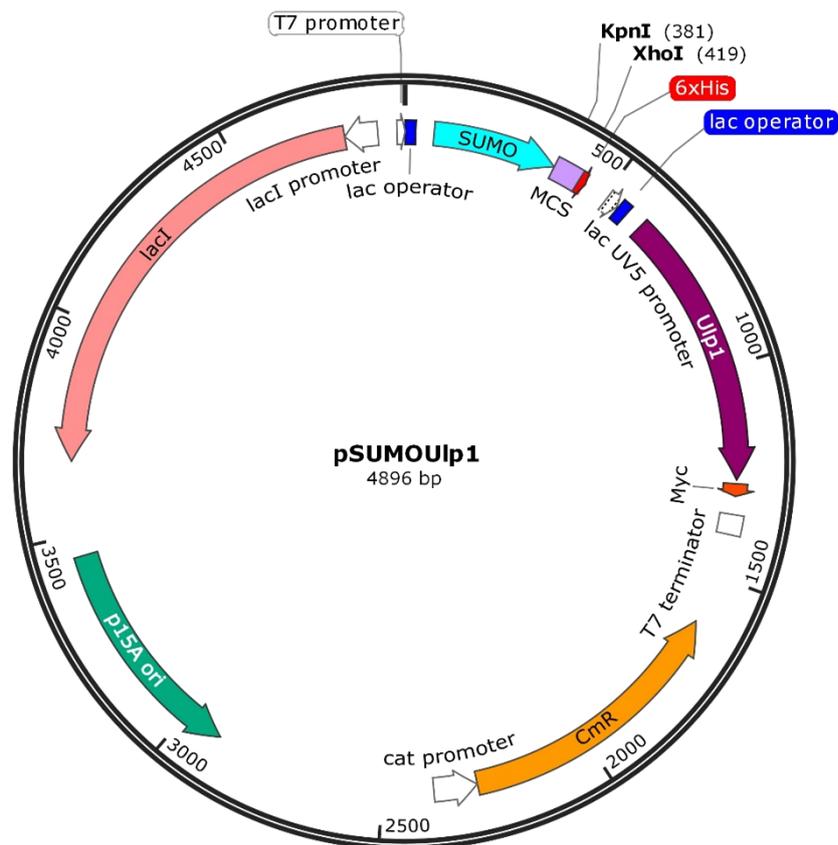


Figura S3. Mapa do vetor de expressão **pSUMOUlp1**, onde estão detalhados componentes como origens de replicação, promotores, operons e o gene de resistência ao antibiótico cloranfenicol (**CmR**, em laranja). A sequência codificando as seis histidinas (**6xHis**, em vermelho) se encontra depois do sítio de clonagem múltipla (**MCS**, em lilás), estando na extremidade C-terminal da proteína recombinante após sua expressão. Os sítios de restrição **KpnI** e **XhoI** foram utilizados para a clonagem das sequências dos genes recombinantes. A etiqueta **SUMO** (em azul) auxilia no dobramento da proteína recombinante fusionada e o gene da enzima **Ulp1** (em roxo) é expresso separadamente e a proteína **Ulp1** produzida promove a clivagem entre a etiqueta **SUMO** e a proteína recombinante em fusão, após reconhecer a estrutura tridimensional correta da **SUMO** (SHIMOKAWA-FALCÃO et al., 2017). Mapa gerado pelo software SnapGene (GSL Biotech LLC).

Tabela S1

Sequência dos pares de oligonucleotídeos utilizados como componentes das reações de qPCR.

Gene/Proteína de origem	Sequência dos oligonucleotídeos
LepA388	<i>F</i> [*] : 5' -GGTCTGGGCTGCCATTGAAT-3' <i>R</i> ^{**} : 5' -TCCTTCTCCACTACAATCCCT-3'
LepA069	<i>F</i> : 5' -TGGGATTCAATTATCAGCGACC-3' <i>R</i> : 5' -AACCAGAAAGTTGAACGGCTC-3'
LipL32	<i>F</i> : 5' -GGCGATTTGGTCAGGCATAA-3' <i>R</i> : 5' -CCGTAATCGCTGAAATGGGA-3'
16S	<i>F</i> : 5' -TTCAGTTGGGCACTCGTAAG-3' <i>R</i> : 5' -CGTGTGTTGCCCTAGACATAA-3'

^{*}*F*: sequências de oligonucleotídeos senso (*forward*);
^{**}*R*: sequências de oligonucleotídeos antisenso (*reverse*).

