

RAQUEL XAVIER LIGEIRO DIAS

REDE DE INTERAÇÃO PROTEÍNA - PROTEÍNA NA RELAÇÃO ENTRE *Rickettsia amblyommatis* E CÉLULAS DO INTESTINO E OVÁRIO DE *Amblyomma sculptum* (BERLESE, 1888)

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

Orientador: Cláudio Lísias Mafra de Siqueira

Coorientador: Tiago Antônio de Oliveira Mendes

**VIÇOSA - MINAS GERAIS
2020**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

D541r
2020
Dias, Raquel Xavier Ligeiro, 1982-
Rede de interação proteína-proteína na relação entre
Rickettsia amblyommatis e células do intestino e ovário de
Amblyomma sculptum (Berlese, 1888)
/ Raquel Xavier Ligeiro Dias.
– Viçosa, MG, 2020.
83 f. : il. (algumas color.) ; 29 cm.

Orientador: Cláudio Lísias Mafra de Siqueira.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Carrapatos. 2. Vetores de doenças. 3. Interações
proteína-proteína - Modelos matemáticos. I. Universidade
Federal de Viçosa. Departamento de Bioquímica e Biologia
Molecular. Programa de Pós-graduação em Bioquímica
Aplicada. II. Título.

CDD 22 ed. 592.42

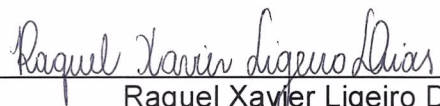
RAQUEL XAVIER LIGEIRO DIAS

REDE DE INTERAÇÃO PROTEÍNA - PROTEÍNA NA RELAÇÃO ENTRE
Rickettsia amblyommatis E CÉLULAS DO INTESTINO E OVÁRIO DE
Amblyomma sculptum (BERLESE, 1888)

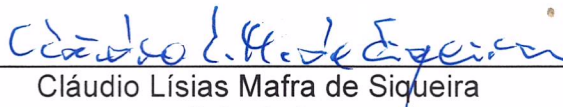
Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Bioquímica Aplicada, para obtenção do título de *Doctor Scientiae*.

APROVADA: 21 de fevereiro de 2020.

Assentimento:



Raquel Xavier Ligeiro Dias
Autora



Cláudio Lísias Mafra de Siqueira
Orientador

Aos meus pais, Roberto e Elizabeth, meu esposo Flávio e minha filhinha Elis.

AGRADECIMENTOS

A Deus, pela dádiva da vida e por me fortalecer diante dos desafios.

A Universidade Federal de Viçosa, por me acolher e proporcionar subsídios para a busca do conhecimento. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

Ao Prof. Dr. Cláudio Lísias Mafra de Siqueira, por todos os anos de orientação que tanto contribuíram para meu aprimoramento como pesquisadora e pessoa.

Ao Prof. Dr. Tiago Antônio de Oliveira Mendes, por toda dedicação e ensinamentos de Bioinformática.

Ao meu esposo, Flávio, por sonhar este sonho comigo, por me acompanhar em todos os momentos sempre me incentivando: sem sua parceria, não conseguiria conciliar todas as atividades com o doutorado.

À minha filhinha Elis, que chegou durante o doutorado para ressignificar minha vida e me permitir experimentar um amor tão puro.

À minha família, que é a base e a força propulsora de todas as minhas conquistas. Em especial, à minha mãe Elizabeth, minha estrela guia no céu a iluminar meu caminho.

Aos professores, alunos e funcionários do Programa de Pós-graduação em Bioquímica Aplicada por me acolherem na instituição de uma forma tão carinhosa, em especial aos doutorandos Gabriel, Fernanda e Flaviane pela parceria no desenvolvimento das pesquisas. Ao Departamento de Tecnologia e Informação, em especial, ao Vinícius, por todo o suporte técnico. Aos amigos pela torcida, e em especial, a Elsa por se fazer tão presente durante esta trajetória.

“Tenho a impressão de ter sido uma criança brincando à beira-mar, divertindo-me em descobrir uma pedrinha mais lisa ou uma concha mais bonita que as outras, enquanto o imenso oceano da verdade continua misterioso diante de meus olhos”.

Isaac Newton

RESUMO

DIAS, Raquel Xavier Ligeiro, D.Sc., Universidade Federal de Viçosa, fevereiro de 2020. **Rede de interação proteína-proteína na relação entre *Rickettsia amblyommatis* e células do intestino e ovário de *Amblyomma sculptum* (Berlese, 1888).** Orientador: Cláudio Lísias Mafra de Siqueira. Coorientador: Tiago Antônio de Oliveira Mendes.

As riquetsias são bactérias intracelulares obrigatórias de hospedeiros vertebrados e invertebrados, diferindo em termos de associação com artrópodes, comportamento dos vetores invertebrados à infecção, na fisiopatologia e desfecho da doença em vertebrados. Entretanto, apesar de saber-se ser o carrapato reservatório e vetor destes microrganismos, os mecanismos moleculares envolvidos na infecção de suas células não foram bem elucidados. Assim, o trabalho teve como objetivo a reconstrução *in silico* da rede de interação proteína-proteína, a partir de proteínas codificadas do genoma da bactéria *Rickettsia amblyommatis* e do carrapato *Ixodes scapularis*. Com o alinhamento das sequências proteicas identificadas, conforme preconizado por diferentes bancos de interações (PSIMAP, iPfam, STRING e PEIMAP), obteve-se uma rede consenso após a normalização dos dados através do cálculo do valor de confiança da interação. Com o auxílio de ferramentas preditivas de localização subcelular, foram selecionadas proteínas de *R. amblyommatis* com alto potencial de interação com o carrapato e de grande importância para o processo infeccioso, sendo identificadas proteínas secretadas, excretadas e, ou, de superfície. Avaliando a rede de proteínas de membrana predita, foram identificadas 3.546 interações, envolvendo 91 proteínas de membrana da bactéria e 980 proteínas do carrapato. Por ontologia gênica foram identificados os processos biológicos significativos envolvidos na modulação da fisiologia do carrapato pela bactéria. Os módulos encontrados revelaram diferentes vias de sinalização, como, p. ex., secreção, fosforilação de aminoácidos e proteínas e alongamento translacional. Com o intuito de validar a rede obtida, foi realizada a integração dos transcriptomas de intestino e ovário de carrapatos *Amblyomma sculptum* em resposta à infecção por *R. amblyommatis*. Com esta estratégia foi possível revelar genes e interações protéicas derivadas de evidências *in silico* que podem auxiliar estudos adicionais quanto aos mecanismos envolvidos na manutenção e transmissão de riquetsias por carrapatos.

Palavras-chave: *Rickettsia amblyommatis*. *Amblyomma sculptum*. Interface vetor-patógeno. Interação proteína-proteína. Modelagem molecular e *docking*.

ABSTRACT

DIAS, Raquel Xavier Ligeiro, D.Sc., Universidade Federal de Viçosa, February, 2020. **Protein-protein interaction network in the relationship between *Rickettsia amblyommatis* and the intestine cells and the ovary of *Amblyomma sculptum* (Berlese, 1888)**. Adviser: Cláudio Lísias Mafra de Siqueira. Co-adviser: Tiago Antônio de Oliveira Mendes.

Rickettsiae are obligatory intracellular bacteria of vertebrate and invertebrate hosts, differing in terms of association with arthropods, behavior of invertebrate vectors to infection, in the pathophysiology and outcome of the disease in vertebrates. However, despite being known to be the tick reservoir and vector of these microorganisms, the molecular mechanisms involved in the infection of its cells have not been well elucidated. Thus, this work aimed the *in silico* reconstruction of the protein-protein interaction network, from proteins encoded in the genome of the bacterium *Rickettsia amblyommatis* and the tick *Ixodes scapularis*. With the alignment of the identified protein sequences, as recommended by different banks of interactions (PSIMAP, iPfam, STRING and PEIMAP), a consensus network was obtained after normalizing the data by calculating the interaction's confidence value. With the aid of predictive tools for subcellular localization, *R. amblyommatis* proteins with high potential for interaction with the tick and of great importance for the infectious process with selected which secreted, excreted and/or surface proteins identified. Evaluating the predicted membrane protein network, 3,546 interactions were identified, involving 91 bacterial membrane proteins and 980 tick proteins. By genetic ontology, significant biological processes involved in the modulation of tick physiology by bacteria were identified. Were found modules revealing different signaling pathways, such as secretion, phosphorylation of amino acids and proteins and translational elongation. In order to validate the network, we integrated the intestinal and ovarian transcriptomes of ticks of *Amblyomma sculptum* in response to infection by *R. amblyommatis*. With this strategy it was possible to reveal genes and protein interactions derived from *in silico* evidence that can help further studies on the mechanisms that involve the maintenance and transmission of *Rickettsia* by ticks.

Keywords: *Rickettsia amblyommatis*. *Amblyomma sculptum*. Vector-pathogen interface. Protein-protein interaction. Modeling and molecular docking.

SUMÁRIO

| | |
|---|----|
| INTRODUÇÃO GERAL..... | 9 |
| CAPÍTULO 1 | |
| REVISÃO DE LITERATURA..... | 12 |
| 1.1. Gênero <i>Rickettsia</i> | 12 |
| 1.2. Riquetsioses | 13 |
| 1.3. Carrapatos como vetores de <i>Rickettsia</i> | 15 |
| 1.4. Interação <i>Rickettsia</i> -carrapato | 18 |
| 1.5. Intestino e ovário: órgãos de interesse para o estudo da relação <i>Rickettsia</i> -carrapato | 20 |
| 1.6. Estudo da relação patógeno-hospedeiro através de redes de interação proteína-proteína..... | 21 |
| 1.7. Referências Bibliográficas..... | 23 |
| CAPÍTULO 2 | |
| DELINEAMENTO EXPERIMENTAL | 29 |
| CAPÍTULO 3 | |
| Rede de interação proteína-proteína na relação entre <i>Rickettsia amblyommatis</i> e células do intestino e ovário de <i>Amblyomma sculptum</i> (BERLESE, 1888)..... | 32 |
| 3.1. INTRODUÇÃO | 34 |
| 3.2. MÉTODOS..... | 36 |
| 3.2.1. Interatoma..... | 36 |
| 3.2.2. Análise do RNA-seq..... | 38 |
| 3.2.3. Modelagem Molecular e <i>Docking</i> | 39 |
| 3.3. RESULTADOS | 41 |
| 3.3.1. Interatoma..... | 41 |
| 3.3.2. Análise do RNA-seq..... | 44 |
| 3.3.3. Modelagem e <i>Docking</i> Molecular | 50 |
| 3.4. DISCUSSÃO | 55 |
| 3.5. REFERÊNCIAS BIBLIOGRÁFICAS | 58 |
| CONSIDERAÇÕES FINAIS | 63 |
| MATERIAL SUPLEMENTAR | 64 |

INTRODUÇÃO GERAL

As riquetsioses se apresentam como doenças febris agudas, cujos sintomas geralmente são inespecíficos, muitas vezes comuns a outras doenças infecciosas, sendo frequente a presença de cefaleias, mialgias e mal estar (FACCINI-MARTÍNEZ et al., 2014; BIGGS et al., 2016). Devido ao número crescente de casos, as pesquisas e a identificação dos agentes riquetsiais patogênicos têm ganhado ênfase nas últimas duas décadas (RAOULT, 2010; EREMEEVA & DASCH, 2015; SANCHEZ-VICENTE et al., 2019).

As infecções são causadas por uma variedade de bactérias do gênero *Rickettsia*. Caracterizando-se como bactérias intracelulares obrigatórias nos hospedeiros vertebrados e invertebrados. Essas bactérias diferem em termos de associação com artrópodes, comportamento do vetor à infecção, fisiopatologia e desfecho da doença (BECHAH et al., 2008). No entanto, o carrapato costuma ser o principal reservatório e vetor de *Rickettsia* pertencente ao grupo da febre maculosa, devido à transmissão transtadial e transovariana (REHÁČEK et al., 1989; SOCOLOVSCHI et al., 2009; TOMASSONE et al., 2018). Os carrapatos podem adquirir a bactéria alimentando-se em hospedeiro riquetsêmico, co-alimentação ou transmissão sexual (MCDADE; NEWHOUSE, 1986; SOCOLOVSCHI et al., 2009). Ao analisar as associações entre carrapato e bactérias do gênero *Rickettsia*, observa-se que o nível de especificidade do vetor varia conforme as espécies de *Rickettsia*, a saber: *Rickettsia cororii* com a espécie *Rhipicephalus sanguineus*, *Rickettsia slovaca* com duas espécies pertencentes ao mesmo gênero *Dermacentor marginatus* e *Dermacentor reticulatus* e *Rickettsia rickettsii* com carrapatos de diferentes gêneros (PAROLA, 2013). *Rickettsia amblyommatis* prevalente em *Amblyomma americanum* e em outros *Amblyomma* spp (S. SANTIBÁÑEZ et al., 2017).

Como parte de seu estilo de vida intracelular obrigatório, os genomas de *Rickettsia* passaram por uma evolução que inclui perda de genes, duplicação de repetição de DNA e transferência horizontal de genes que moldaram a evolução do seu genoma. A perda gênica envolveu diversas vias metabólicas, dentre as quais, biossíntese de aminoácidos, ATP, lipossacarídeos e componentes da parede celular e reguladores transcricionais, mantendo, entretanto uma alta preservação dos módulos toxina-antitoxina, recombinação e proteínas de reparo do DNA (NEBBAK et

al., 2019). Tal mudança no perfil genômico permite que a bactéria manipule os mecanismos moleculares e celulares dos hospedeiros através de interações moleculares, a fim de se aproveitar da maquinaria das células hospedeiras e propiciar suas infecções. Neste processo, temos as interações moleculares entre agentes patogênicos e seus hospedeiros como fundamentais para iniciar e manter o processo de infecção (McDADE & NEWHOUSE, 1986; NARRA et al., 2020).

Ao considerar que é essencial à sobrevivência bacteriana infectar células do carrapato, compreender a biologia da interação *Rickettsia*-carrapato em nível molecular é um grande desafio, visto o envolvimento de maquinarias celulares dinâmicas, diversos e diferentes complexos proteicos e a existência de um sistema de regulação sensível e robusto. Desta maneira, a estratégia de reconstrução de redes de interação proteína-proteína apresenta-se como uma ferramenta importante para elucidar os mecanismos moleculares e funcionais que caracterizam a replicação eficiente do patógeno e a sua manutenção em longo prazo no ambiente celular, bem como sua transmissão.

Assim, o presente estudo teve por objetivo reconstruir a rede de interação proteína-proteína a partir dos genomas de *Rickettsia amblyommatis* e de *Ixodes scapularis* e integrar aos transcriptomas de intestino e ovário de *Amblyomma sculptum* em resposta a infecção por *Rickettsia amblyommatis*, abordando os aspectos moleculares e funcionais envolvidos nesta interface.

REFERÊNCIAS BIBLIOGRÁFICAS

- BECHAH, Y. et al. Rickettsial diseases: from *Rickettsia*–arthropod relationships to pathophysiology and animal models. **Future Microbiology**, v. 3, n. 2, p. 223-236, 2008.
- BIGGS, H. M. et al. Diagnosis and management of tickborne rickettsial diseases: Rocky Mountain spotted fever and other spotted fever group rickettsioses, ehrlichioses, and anaplasmosis-United States: A practical guide for health care and public health professionals. **Morbidity and Mortality Weekly Report: Recommendations and Reports**, v. 65, n. 2, p. 1-44, 2016.
- EREMEEVA, M. E.; DASCH, G. A. Challenges posed by tick-borne rickettsiae: eco-epidemiology and public health implications. **Frontiers in Public Health**, v. 3, n. 55, p. 55, 2015.
- FACCINI-MARTÍNEZ, Á. A. et al. Syndromic classification of rickettsioses: an approach for clinical practice. **International Journal of Infectious Diseases**, v. 28, p. 126-139, 2014.
- MCDADE, J. E.; NEWHOUSE, V. F. Natural history of *Rickettsia rickettsii*. **Annual Reviews in Microbiology**, v. 40, n. 1, p. 287-309, 1986.
- NARRA, H.P. et al. Recent research milestones in the pathogenesis of human rickettsiosis and future opportunities. **Microbiology of the Future**, v. 15, p. 753-765, 2020.
- NEBBAK, A. et al. Co-infection of bacteria and protozoan parasites in *Ixodes ricinus* nymphs collected in the Alsace region, France. **Ticks and Tick-borne Diseases**, v. 10, n. 6, p. 101241, 2019.
- PAROLA, P. et al. Update on tick-borne rickettsioses around the world: a geographic approach. **Clinical Microbiology Reviews**, v. 26, n. 4, p. 657-702, 2013.
- RAOULT, D. Emerging rickettsioses reach the United States. **Clinical Infectious Diseases**, v. 51, n. 1, p. 121-122, 2010.
- REHÁČEK, J. Ecological relationships between ticks and rickettsiae. **European Journal of Epidemiology**, v. 5, p. 407-413, 1989.
- SANTIBÁÑEZ, S. et al. Isolation of *Rickettsia amblyommatis* in HUVEC line. **New microbes and new infections**, v. 21, p. 117–121, 2017.
- SOCOLOVSCHI, C. et al. Transovarial and trans-stadial transmission of *Rickettsiae africae* in *Amblyomma variegatum* ticks. **Clinical Microbiology and Infection**, v. 15, p. 317-318, 2009.

CAPÍTULO 1

REVISÃO DE LITERATURA

1.1. Gênero *Rickettsia*

Rickettsiales é uma ordem de bactérias gram negativas classificada em três famílias: Anaplasmataceae, Rickettsiaceae, Holosporaceae. A maioria das bactérias pertencentes às famílias Anaplasmataceae e Rickettsiaceae vivem em associação com hospedeiros de artrópodes, sendo as quem compõem a família Holosporaceae, em sua maioria, simbioses protistas (DUMLER et al., 2001). O gênero *Rickettsia*, pertence à família Rickettsiaceae e compreende bactérias parasitas intracelulares obrigatórias devido à perda de vias metabólicas e sintéticas necessárias para a vida extracelular (AMMERMAN & BEIER-SEXTON, 2008). Essas bactérias são descritas principalmente como cocobacilos e conhecidas por causar graves doenças em humanos e outros mamíferos (YU & WALKER, 2006; BOUDESA & PAROLA, 2007).

Todas as espécies de *Rickettsia* associadas a vertebrados são vetorizadas por artrópodes (carrapatos, pulgas, piolhos e ácaros) como parte do seu ciclo de vida, com várias encontradas exclusivamente em artrópodes ainda sem hospedeiro secundário conhecido (WEINERT et al., 2009). Multiplicando-se por fissão binária, as riquetsias colonizam predominantemente glândulas salivares e ovários de artrópodes hospedeiros (RAOULT & ROUX, 1997; BOUDESA & PAROLA, 2007), com algumas espécies sendo descritas infectando e se multiplicando em células do intestino, tubos de Malpighi e hemolinfa (YU & WALKER, 2006).

Fundamentados em padrões morfológicos, moleculares, antigênicos e ecológicos, as espécies de *Rickettsia* são classificadas em três grupos: Grupo Ancestral (GA), agrupando as espécies associadas a carrapatos e algumas associadas a anelídeos (sanguessuga); Grupo Tifo (GT); compreendendo aquelas associadas com piolhos e pulgas; e o Grupo das Febres Maculosas (GFM), com espécies primariamente associadas a carrapatos, com exceções de algumas bactérias associadas a pulgas e ácaros (PAROLA et al., 2009).

Diversas doenças infecciosas causadas por *Rickettsia*, conhecidas como riquetsioses foram identificadas em todo o mundo, inclusive apresentando caráter endêmico, representando risco variado à saúde humana e de animais (PAROLA et al., 2013). No Brasil, a riquetsiose mais conhecida é a Febre Maculosa Brasileira, causada por riquetsias do GFM (LABRUNA et al., 2014). No período entre 2007 –

2013, foram confirmados 807 casos de febre maculosa pelo Ministério da Saúde (BRASIL, 2014a). Além disso, foram registradas 411 mortes entre 2007 e 2015, principalmente na região sudeste do país, onde a taxa de letalidade foi de 55% (DE OLIVEIRA et al., 2016). Esta situação, agravada a partir dos anos 1980, levou a inserção dessa riquetsiose na lista de agravos de saúde de notificação compulsória, o que permitiu um melhor monitoramento desta enfermidade pelas autoridades, podendo desta maneira, antever possíveis surtos devido a estes agentes (BRASIL, 2001; BRASIL, 2014b).

1.2. Riquetsioses

Como consequência do número crescente de casos das riquetsioses, as pesquisas e identificação dos agentes patogênicos do gênero ganhou ênfase nas últimas décadas (RAOULT, 2010; EREMEEVA & DASCH, 2015). As espécies de *Rickettsia* mantêm ciclos enzoótico (silvestre) e zoonótico (urbano e rural) na natureza, como muitas destas sendo agentes com potencial zoonoses transmitidas especialmente por carrapatos, piolhos, pulgas, ácaros e outros artrópodes. Na maioria dos casos, acredita-se que os humanos sejam hospedeiros acidentais destes microorganismos (FACCINI-MARTÍNEZ et al., 2014).

Apesar de sua baixa morbidade, as riquetsioses caracterizam-se como agravos de saúde com elevada mortalidade, sendo um sério problema de saúde pública em todo o mundo, requerendo assim, grande atenção por parte das autoridades e profissionais de saúde. Faz-se fundamental considerar suspeita de infecção por *Rickettsia* quando os pacientes apresentam seus mais frequentes sinais clínicos, laboratoriais e sintomas, diversos destes comuns a outros quadros infecciosos, em especial aqueles caracterizados por febres hemorrágicas, sejam virais ou bacterianas. Mesmo sendo um importante parâmetro a ser considerado quando da avaliação do caso, sua inclusão dentre o conjunto de enfermidades a se realizar o diagnóstico diferencial, independe do relato de exposição a carrapatos ou outros artrópodes, uma vez que o contato com o vetor pode não ter sido percebido ou ser lembrado. Infecções por *Rickettsia* são mais comuns nos períodos mais quentes e em indivíduos expostos às atividades ao ar livre, devido ao possível contato com carrapatos, piolhos e pulgas (SNOWDEN & KIN, 2019).