Tabela S2

Sequência dos pares de oligonucleotídeos utilizados na amplificação das sequências nucleotídicas referentes às proteínas recombinantes desenhadas, para clonagem e posterior expressão em bactérias.

Gene/Proteína de origem	Proteína recombinante	Sequência dos oligonucleotídeos
<u>LepA962</u>	LepA962_Nt	<i>F</i> [*] : 5' -CCG <u>CTCGAG</u> GAACGTCTTTCCGTTCCCTTCC-3' <i>R</i> ^{**} : 5' -CCC <u>AAGCTT</u> <u>TTA</u> TTCCGTGCTGATTCTTGATTT-3'
	LepA962_Phg	<i>F</i> : 5' -CCG <u>CTCGAG</u> CCGGATGAACGCCTTCTGGCT-3' <i>R</i> : 5' -CCC <u>AAGCTT</u> <u>TTA</u> TCTTTGTCTCATCACTTCTAC-3'
<u>LepA388</u>	LepA388_Blec	<i>F</i> : 5' -CCG <u>CTCGAG</u> CAAAGACCTACTGATAAACCA-3' <i>R</i> : 5' -CCC <u>AAGCTT</u> <u>TTA</u> ATCTTTTAACCGATAACGCTT-3'
	LepA388_Phg	<i>F</i> : 5' -CCG <u>CTCGAG</u> TATAATTGGTATGCCTATATT-3' <i>R</i> : 5' -CCC <u>AAGCTT</u> <u>TTA</u> ATAGAGCGATCCACTCACTGG-3'
	LepA388_D1	<i>F</i> : 5' -CCG <u>CTCGAG</u> CAAAGACCTACTGATAAACCA-3' <i>R</i> : 5' -CCC <u>AAGCTT</u> <u>TTA</u> AACACCCCAATTAGGTCCATA-3'
	LepA388_Cards	<i>F</i> : 5' -CCG <u>CTCGAG</u> CAAAGACCTACTGATAAACCA-3' <i>R</i> : 5' -CCC <u>AAGCTT</u> <u>TTA</u> GTAACGGACCCAATTTAATAA-3'
	LepA388_Ct	<i>F</i> : 5' -CCG <u>CTCGAG</u> TTACTTCATACCTTCCAGATG-3' <i>R</i> : 5' -CCC <u>AAGCTT</u> <u>TTA</u> AGAGATCATGAAATCAAAAAT-3'
	EGFP-LepA388_Cards (vetor pAZ)	<i>F</i> : 5' -AGAT <u>GGATCC</u> CAAAGACCTACTGATAAACCA-3' <i>R</i> : 5' -CCC <u>AAGCTT</u> <u>TTA</u> GTAACGGACCCAATTTAATAA-3'

**F*: sequências de oligonucleotídeos senso (*forward*);

***R*: sequências de oligonucleotídeos antisenso (*reverse*);

CTCGAG: sequência do sítio de restrição da enzima **XhoI**;

AAGCTT: sequência do sítio de restrição da enzima **HindIII**;

GGATCC: sequência do sítio de restrição da enzima **BamHI**.

LISTA DE ANEXOS

A – CERTIFICADO DA COMISSÃO DE ÉTICA (CEUAIB)

B – SÚMULA CURRICULAR

C – CARTA DE CONFIRMAÇÃO DE PRESIDENTE DO EVENTO BIOLOGY OF SPIROCHETES (GRS)

D – IMPRESSÃO DO SITE REFERENTE AO EVENTO BIOLOGY OF SPIROCHETES (GRS) DE 2022

E – RESEARCH ARTICLE: NEW STRATEGIES FOR *LEPTOSPIRA* VACCINE DEVELOPMENT BASED ON LPS REMOVAL

F – RESEARCH ARTICLE: PHAGOCYTOSIS OF *LEPTOSPIRA* BY LEUKOCYTES FROM MICE WITH DIFFERENT SUSCEPTIBILITY TO LEPTOSPIROSIS AND POSSIBLE ROLE OF CHEMOKINES