Os artrópodes podem adquirir as riquetsias por dois mecanismos: transmissão transtadial (transmissão horizontal) e transmissão transovariana (transmissão vertical). A transmissão transtadial ocorre quando um patógeno permanece com o vetor de um estágio do ciclo de vida para o próximo, começando com o consumo de uma refeição sanguínea de um animal hospedeiro infectado. E a transmissão transovariana, ocorre pela transmissão de patógenos causadores de doenças, do artrópode parental, para o artrópode da prole, isto é, da mãe infectada para sua progênie (RAOULT et al., 2005). A transmissão para humanos e animais é feita pelo vetor diretamente através de uma picada ou através da inoculação na pele ou mucosa de material fecal infectado (RICHARDS, 2012).

Os mecanismos moleculares referentes à patogenicidade das riquetsias ainda precisam ser estabelecidos com precisão (UCHIYAMA, 2012). Entretanto, é de conhecimento, que no interior do organismo a bactéria infecta prioritariamente as células endoteliais que revestem os pequenos e médios vasos sanguíneos, causando sintomas sistêmicos e alta mortalidade. A agressão endotelial gera inflamação disseminada, com perda da função da barreira e alteração da permeabilidade vascular em todo o corpo (MAHAJAN, 2012). Os mecanismos envolvidos na rápida entrada da bactéria na célula e na regulação negativa das vias imunológicas, permitindo a persistência da infecção, estão sendo estudados com o intuito de identificar novas fórmulas terapêuticas eficazes e assim, reduzir os agravos provocados por essas infecções (SNOWDEN & KIN, 2019).

O período de incubação das riquetsioses costuma ser de quatro a dez dias, com a doença apresentando quadros febris agudos, cujos sintomas geralmente são inespecíficos e muitas vezes, comuns a outras doenças infecciosas, sendo frequente a presença de cefaleias, mialgias, e mal estar. Outros sintomas que podem estar presente são: linfadenopatia, alterações do sistema nervoso central, como confusão ou rigidez nuchal, escara no local da inoculação, hepatite, vômito e instabilidade cardiovascular (FACCINI-MARTÍNEZ et al., 2014; BIGGS et al., 2016).

Alguns hábitos estão associados com relatos de riquetsioses, como trabalho em zona rural ou moradia de risco em região periurbana. Esses fatores aumentam a possibilidade de contato de seres humanos com reservatórios naturais (capivaras, gambás e outros roedores), tendo hospedeiros amplificadores (cavalos e cães) como

papel importante na disseminação dos carrapatos infectados com riquetsias do ciclo enzoótico para o zoonótico (ANGERAMI et al., 2012).

1.3. Carrapatos como vetores de *Rickettsia*

Os carrapatos são ectoparasitas e vetores biológicos de múltiplos patógenos humanos e animais, capazes de transmitir uma maior diversidade de agentes infecciosos do que qualquer outro grupo de artrópodes. As doenças transmitidas por carrapatos resultam em uma morbidade significativa com milhares de mortes de humanos e animais anualmente (GULIA-NUSS et al., 2016). Entre os patógenos transmitidos por carrapatos, os pertencentes ao gênero *Rickettsia*, são os principais, ocasionando surtos em todo mundo (JONES et al., 2008).

Várias espécies de carrapatos já foram descritas como hospedeiros invertebrados e reservatórios destes organismos, as quais se encontram distribuídas em três famílias: Ixodidae, Argasidae e Nuttalliellidae (GUGLIELMONE et al., 2010). Os representantes das famílias Ixodidae e Argasidae são conhecidos como carrapatos duros e moles, respectivamente, apresentando importância para a saúde pública. A família Nuttalliellidae é monotípica e restrita à África do Sul e Tanzânia (BEDFORD, 1931; KEIRANS et al., 1976). As famílias Ixodidae e Argasidae têm hábitos hematófagos, alimentando-se de uma grande variedade de vertebrados terrestres, incluindo humanos, animais domésticos e silvestres (MASTROPAOLO et al., 2014; SPONCHIADO et al., 2015). Os argasídeos são carrapatos que durante seu ciclo de vida se alimentam mais de uma vez. As fêmeas efetuam várias posturas, alternando oviposição com repasto sanguíneo. Cada postura pode variar de 50 a 600 ovos, números significativamente inferiores quando comparado aos ixodídeos (aproximadamente 23.000 ovos). Outra diferença entre estas duas famílias é que nos ixodídeos, a alimentação dura um tempo maior e é realizada uma única vez em todo o seu ciclo de vida (KLOMPEN, 2005).

O gênero *Ixodes*, pertencente à família Ixodidae, inclui múltiplas espécies de importância médica e veterinária, servindo principalmente como vetores de agentes como: *Rickettsia helvetica*, *Rickettsia australis*, *Borrelia burgdorferi*, *Anaplasma phagocytophilum*, *Borrelia miyamotoi*, *Babesia microti* e Powassan virus (NELDER et al., 2016). No Brasil, até o momento, a fauna ixodídica é representada por 71 espécies (SPRONG et al., 2009; MUÑOZ-LEAL et al., 2017).

Em meio os tipos de vetores artrópodes, os carrapatos são os mais comuns na transmissão de riquetsioses (PAROLA et al., 2013). Destas, as febres maculosas são as de maior impacto e incidência mundial. Essas zoonoses estão entre as mais antigas doenças transmitidas por vetores conhecidas e inclui a Febre Maculosa das Montanhas Rochosas (FMMR), principal riquetsiose, tendo a *Rickettsia rickettsii* como agente etiológico e apresentando distribuição vasta na América (WILSON & CHOWNING, 1902; CHILDS et al., 2007). Desde o século passado os carrapatos têm sido associados à FMMR, tendo início quando estudos relacionaram casos da doença a relatos de pacientes picados por carrapato. A partir de então, pesquisas foram realizadas buscando a comprovação da importância dos carrapatos na transmissão de doenças a seres humanos. Em consequência, o gênero *Rickettsia* foi mais bem elucidado (WILSON & CHOWNING, 1902; PAROLA et al., 2009). No Brasil, a Febre Maculosa Brasileira (FMB), também causada pela *R. rickettsii*, é a mais patogênica entre outras riquetsioses dentro do GFM, causando problemas severos à saúde pública (PAROLA et al., 2009).

No GFM existem cerca de 22 espécies de *Rickettsia* conhecidas, dessas, algumas estão relacionadas a transmissão por carrapatos vetores: FMMR e FMB (*R. rickettsii*), febre maculosa do mediterrâneo ou febre botonosa (*R. conorii*), febre africana da picada do carrapato (*R. africae*), tifo do carrapato de Queensland (*R. australis*), tifo da Ilha Flinders (*R. honei*), tifo siberiano ou do norte da Ásia (*R. sibirica*), febre maculosa oriental (*R. japonica*), tifo (*R. slovaca*), linfangite expansiva (*R. sibirica mongolotimonae*), riquetsiose europeia (*R. helvetica*), febre maculosa de Parker (*R. parkeri*), *R. amblyommatidis*, dentre outras (MARTINS et al., 2014).

Os carrapatos transmissores variam de acordo com a região e o organismo, tendo como exemplo: *Dermacentor variabilis* (carrapato americano), *Dermacentor andersoni* (carrapato de madeira das montanhas rochosas), *Amblyomma americanum* (carrapato de estrela solitária) associados à maioria dos casos de febre maculosa nos Estados Unidos; espécies do complexo *Amblyomma cajennense* tem sido associado à febre maculosa na América do Sul; e *Amblyomma hebraeum* ou *Amblyomma variegatum* na África do Sul (SNOWDEN & KIN, 2019).

Estudos demonstraram que carrapatos do gênero *Amblyomma* são alguns dos responsáveis pela manutenção das riquetsias na natureza (AZAD, 1988). Muitas

espécies do gênero *Rickettsia* são consideradas como simbiontes de invertebrados, transmitidas verticalmente, o que sugere que os artrópodes vetores funcionam como reservatórios ou amplificadores destes organismos na natureza (JONGEJAN & UILENBERG, 2004). Dentre as espécies, destaca-se o *Amblyomma sculptum*, por ser uma espécie de relevância epidemiológica e zoonótica em uma extensa região do continente americano, do Sul dos Estados Unidos até o Norte da Argentina (NAVA et al., 2014; RAMOS et al., 2014). No Brasil, o *Amblyomma sculptum* (Complexo *A. cajennense*) é o carrapato que apresenta maior incidência de parasitismo em humanos, sendo considerado o principal vetor da Febre Maculosa Brasileira, um agravo de notificação compulsória de elevada letalidade (NUNES et al., 2008).

Na interação carrapato-hospedeiro, o processo de detecção do hospedeiro pelo vetor requer um aparato sensorial peculiar composto por sensilas (quimiossensila, mecanossensila, fotossensila e termossensila) que monitoram o ambiente externo e interno. A quimiorrecepção ocorre presumivelmente através do órgão de Haller, localizado no dorso do tarso no primeiro par de patas, sendo importante para a identificação do potencial hospedeiro (SONENSHINE & ROE, 2013; CARR & ROE, 2016; GULIA-NUSS et al., 2016). Estudando-se a anotação do genoma de *I. scapularis*, verificou-se que aproximadamente 57% das sequências codificadoras de proteínas (20.480 sequências) estão associadas a interações do carrapato e seus hospedeiros. Dentre eles, genes importantes para a detecção e permanência no hospedeiro, alimentação prolongada, síntese de cutícula, digestão da hemoglobina, desintoxicação e vitelogênese (GULIA-NUSS et al., 2016).

O ciclo de vida do carrapato inclui longos períodos de desenvolvimento no hospedeiro e hematofagia, bem como semanas ou meses que o carrapato passa no ambiente externo, fora do hospedeiro, sem alimentação. Inicialmente, o carrapato cria uma cavidade e injeta sua saliva no hospedeiro. Produto das glândulas salivares a saliva do carrapato contém uma mistura complexa de peptídeos e outros compostos diretamente envolvidas na manutenção do equilíbrio hidrodinâmico do carrapato, produzindo e secretando inúmeras substâncias que atuam desde a facilitação de sua fixação na pele do hospedeiro vertebrado, bem como na modulação de respostas imune e hemostática deste último, limitam sua resposta hemostática, inflamatória e imunológica durante a hematofagia

(FRANCISCHETTI et al., 2009; SONENSHINE & ROE, 2013). Peptídeos antimicrobianos, lipocalinas, metaloproteinases e proteínas de domínio Kunitz são apenas alguns dos compostos presentes na saliva dos carrapatos. Além desses compostos, destacam-se também proteínas ricas em resíduos de glicina que auxiliam na fixação do hipostômio do parasito ao sítio de alimentação na pele do hospedeiro, além de atuarem como principal via de transmissão de riquetsias (MARUYAMA et al., 2010).

1.4. Interação *Rickettsia*-carrapato

A relação *Rickettsia*-carrapato tem sido alvo de diversos estudos, em especial quanto ao papel epidemiológico dos carrapatos como vetores. Entretanto, há muito se fala que pouca atenção tem sido dirigida para elucidar o envolvimento entre *Rickettsia* e as células, tecidos e órgãos dos carrapatos (BEDFORD, 1931). Atualmente, tem-se que para a escolha de novos alvos moleculares no desenvolvimento de medidas de controle para carrapatos e doenças associadas, esta deverá ser baseada no entendimento das interações moleculares entre hospedeiros, carrapatos e agentes patogênicos. Desta maneira, identificando-se e caracterizando-se as interações que promovam a sobrevivência, a disseminação e a transmissão de patógenos pelo carrapato, desenvolver novas estratégias de controle e prevenção para as riquetsioses (DE LA FUENTE et al., 2016).

Apesar de todos os esforços neste sentido, compreender as interações envolvidas na relação entre o patógeno e o vetor a nível molecular ainda é um grande desafio, visto envolver maquinarias celulares dinâmicas, diferentes complexos proteicos e um sistema de regulação sensível e robusto (RAMAPRASAD et al., 2012). Nos últimos anos, tal estratégia de estudo tem sido favorecida em decorrência dos avanços no sequenciamento de nova geração e o rápido desenvolvimento de bancos de dados genômicos para muitos agentes patogênicos e seus vetores, o que possibilita a aplicação da biologia de sistemas para o estudo das interações parasita-hospedeiro. Ao longo da última década, essas abordagens levaram à descoberta de muitos processos biológicos importantes que governam a doença parasitária, proporcionando informações necessárias para impulsionar pesquisas para a descoberta de drogas e acelerar o desenvolvimento de novas terapias antiparasitárias (JONES et al., 2008). Assim, a busca pelo entendimento das

relações moleculares entre patógenos e seus vetores podendo gerar conhecimento para desenvolver novas estratégias, mais precisas, para o controle de carrapatos e das riquetsioses.

Sabe-se que a coevolução de riquetsias com artrópodes é responsável por muitas características da relação patógeno-vetor, incluindo a replicação eficiente do patógeno, sua manutenção em longo prazo da infecção e transmissão transestadial e transovariana (GIMÉNEZ, 1964). Capazes de se multiplicar em células dos intestinos, ovários, glândulas salivares, túbulos de Malpighi e hemolinfa, tais bactérias penetram nas células hospedeiras de invertebrados através de receptores específicos de membrana (GUTIÉRREZ et al., 2014), podem também se multiplicar no interior das células epiteliais de vasos sanguíneos de seus hospedeiros vertebrados por divisão binária. Mantidas no carrapato na ausência de hematofagia por sucessivas gerações, tem sua virulência reativada por fatores nutricionais presentes no sangue de hospedeiros vertebrados, ou quando são expostas à temperatura adequada (em torno de 28°C) (WALKER, 1989; RAOULT & ROUX, 1997).

Estudos evidenciaram que as bactérias do gênero *Rickettsia* desenvolveram várias estratégias para se adaptar a diferentes condições ambientais, incluindo aqueles dentro de seus vetores artrópodes e hospedeiros vertebrados (SOCOLOVSKI et al., 2009). Em relação aos aspectos moleculares, estudos indicam que a temperatura pode modular o perfil transcriptômico de *Rickettsia* em carrapatos (DREHER-LESNICK et al., 2008; ELLISON et al., 2009; GALLETTI et al., 2013) e em culturas de células (POLICASTRO et al., 1997).

Assim, identificar genes diferencialmente expressos na interação *Rickettsia*-carrapato é fundamental para obter informações importantes na elucidação dos mecanismos de virulência deste patógeno (CERAUL et al., 2011). Estudos evidenciam que a infecção de carrapatos por bactérias modula a sua expressão gênica, sendo alguns genes diferencialmente expressos com a finalidade de controlar a infecção no vetor e a sua transmissão para o hospedeiro vertebrado (ZIVKOVIC et al., 2010).

Em estudos realizados por Dreher-Leshnick et al. (2006), foi observada indução pronunciada do gene da GTS (gene associado ao estresse) em carrapatos *Dermacentor variabilis* infectados por *Rickettsia montanensis*. Como a infecção por

riquétias ocorre principalmente pela ingestão de sangue infectado, estes autores sugeriram a existência de relação entre esse gene e a resposta contra a infecção. Posteriormente, como demonstrado por Ceraul et al. (2011), verificou-se que um inibidor tipo Kunitz apresenta atividade bacteriostática contra *R. montanensis*.

1.5. Intestino e ovário: órgãos de interesse para o estudo da relação *Rickettsia*-carrapato

O sucesso da ação do carrapato sobre uma variedade de hospedeiros vertebrados e, por conseguinte, a transmissão de patógenos com potencial zoonótico está intimamente relacionada à sua morfologia, bem como com a fisiologia dos seus órgãos internos, sobretudo glândulas salivares, intestino e ovário. Desta maneira, um melhor entendimento dos processos moleculares envolvidos na interação patógeno-vetor nestes órgãos, vem a se postar como etapa fundamental para a compreensão da biologia dos carrapatos, bem como dos eventos moleculares que regem seu papel como agente vetor de *Rickettsia* (GULIA-NUSS et al., 2016; MOREIRA et al., 2017).

Os carrapatos são considerados os vetores de doenças mais relevantes em animais domésticos e selvagens uma vez que se alimentam de sangue, sendo por sua alimentação que ocorre sua infecção, bem como a transmissão de patógenos (GODDARD, 2009; ANDREOTTI et al., 2011). Estudos têm mostrado que o intestino de carrapatos e de outros artrópodes atuam como um ponto de entrada de diversas bactérias se mostrando importante na sobrevivência do patógeno (MARTIN & SCHMIDTMANN, 1998; ANDREOTTI et al., 2011; BOISSIÈRE et al., 2012; BRODERICK & LEMAITRE, 2012). O entendimento da microbiota intestinal de artrópodes que atuam como vetores pode contribuir no que diz respeito ao controle e prevenção de doenças transmitidas por vetores.

Ao examinar o microbiota intestinal do carrapato *I. scapularis*, parasitas de camundongos, Narasimhan et al. (2014), observaram que a bactéria *Borrelia burgdorferi*, agente da doença de Lyme, entra no intestino do artrópode, ainda na fase larval, junto com a alimentação. A seguir, a larva infectada se desenvolve em ninfa e se alimenta do seu hospedeiro vertebrado, o que permite que a *B. burgdorferi* migre do intestino para as glândulas salivares e seja transmitido ao hospedeiro junto com a saliva do carrapato (NARASIMHAN et al., 2014).

Bactérias como as do gênero *Rickettsia*, além de se alojarem nas células que revestem o intestino do carrapato, também podem invadir outros órgãos e tecidos, como por exemplo, ovários (HARDSTONE YOSHIMIZU & BILLETER, 2018). Considerado o órgão mais importante na associação carrapato -*Rickettsia*, o ovário do carrapato desempenha papel fundamental na transmissão transovariana deste patógeno no ambiente (SOARES et al., 2012). Uma vez que ocorre a infecção dos ovos pela bactéria, tem-se a infecção da prole, resultando na proliferação e manutenção de carrapatos infectados no ambiente (LABRUNA et al., 2002).

Ensaio demonstraram que a infecção por *R. amblyommatis* em ovários de carrapatos *A. sculptum* resultou, de forma direta e indiretamente, na diminuição da expressão de muitos genes, sendo destes *housekeeping*, envolvidos na síntese de proteínas, especialmente aminoacil-tRNAs; no metabolismo de carboidratos; e na síntese de componentes do citoesqueleto (MAFRA, C. Dados não publicados).

1.6. Estudo da relação patógeno-hospedeiro através de redes de interação proteína-proteína

Os patógenos manipulam os mecanismos moleculares e celulares dos hospedeiros através de interações moleculares, a fim de aproveitar a maquinaria das células hospedeiras e propiciar as infecções. As interações moleculares entre agentes patogênicos e seus hospedeiros são fundamentais para iniciar e manter o processo de infecção (DURMUS, et al., 2015). Ao contrário de uma abordagem tradicional que considera separadamente o hospedeiro ou o patógeno, uma abordagem baseada na biologia de sistemas pode fornecer uma compreensão mais completa das possíveis interações patógeno-hospedeiro (NOURANI, et al., 2015). Os métodos computacionais utilizam principalmente informações de sequência, estrutura proteica e interações para análise e previsão da rede de interação proteína-proteína (DURMUS, et al., 2013). Por meio de uma abordagem integrativa através das redes de interação proteína-proteína é possível desvendar as relações moleculares complexas nos sistemas vivos, ou seja, mapear as interações físicas de proteínas para proteínas. As proteínas funcionam não apenas como entidades únicas, mas exibem seus papéis interagindo com outros componentes celulares, como é visto na sinalização e transdução de sinal, arranjo e reorganização do

citoesqueleto, reconhecimento biológico do sistema imunológico, regulação enzimática e expressão gênica (OTTMAN, 2013).

Nos carrapatos, as análises do transcriptoma identificaram proteínas que atuam em quase todas as últimas vias, mas não foram realizados estudos para validar as vias e para entender melhor as interações proteína-proteína que ocorrem nessas vias. Sem essa informação, nossa compreensão da biologia do carrapato e dos alvos moleculares que podem ser avaliados permanece ineficaz (DYER, et al., 2010). Dessa forma, novas abordagens necessitam ser repensadas com base nas interações moleculares entre hospedeiros, carrapatos e agentes patogênicos no intuito de descobrir moléculas-alvo que poderiam ser testadas em vacinas ou terapias para a intervenção dos ciclos de carrapato-patógeno (FUENTE, et al., 2016).

1.7. Referências Bibliográficas

- AMMERMAN, N. C.; BEIER-SEXTON, M.; AZAD, A. F. Laboratory maintenance of *Rickettsia rickettsii*. **Current Protocols in Microbiology**, v. 11, n. 1, p. 3A. 5.1-3A. 5.21, 2008.
- ANDREOTTI, R. et al. Assessment of bacterial diversity in the cattle tick *Rhipicephalus (Boophilus) microplus* through tag-encoded pyrosequencing. **BMC Microbiology**, v. 11, n. 1, p. 6, 2011.
- ANGERAMI, R. N. et al. Features of Brazilian spotted fever in two different endemic areas in Brazil. **Ticks and Tick-borne Diseases**, v. 3, n. 5-6, p. 346-348, 2012.
- AZAD, A. F. Relationship of vector biology and epidemiology of louse-and flea-borne rickettsioses. **Biology of Rickettsial Diseases**, v. 1, p. 51-62, 1988.
- BEDFORD, G. A. H. *Nuttalliella namaqua*, a new genus and species of tick. **Parasitology**, v. 23, n. 2, p. 230-232, 1931.
- BIGGS, H. M. et al. Diagnosis and management of tickborne rickettsial diseases: Rocky Mountain spotted fever and other spotted fever group rickettsioses, ehrlichioses, and anaplasmosis-United States: A practical guide for health care and public health professionals. **Morbidity and Mortality Weekly Report: Recommendations and Reports**, v. 65, n. 2, p. 1-44, 2016.
- BOISSIÈRE, A. et al. Midgut microbiota of the malaria mosquito vector *Anopheles gambiae* and interactions with *Plasmodium falciparum* infection. **PLoS Pathogens**, v. 8, n. 5, 2012.
- BOUDESA, A.; PAROLA, P. *Rickettsia*. **Revue Francophone des Laboratoires**, v. 2007, n. 391, p. 23-32, 2007.
- BRASIL. Ministério da Saúde/SVS. Sistema de Informação de Agravos de Notificação - Sinan Net. FEBRE MACULOSA - Casos confirmados notificados no Sistema de Informação de Agravos de Notificação - SinanNet. Disponível em: <http://dtr2004.saude.gov.br/sinanweb/tabnet/dh?sinannet/fmaculosa/bases/febremaculosabrnet.def>. Acesso em: 02/02/2020.
- BRASIL. Ministério da Saúde. Gabinete do Ministro. Portaria nº 1.271, de 6 de junho de 2014. Define a Lista Nacional de Notificação Compulsória de doenças, agravos e eventos de saúde pública nos serviços de saúde públicos e privados em todo o território nacional, nos termos do anexo, e dá outras providências. **Diário Oficial da União**, Brasília, DF, 06 jun. 2014.
- BRASIL. Ministério da Saúde. Gabinete do Ministro. Portaria nº 1943/GM, de 18 de outubro de 2001. Define a relação de doenças de notificação compulsória para todo território nacional. Ministério da Saúde. **Diário Oficial da União**, Brasília, DF, 24 out. 2001.
- BRODERICK, N. A.; LEMAITRE, Bruno. Gut-associated microbes of *Drosophila melanogaster*. **Gut Microbes**, v. 3, n. 4, p. 307-321, 2012.

CARR, A. L.; ROE, M. *Acarine attractants: chemoreception, bioassay, chemistry and control*. **Pesticide Biochemistry and Physiology**, v. 131, p. 60-79, 2016.

CERAUL, S. M. et al. A Kunitz protease inhibitor from *Dermacentor variabilis*, a vector for spotted fever group rickettsiae, limits *Rickettsia montanensis* invasion. **Infection and Immunity**, v. 79, n. 1, p. 321-329, 2011.

CHILDS, J. E.; PADDOCK, Christopher D. Rocky Mountain spotted fever. In: **Rickettsial Diseases**. CRC Press, p. 109-128, 2007.

DE LA FUENTE, J. et al. Strategies for new and improved vaccines against ticks and tick-borne diseases. **Parasite Immunology**, v. 38, n. 12, p. 754-769, 2016.

DE OLIVEIRA, S. V. et al. An update on the epidemiological situation of spotted fever in Brazil. **Journal of Venomous Animals and Toxins Including Tropical Diseases**, v. 22, n. 1, p. 1-8, 2016.

DREHER-LESNICK, S. M. et al. Differential expression of two glutathione S-transferases identified from the American dog tick, *Dermacentor variabilis*. **Insect Molecular Biology**, v. 15, n. 4, p. 445-453, 2006.

DREHER-LESNICK, S. M. et al. Genome-wide screen for temperature-regulated genes of the obligate intracellular bacterium, *Rickettsia typhi*. **BMC Microbiology**, v. 8, n. 1, p. 61, 2008.

DUMLER, J. S. et al. Reorganization of genera in the families *Rickettsiaceae* and *Anaplasmataceae* in the order *Rickettsiales*: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. **International Journal of Systematic and Evolutionary Microbiology**, v. 51, n. 6, p. 2145-2165, 2001.

DYER, M. D. et al. The human-bacterial pathogen protein interaction networks of *Bacillus anthracis*, *Francisella tularensis*, and *Yersinia pestis*. **PLoS One**, v. 5, n. 8, p. e12089, 2010.

DURMUS, T. et al. Systems biology of pathogen-host interaction: networks of protein-protein interaction within pathogens and pathogen-human interactions in the post-genomic era. **Biotechnology Journal**, v. 8, n. 1, p. 85-96, 2013.

DURMUS, S. et al. A review on computational systems biology of pathogen-host interactions. **Frontiers in Microbiology**, v. 6, n. 235, p. 1-19, 2015.

ELLISON, D. W. et al. Limited transcriptional responses of *Rickettsia rickettsii* exposed to environmental stimuli. **PLoS One**, v. 4, n. 5, 2009.

EREMEEVA, M. E.; DASCH, G. A. Challenges posed by tick-borne rickettsiae: eco-epidemiology and public health implications. **Frontiers in Public Health**, v. 3, p. 55, 2015.

FACCINI-MARTÍNEZ, Á. A. et al. Syndromic classification of rickettsioses: an approach for clinical practice. **International Journal of Infectious Diseases**, v. 28, p. 126-139, 2014.

FRANCISCHETTI, I. M. B et al. The role of saliva in tick feeding. **Frontiers in Bioscience: a Journal and Virtual Library**, v. 14, p. 2051, 2009.

FUENTE, J. et al. Strategies for new and improved vaccines against ticks and tick-borne diseases. **Parasite Immunology**, v.38, p.754-769, 2016.

GALLETTI, Maria F. B. M. et al. Natural blood feeding and temperature shift modulate the global transcriptional profile of *Rickettsia rickettsii* infecting its tick vector. **PLoS One**, v. 8, n. 10, 2013.

GIMÉNEZ, D. F. Staining rickettsiae in yolk-sac cultures. **Stain Technology**, v. 39, n. 3, p. 135-140, 1964.

GODDARD, J. Infectious diseases and arthropods. **Springer Science & Business Media**, 2009.

GUGLIELMONE, A. A. et al. The *Argasidae*, *Ixodidae* and *Nuttalliellidae* (*Acari: Ixodida*) of the world: a list of valid species names, n. 6. p. 1-28, 2010.

GULIA-NUSS, M. et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. **Nature Communications**, v. 7, n. 1, p. 1-13, 2016.

GUTIÉRREZ, R. et al. The effect of ecological and temporal factors on the composition of *Bartonella* infection in rodents and their fleas. **The ISME Journal**, v. 8, n. 8, p. 1598-1608, 2014.

HARDSTONE YOSHIMIZU, M.; BILLETER, S. A. Suspected and confirmed vector-borne rickettsioses of North America associated with human diseases. **Tropical Medicine and Infectious Disease**, v. 3, n. 1, p. 2, 2018.

JONES, K. E. et al. Global trends in emerging infectious diseases. **Nature**, v. 451, n. 7181, p. 990-993, 2008.

JONGEJAN, F.; UILENBERG, G. The global importance of ticks. **Parasitology**, v. 129, n. S1, p. S3-S14, 2004.

KEIRANS, J. E.; CLIFFORD, C. M.; CORWIN, D. *Ixodes sigelos*, n. sp. (*Acarina: Ixodidae*), a parasite of rodents in Chile, with a method for preparing ticks for examination by scanning electron microscopy. **Acarologia**, v. 18, n. 2, p. 217-225, 1976.

KLOMPEN, H. Ticks, the *Ixodida* In: MARQUARDT, W. C. et al. Biology of Disease Vectors. 2nd ed Colorado: **Elsevier Academic Presss**, p. 45-55, 2005.

LABRUNA, M. B. et al. Genetic identification of rickettsial isolates from fatal cases of Brazilian spotted fever and comparison with *Rickettsia rickettsii* isolates from the American continents. **Journal of Clinical Microbiology**, v. 52, n. 10, p. 3788-3791, 2014.

LABRUNA, M. B. et al. Seasonal dynamics of ticks (*Acari: Ixodidae*) on horses in the state of São Paulo, Brazil. **Veterinary Parasitology**, v. 105, n. 1, p. 65-77, 2002.

- MAHAJAN, S. K. Rickettsial diseases. **Journal of the Association of Physicians of India**, v. 60, p. 37-44, 2012.
- MARTIN, P. A. W; SCHMIDTMANN, E. T. Isolation of aerobic microbes from *Ixodes scapularis* (Acari: Ixodidae), the vector of Lyme disease in the eastern United States. **Journal of Economic Entomology**, v. 91, n. 4, p. 864-868, 1998.
- MARTINS, T. F. et al. New tick records from the state of Rondônia, western Amazon, Brazil. **Experimental and Applied Acarology**, v. 62, n. 1, p. 121-128, 2014.
- MARUYAMA, S. R. et al. The expression of genes coding for distinct types of glycine-rich proteins varies according to the biology of three metastriate ticks, *Rhipicephalus* (Boophilus) *microplus*, *Rhipicephalus sanguineus* and *Amblyomma cajennense*. **BMC Genomics**, v. 11, n. 1, p. 363, 2010.
- MASTROPAOLO, M. et al. The ticks (Acari: Ixodida: Argasidae, Ixodidae) of Bolivia. **Ticks and Tick-Borne Diseases**, v. 5, n. 2, p. 186-194, 2014.
- MOREIRA, H. N. S. et al. A deep insight into the whole transcriptome of midguts, ovaries and salivary glands of the *Amblyomma sculptum* tick. **Parasitology International**, v. 66, n. 2, p. 64-73, 2017.
- MUÑOZ-LEAL, S. et al. Description of a new soft tick species (Acari: Argasidae: Ornithodoros) associated with stream-breeding frogs (Anura: Cycloramphidae: Cycloramphus) in Brazil. **Ticks and Tick-Borne Diseases**, v. 8, n. 5, p. 682-692, 2017.
- NARASIMHAN, S. et al. Gut microbiota of the tick vector *Ixodes scapularis* modulate colonization of the Lyme disease spirochete. **Cell Host & Microbe**, v. 15, n. 1, p. 58-71, 2014.
- NAVA, S. et al. Reassessment of the taxonomic status of *Amblyomma cajennense* () with the description of three new species, *Amblyomma tonelliae* n. sp., *Amblyomma interandinum* n. sp. and *Amblyomma patinoi* n. sp., and reinstatement of *Amblyomma mixtum*, and *Amblyomma sculptum* (Ixodida: Ixodidae). **Ticks and Tick-Borne Diseases**, v. 5, n. 3, p. 252-276, 2014.
- NELDER, M. P. et al. Human pathogens associated with the blacklegged tick *Ixodes scapularis*: a systematic review. **Parasites & Vectors**, v. 9, n. 1, p. 265, 2016.
- NOURANI, E. et al. Computational approaches for prediction of pathogen-host protein–protein interactions. **Frontiers in Microbiology**, v.6, p. 1-10, 2015.
- NUNES, P. H.; BECHARA, G. H.; CAMARGO-MATHIAS, M. I. Morphological changes in the salivary glands of *Amblyomma cajennense* females (Acari: Ixodidae) in different feeding stages on rabbits at first infestation. **Experimental and Applied Acarology**, v. 45, n. 3-4, p. 199, 2008.
- OTTMAN, C. Protein–protein interactions: an overview. In MANNHOLD R et al. (eds): **Protein–Protein Interactions in Drug Discovery**. Singapore, Wiley-VCH Verlag GmbH & Co, p. 1-19, 2013.

PAROLA, P. et al. Update on tick-borne rickettsioses around the world: a geographic approach. **Clinical Microbiology Reviews**, v. 26, n. 4, p. 657-702, 2013.

PAROLA, P.; LABRUNA, M. B.; RAOULT, D.; Tick-borne rickettsioses in America: unanswered questions and emerging diseases. **Current Infectious Disease Reports**, v. 11, n. 1, p. 40-50, 2009.

POLICASTRO, P. F. et al. *Rickettsia rickettsii* growth and temperature-inducible protein expression in embryonic tick cell lines. **Journal of Medical Microbiology**, v. 46, n. 10, p. 839-845, 1997.

RAMAPRASAD, A.; PAIN, A.; RAVASI, T. Defining the protein interaction network of human malaria parasite *Plasmodium falciparum*. **Genomics**, v. 99, n. 2, p. 69-75, 2012.

RAMOS, V. N. et al. Ticks on humans in the Pantanal wetlands, Brazil. **Ticks and Tick-Borne Diseases**, v. 5, n. 5, p. 497-499, 2014.

RAOULT, D. Emerging rickettsioses reach the United States. **Clinical Infectious Diseases**, v. 51, n. 1, p. 121-122, 2010.

RAOULT, D. et al. Naming of rickettsiae and rickettsial diseases. **Annals of the New York Academy of Sciences**, v. 1063, n. 1, p. 1-12, 2005.

RAOULT, D.; ROUX, V. Rickettsioses as paradigms of new or emerging infectious diseases. **Clinical Microbiology Reviews**, v. 10, n. 4, p. 694-719, 1997.

RICHARDS, A. L. Worldwide detection and identification of new and old rickettsiae and rickettsial diseases. **FEMS Immunology & Medical Microbiology**, v. 64, n. 1, p. 107-110, 2012.

SNOWDEN, J.; KING, K. C. Rickettsial Infection. **Europe PMC**, 2019.

SOARES, J. F. et al. Experimental infection of the tick *Amblyomma cajennense*, Cayenne tick, with *Rickettsia rickettsii*, the agent of Rocky Mountain spotted fever. **Medical and Veterinary Entomology**, v. 26, n. 2, p. 139-151, 2012.

SOCOLOVSCHI, C. et al. Transovarial and trans-stadial transmission of *Rickettsiae africae* in *Amblyomma variegatum* ticks. **Clinical Microbiology and Infection**, v. 15, p. 317-318, 2009.

SONENSHINE, D. E.; ROE, R. M.(Ed.). **Biology of Ticks**. Oxford University Press, 2013.

SPONCHIADO, J. et al. Association patterns of ticks (*Acari: Ixodida: Ixodidae, Argasidae*) of small mammals in Cerrado fragments, western Brazil. **Experimental and Applied Acarology**, v. 65, n. 3, p. 389-401, 2015.

SPRONG, H. et al. *Ixodes ricinus* ticks are reservoir hosts for *Rickettsia helvetica* and potentially carry flea-borne *Rickettsia* species. **Parasites & Vectors**, v. 2, n. 1, p. 41, 2009.

UCHIYAMA, T. Tropism and pathogenicity of rickettsiae. **Frontiers in Microbiology**, v. 3, p. 230, 2012.

WALKER, D. H. Rocky Mountain spotted fever: a disease in need of microbiological concern. **Clinical Microbiology Reviews**, v. 2, n. 3, p. 227-240, 1989.

WEINERT, L. A. et al. Evolution and diversity of *Rickettsia* bacteria. **BMC Biology**, v. 7, n. 1, p. 6, 2009.

WILSON, L. B.; CHOWNING, WM. M. The so-called spotted fever of the rocky mountains.: a preliminary report to the montana state board of health. **Journal of the American Medical Association**, v. 39, n. 3, p. 131-136, 1902.

YU, X. J.; WALKER, D. H. The Order *Rickettsiales*. In: The Prokaryotes: An Evolving Electronic Resource for the Microbiology Community. 2003.

ZIVKOVIC, Z. et al. Differential expression of genes in salivary glands of male *Rhipicephalus* (Boophilus) *microplus* in response to infection with *Anaplasma marginale*. **BMC Genomics**, v. 11, n. 1, p. 186, 2010.

CAPÍTULO 2

DELINEAMENTO EXPERIMENTAL

A reconstrução da rede de interação proteína-proteína foi realizada a partir de proteínas codificadas nos genomas do carrapato *I. scapularis* e da bactéria *R. amblyommatis*. As proteínas destes organismos foram alinhadas conforme o preconizado por cada banco de interações: PSIMAP, iPfam, STRING e PEIMAP. Em seguida, os resultados foram concatenados de modo a estabelecer as interações entre *I. scapularis* e *R. amblyommatis*.

Uma rede consenso foi construída após a normalização dos dados através do cálculo do valor de confiança para cada interação. De modo a selecionar as proteínas de *R. amblyommatis* com alto potencial de interação com o carrapato *I. scapularis* e de grande importância para o processo infeccioso foram recuperadas na rede consenso as proteínas secretadas, excretadas e, ou, de superfície através de preditores de localização subcelular, a saber: SignalP 4.1, TMHMM 2.0, PsortB 3.0.4, LipoB 1.0a e Phobius 1.0.1.

Uma vez predita a rede de interação proteína-proteína para os organismos em estudo, realizou-se a anotação funcional das proteínas do carrapato baseada na Ontologia Gênica categorizadas em: processo biológico, componente celular e função molecular, de modo a identificar proteínas envolvidas na modulação da fisiologia do *I. scapularis*.

Por se tratar de maquinarias celulares dinâmicas, a análise do interactoma baseado nos genomas dos organismos integrada aos transcriptomas de ovário e intestino de fêmeas de *A. sculptum* antes e após a infecção por *R. amblyommii* permitiu refinar e selecionar as interações que melhor caracterizam o processo de infecção.

Para a análise do RNA-seq de ovário e intestino seguiram as seguintes etapas: controle de qualidade (FastQC/Trimmomatic), montagem de novo (Trinity) e alinhamento das *reads* e expressão diferencial dos genes nos diferentes órgãos (Kallisto).

A análise da expressão diferencial em ovário e intestino permitiu selecionar proteínas mais representativas que foram encaminhadas a modelagem molecular para a predição da estrutura tridimensional (Phyre2), análise da qualidade da estrutura tridimensional (PROCHECK, ERRAT, VERIFY 3D, ProSA e Ramachandran

plot), minimização de energia dos modelos (*Yasaha Energy Minimization*), seguida novamente da análise da qualidade da estrutura tridimensional e docking (ClusPro) entre as proteínas de bactéria e do carrapato (Figura 1).

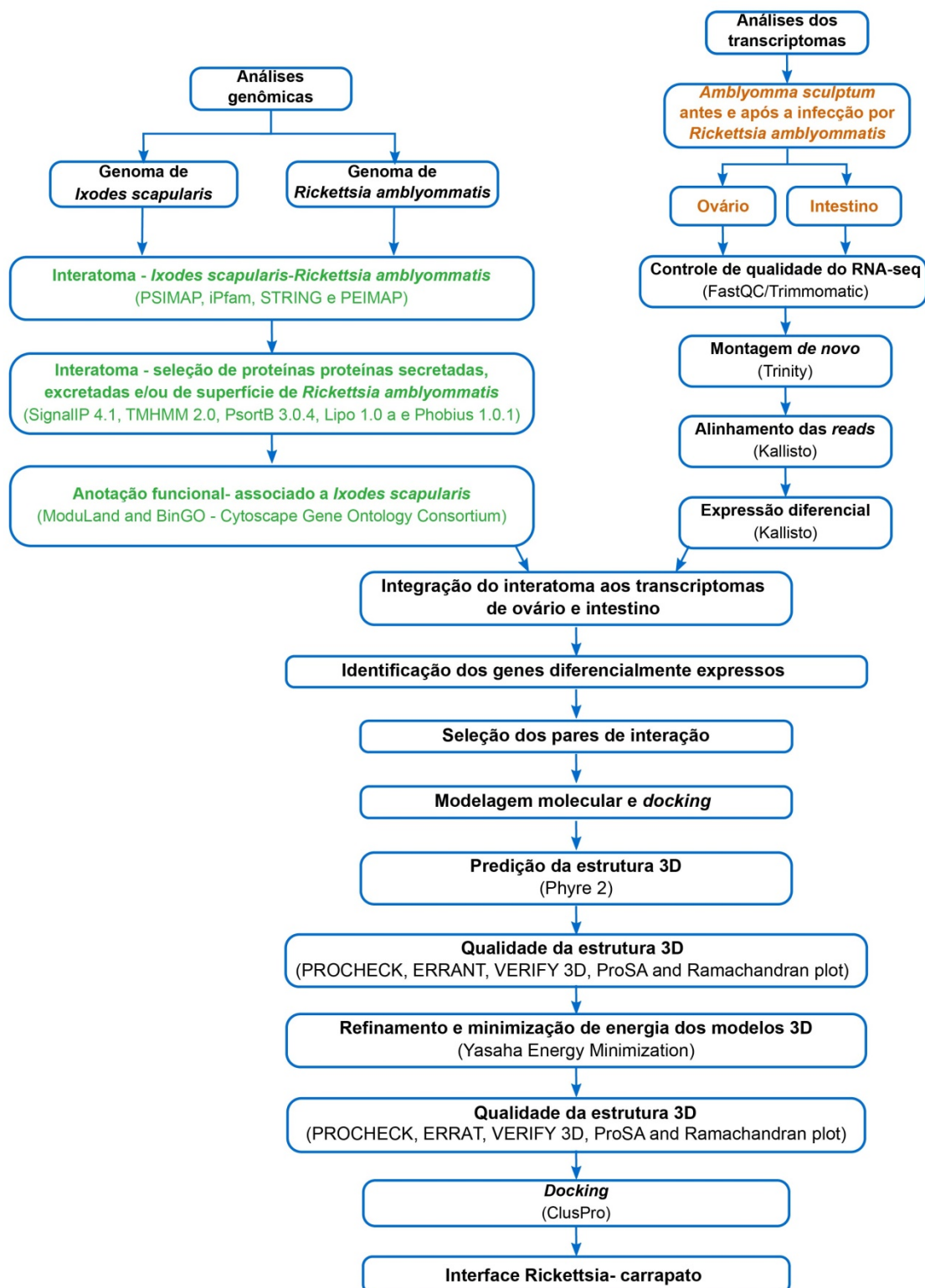


Figura 1. A predição da rede de interação proteína-proteína na relação entre *Rickettsia amblyommatis* e células do intestino e ovário de *Amblyomma sculptum* (Berlese, 1888) foi realizado

em colaboração a trabalhos desenvolvidos pelo grupo de pesquisa, desta forma o diagrama em cores representa em verde o estudo da Interação funcional, expressão de genes e filogenia entre proteínas de membrana de *Rickettsia* spp. e carrapatos; em laranja o estudo da anotação e montagem de transcriptomas de intestino médio e ovários do carrapato *Amblyomma sculptum*, antes e após a infecção por *Rickettsia amblyommatis*; e, em preto, o estudo apresentado na presente tese.

Dessa forma, o estudo enfatizou a importância da rede de interação proteína-proteína na compreensão da interface patógeno-vetor e no desenvolvimento de novas pesquisas e novas aplicações biotecnológicas, bem como o auxílio na seleção de proteínas para o desenvolvimento de drogas, inclusive para inibir interações específicas.

CAPÍTULO 3

Rede de interação proteína-proteína na relação entre *Rickettsia amblyommatis* e células do intestino e ovário de *Amblyomma sculptum* (BERLESE, 1888)

DIAS, R. X. L.¹; GOMES, G. G.¹; MENDES, T. A. O.¹; MAFRA, C.^{1*}

¹ Departamento de Bioquímica e Biologia Molecular, Universidade Federal de Viçosa, Viçosa, Minas Gerais, 36570-900, Brasil.

*mafra@ufv.br

RESUMO

O estudo propôs a reconstrução da rede de interação proteína-proteína a partir das proteínas codificadas nos genomas do carrapato *Ixodes scapularis* e da bactéria *Rickettsia amblyommatis* usando quatro métodos: PSIMAP, PEIMAP, String e iPfam. Combinando os resultados destes métodos, calculou-se uma rede de alta confiança (índice de confiança >0,70) com 3.546 interações envolvendo 91 proteínas de membrana da bactéria e 980 proteínas do carrapato. A partir da Ontologia Gênica (GO) foram preditos processos biológicos significativos (valor p ajustado <0,05) envolvidos na modulação da fisiologia do carrapato pela bactéria. Os módulos encontrados revelam diferentes vias de sinalização, a exemplo, secreção, alongamento translacional e fosforilação de aminoácidos e proteínas. De modo a validar a rede, integraram-se os transcriptomas de intestino e ovário de carrapatos *Amblyomma sculptum* em resposta à infecção por *Rickettsia amblyommatis*. A análise da expressão diferencial revelou que a presença da *Rickettsia amblyommatis* negativou a expressão de genes no intestino e ovário. A análise integrativa revelou 12 genes do carrapato expressos diferencialmente durante a infecção e 42 interações proteicas riquetsia-carrapato. Com base em estudos anteriores, que já haviam demonstrado a importância das proteínas da riquetsia para a infecção, selecionou-se 4 proteínas, a saber: Era de ligação ao GTP, LepA de ligação ao GTP, proteína do tipo PPIC com domínio PPIASE e ecotina e as respectivas proteínas do carrapato correspondentes: calmodulina putativa, gap1 putativa, proteína hipotética conservada, proteína hidrolase hipotética e proteína putativa de ativação da fosfolipase A2. A predição das estruturas tridimensionais descritas foi realizada pelo aplicativo Phyre2, a análise de qualidade pelos servidores PROCHECK, ERRAT, VERIFY 3D, ProSA e Ramachandran plot, e por fim a minimização de energia dos modelos pelo aplicativo *Yasaha Energy Minimization*. A análise revelou modelos de boa qualidade. A partir da análise de *docking* molecular foram obtidos cinco potenciais complexos proteicos: calmodulina putativa-proteína do tipo PPIC com domínio PPIASE, proteína hipotética conservada-proteína do tipo PPIC com domínio PPIASE, proteína hipotética conservada e proteína LepA de ligação ao GTP, proteína hipotética hidrolase-ecotina, proteína putativa de ativação da fosfolipase A2 e

Proteína Era de ligação ao GTP. Assim, o estudo revelou genes, interações proteicas e modelos 3D de boa qualidade derivadas de evidências *in silico* que podem auxiliar estudos adicionais sobre os mecanismos que envolvem a manutenção e transmissão de riquetsias por carrapatos.

ABSTRACT

This study proposed the reconstruction of the protein-protein interaction network from the proteins encoded in the genomes of the tick *Ixodes scapularis* and the bacterium *Rickettsia amblyommatis* using four validated methods: PSIMAP, PEIMAP, String and iPfam. Combining the results of these methods, we calculated a highly reliable network (confidence index > 0.70) with 3,546 interactions involving 91 bacterial membrane proteins and 980 tick proteins. From the Genetic Ontology (GO), the significant biological processes (adjusted p-value <0.05) involved in the modulation of the tick's physiology by the bacteria were predicted. The found modules reveal different signaling pathways such as secretion, translational elongation, phosphorylation of amino acids and proteins. In order to validate the network, we integrated the intestine and the ovary transcriptomes of ticks of *Amblyomma sculptum* in response to infection by *Rickettsia amblyommatis*. The comparative analysis of the levels of the differential expression revealed that the presence of *Rickettsia amblyommatis* decreased the expression of genes in the intestine and the ovary. Integrative analysis revealed 12 tick genes expressed during infection and 42 *Rickettsia*-tick protein interactions. Based on previous studies that had already demonstrated the importance of *Rickettsia* proteins for infection, we selected 4 proteins, namely: GTP-binding Era, GTP-binding LepA, PPIC-type protein with PPIASE domain and ecotine and the respective corresponding tick proteins: putative calmodulin, putative gap1, conserved hypothetical protein, hypothetical hydrolase protein and putative phospholipase A2 activation protein. The prediction of the three-dimensional structures described was performed by the Phyre2 application, the quality analysis by the PROCHECK, ERRAT, VERIFY 3D, ProSA and Ramachandran plot servers and finally the energy minimization of the models by the Yasaha Energy Minimization application. The analysis revealed good quality models. From the molecular docking analysis, five potential complexes were obtained: putative calmodulin-PPIC-type protein with PPIASE domain, conserved hypothetical protein-PPIC-type protein with PPIASE domain, conserved hypothetical protein and GTP-binding LepA protein, hypothetical hydrolase protein -ecotin, putative phospholipase A2 activation protein and GTP-binding Era protein. Thus, the study revealed genes, protein interactions and good quality 3D models derived from *in silico* evidence that can help additional studies on the mechanisms involving the maintenance and transmission of *Rickettsia* by ticks.

3.1. INTRODUÇÃO

As riquetsioses são doenças infecciosas causadas por bactérias do gênero *Rickettsia*, com algumas de suas espécies tendo sido apontadas como causadoras de enfermidades muitas vezes desconhecidas ou emergentes (PAROLA et al., 2013). Reportadas em diversas regiões do mundo, estão associadas a artrópodes reservatórios vetores, mantidos na natureza por hospedeiros que atuam como amplificadores de riquetsia (JONGEJAN & UILENBERG, 2004; PAROLA et al., 2013). A percepção da importância das riquetsioses tem aumentado e diversos estudos estão focados em entender a atuação desses agentes patogênicos (RAOULT, 2010; EREMEEVA & DASCH, 2015).

Considerados os vetores artrópodes mais comuns na transmissão de riquetsioses (PAROLA et al., 2013), os carrapatos atuam dispersando estes agentes diretamente através de uma picada para a realização de hematofagia ou através da inoculação de material fecal infectado na pele ou mucosa em humanos e animais (RICHARDS, 2012). Dentre as doenças causadas por riquetsias e transmitidas por carrapatos, as febres maculosas são as de maior impacto e incidência mundial. No Brasil, a Febre Maculosa Brasileira (FMB) é a mais patogênica dentre outras riquetsioses, causando problemas severos à saúde pública. Essa doença é causada pela *Rickettsia rickettsii*, tendo como principal vetor o carrapato *Amblyomma sculptum* (Complexo *Amblyomma cajennense*), conhecido popularmente como carrapato estrela (NUNES et al., 2008; PAROLA et al., 2009).

Para a escolha de novos alvos moleculares no desenvolvimento de medidas de controle para carrapatos e das doenças associadas a esse vetor, faz-se necessário que pesquisas abrangendo estes agentes e enfermidades estejam voltadas para entender as interações moleculares entre hospedeiros, carrapatos e agentes patogênicos (DE LA FUENTE et al., 2016). Desta maneira, avaliar a ação do carrapato sobre uma variedade de hospedeiros vertebrados e, por consequência, como a transmissão de zoonoses está intimamente relacionada à sua morfologia, bem como a fisiologia dos órgãos internos, sobretudo glândulas salivares, intestino e ovário. Assim, um melhor entendimento dos processos moleculares envolvidos na interação patógeno-vetor nestes órgãos, como etapa fundamental para compreensão da biologia destes ácaros, bem como dos eventos moleculares que

regem seu papel como agente vetor de microorganismos do gênero *Rickettsia* (GULIA-NUSS et al., 2016; MOREIRA et al., 2017).

Com a expansão das pesquisas genômicas, incrementada pelo desenvolvimento de novas tecnologias - dentre estas, o sequenciamento de nova geração e o rápido desenvolvimento de bancos de dados genômicos para vários agentes patogênicos e seus vetores, possibilitando uma abordagem baseada na biologia de sistemas para o estudo das interações patógeno-vetor têm-se avançado na compreensão das interações *Rickettsia*-carrapato. Essas abordagens têm resultado no entendimento de muitos processos biológicos importantes que governam a doença parasitária, proporcionando informações com o potencial de impulsionar descobertas de novas drogas, além de acelerar o desenvolvimento de novas terapias (FLÓREZ et al., 2010).

Essenciais para o estudo da função biológica, as interações proteína-proteína são consideradas como uma classe emergente de alvos terapêuticos (DÖMLING, 2008), tendo sido utilizadas para a compreensão de doenças infecciosas como tuberculose, malária e filariose associadas, respectivamente, aos agentes *Mycobacterium tuberculosis* (RAMAN & CHANDRA, 2008), *Plasmodium falciparum* (La COUNT et al., 2003) e *Brugia malayi* (TAYLOR et al., 2011), a seleção de alvos moleculares (KAR et al., 2009, HASE et al., 2009, HWANG et al., 2006) e ao desenvolvimento de fármacos (MULLARD, 2012). Estudos demonstraram a coevolução de interações moleculares entre carrapatos e os patógenos os quais transmitem, envolvendo a conservação de características genéticas que medeiam o desenvolvimento e sobrevivência do patógeno em seu vetor (de la FUENTE et al., 2010).

A reconstrução de redes de interação proteína-proteína usando a abordagem computacional, auxilia na compreensão dos mecanismos moleculares e funcionais que caracterizam a relação patógeno-vetor, incluindo a replicação eficiente do patógeno e sua manutenção em longo prazo no ambiente celular, bem como a sua transmissão. Por meio de uma abordagem integrativa através das redes de interação proteína-proteína, torna-se possível desvendar relações moleculares complexas nos sistemas vivos, mapeando as interações físicas de proteínas para proteínas (DURMUS & ULGEN, 2013; OTTMAN, 2013). Neste sentido, este estudo almejou traçar as interações dinâmicas envolvidas na relação *R. amblyommatis* - *A.*

sculptum, com vista à compreensão dos mecanismos moleculares e funcionais que caracterizam tal interação.

3.2. MÉTODOS

3.2.1. Interatoma

Para a reconstrução do interatoma, foram utilizados os genomas depositados de *I. scapularis* (Taxonomy ID 6945) com 21.259 proteínas e de *R. amblyommatidis str Ac/Pa ASM96467v1* (Taxonomy ID) 1359164 com 1.811 proteínas, ambas depositadas respectivamente, no Ensembl Metazoa e Ensembl Bacteria versão 38 (KERSEY et al., 2010).

A reconstrução da rede de interação proteína-proteína para estes organismos foi realizada a partir de quatro bancos de interações experimentalmente validadas entre proteínas de diferentes organismos, quais sejam: PEIMAP (KIM, 2008; VON-MERING et al., 2003), STRING, (PARK, 2005), PSIMAP (FINN, 2006), iPfam (FINN, 2006; BARABÁSI et al., 2004). Para o alinhamento contra a base de dados de proteínas depositadas no PEIMAP e no STRING, utilizou-se o aplicativo BLASTP (ALTSCHUL, 1997), considerando interações de proteínas significativas com alinhamentos com um mínimo de 40% de identidade e 70% de cobertura (KIM, 2008). Para a base de dados SCOP (HUBBARD et al., 1997) utilizada pelo PSIMAP, o PSI-BLAST foi o método de alinhamento (ALTSCHUL, 1997), considerando *e-value* de 1×10^{-4} (KIM, 2008; VON-MERING et al., 2003) para predição de interações proteicas. A fim de identificar a presença de domínios no iPfam, utilizou-se a ferramenta HMM Pfam baseada em Modelos Ocultos de Markov para análise de sequências biológicas (KROGH et al., 1994). Por fim, utilizamos o banco de dados STRING (Ferramenta de Pesquisa para Recuperação de Genes / Proteínas Interagentes) o qual é dedicado a associações funcionais entre proteínas, em escala global. Todas as fontes de evidência de interação são comparadas e calibradas em relação ao conhecimento prévio, utilizando os agrupamentos funcionais de alto nível fornecidos pelo Kyoto Enciclopédia de Genes e Genomas (KEEG) (YANAI et al., 2001; SNEL et al., 2000; SZKLARCZYK, 2011).

Para integrar os dados de predição foi realizado um cálculo de confiança (VON-MERING et al., 2003; FINN, 2006; FLÓREZ et al., 2010) que assume independência entre os quatro métodos, o qual foi calculado de acordo com a fórmula:

$$score = 1 - \prod_{i \in M} (1 - R_i)^n$$

onde, *score* é o valor de confiança, *M* é o conjunto de métodos utilizados (PEIMAP, STRING, PSIMAP e iPfam), *R_i* é o valor de confiança de cada método *i*, e *n* é o número de vezes que uma interação foi predita pelo método *i*. O valor de confiança do PEIMAP reflete a confiabilidade de cada método experimental para detectar interações entre proteínas, tendo sido previamente estimado (VON-MERING et al., 2003; FLÓREZ et al., 2010).

O valor de confiança do iPfam é a soma entre os dois *bit score* da assinatura dos domínios do Pfam do par proteínas com potencial interação. O valor de confiança do PSIMAP é o inverso da distância entre os pares de domínios estruturais que estão interagindo no PDB. O valor de confiança do STRING é calculado pela própria base de dados. O valor de confiança final foi combinando e normalizado gerando valores entre 0 e 1. Interações preditas foram consideradas de alta confiabilidade para valores de confiança combinada acima de 0,7 (VON-MERING et al., 2003).

Para a identificação de potenciais proteínas secretadas/excretadas ou de superfície de *R. amblyommatidis* foram utilizados cinco aplicativos preditores de localização subcelular, quais sejam: SignalP 4.1, TMHMM 2.0c, LipoB 1.0a, Psortb 3.0.4 e Phobius 1.01. O SignalP 4.1 (PETERSEN, 2011) utiliza-se de uma combinação de diferentes redes neurais para prever a presença do peptídeo sinal na região amino terminal e a localização do sítio de clivagem que media a translocação de uma proteína do citoplasma para a membrana. O algoritmo TMHMM 2.0 (SONNHAMMER et al., 1998) identifica potenciais hélices transmembrana em sequências proteicas utilizando um modelo oculto de Markov (KROGH et al., 1994). O PSORTb versão 3.0.4 (GARDY et al., 2005) gera resultados de predição para cinco localizações em bactérias Gram-negativas: citoplasmática, membrana interna, periplasmática, membrana externa e extracelular. O LipoB versão 1.0a (YU et al., 2004) produz previsões de lipoproteínas, discriminando entre peptídeos de sinal de lipoproteína, outros peptídeos de sinal e hélices de membrana n-terminal em bactérias Gram-negativas. O Phobius 1.01 foi também usado para identificar peptídeos sinais e domínio transmembrana utilizando um algoritmo baseado em

modelo oculto de Markov, o qual, quando combinado ao SignalP e TMHMM, reduz substancialmente erros de predição (KALL et al., 2005).

A anotação de ontologia gênica (GO) (CAMON, 2004) para função molecular e processo biológico de proteínas da interface *R. amblyommatis* vs *I. scapularis* foi obtida por meio do servidor AgBase-GOanna 2.2 (McCARTHY, 2006; MENDES, 2013). Inicialmente o proteoma de *I. scapularis* foi submetido ao BLASTP contra as anotações GO do banco de dados SwissProt com e-value menor que 10^{-5} , tamanho da palavra 3, filtro de baixa complexidade ligado e matriz de substituição BLOSUM62 (MENDES, 2013). Nesta etapa, foram consideradas as proteínas pertencentes a todos os grupos funcionais para os quais elas fossem assinaladas e classificadas com múltiplas assinaturas de GO. Posteriormente, os dados para funções moleculares foram curados manualmente para eliminar redundâncias de proteínas (MENDES, 2013).

Uma rede consenso contendo todas as interações preditas para a interface *R. amblyommatis* vs *I. scapularis* foi construída e visualizada no aplicativo *Cytoscape* versão 3.5.1 (SAITO, 2012). Para identificação de módulos biologicamente funcionais, inicialmente a rede foi dividida em módulos através do pacote APCluster, o qual utiliza o método de afinidade de propagação para identificar módulos coerentes dentro de uma rede (FREY & DUECK, 2007). A seguir, o termo de ontologia gênica mais significativo para cada módulo foi identificado utilizando o método de distribuição hipergeométrica implementado no pacote BINGO (MAERE et al., 2005) considerando termos enriquecidos para $p < 0,05$.

3.2.2. Análise do RNA-seq

RNA-seq de intestino e ovário de fêmeas de *A. sculptum* com e sem infecção por *R. amblyommatis* foram obtidos em colaboração ao estudo: Anotação e montagem de transcriptomas de intestino médio e ovários do carrapato *A. sculptum*, antes e após a infecção por *R. amblyommatis* (MAFRA, C. Dados não publicados). Para controle de qualidade das *reads* utilizou-se o aplicativo Trimmomatic com o intuito de remover as sequências do adaptador Illumina (BOLGER et al., 2014). A qualidade dos dados produzidos foi analisada usando o FastQC pelo índice de qualidade Phred (VAN DER AUWERA et al., 2013), sendo descartadas as leituras com índices menores que 15. A montagem *de novo* e pseudo alinhamento das *reads*

foram realizadas por meio do Trinity 2.2.0 e Kallisto, respectivamente (GRABHERR et al., 2011; BRAY et al., 2016). A abundância de transcrições foi calculada como transcritos por milhão de leituras mapeadas (TPM), o que reflete a abundância nas amostras do intestino e ovário por normalização do comprimento do RNA e do número total de leitura (BRAY et al., 2016).

A partir da montagem 'de novo' dos transcriptomas por meio do Trinity 2.2.0 foram identificados os genes diferencialmente expressos de *A. sculptum* sob condições de infecção e não infecção por *R. amblyommatis*. Em seguida, foi realizada a identificação das proteínas codificadas por genes com expressão diferenciada no modelo de interatoma construído entre as proteínas da interface de *R. amblyommatis* e *I. scapularis*, com o objetivo de delinear os mecanismos moleculares que ocorrem no ovário e intestino de *A. sculptum* em resposta à infecção por *R. amblyommatis*.

Ao considerar estudos anteriores, selecionaram-se 4 proteínas de riquétsia que formaram 7 pares de interação bactéria-carrapato, cuja relevância no processo de infecção já havia sido demonstrada (THEPPARIT et al., 2010, UCHIYAMA, 2012, STAPELS et al., 2015, PETCHAMPAI et al., 2015).

3.2.3. Modelagem Molecular e *Docking*

O servidor Phyre2 (KELLEY & STERNBERG, 2009) foi utilizado para modelagem da estrutura tridimensional das proteínas das ISCW001070-PA.1, ISCW002948-PA.1, ISCW003706-PA.1, ISCW004280-PA.1, ISCW008317-PA.1, KJV61538.1, KJV61792.1, KJV62290.1 e KJV62.746-1.

O refinamento estrutural e a minimização de energia dos modelos previstos foram realizados por meio servidor de Minimização de Energia *Yet Another Scientific Reality Artificial Application* (YASARA) (KRIEGER et al., 2009). A qualidade dos modelos foi mensurada antes e após refinamento e minimização de energia por meio da avaliação da geometria e estereoquímica, distribuição de energia e outras características dos modelos 3D usando PROCHECK (<https://servicesn.mbi.ucla.edu/PROCHECK/>) (WIEDERSTEIN e SIPPL, 2007), ProSA-web (<https://prosa.services.came.sbg.ac.at/prosa.php>) (WALLNER e ELOFSSON, 2003), Verify-3D (http://services.mbi.ucla.edu/Verify_3D/) (MORRIS et al., 2009) e ERRAT (<http://services.mbi.ucla.edu/ERRAT/>) (HESS et al., 2008).

O *docking* dos pares de interação *Rickettsia*-carrapato foi realizado por meio do aplicativo ClusPro (KOZAKOV et al., 2013; KOZAKOV et al., 2017). O modo de ligação entre as proteínas foi visualizado e analisado usando o programa PyMOL *Molecular Graphics System*, versão 2.0 (<http://www.pymol.org/>).

Para as simulações por dinâmica molecular foram selecionadas as conformações de menor energia de cada *docking* realizado. Nessa etapa foi empregado o pacote de simulação de dinâmica molecular NAMD 2.12, utilizando o campo de força CHARMM27, que pode ser empregado para dados de proteínas, modelo TIP3P de água e lipídios. A inserção e orientação dos complexos proteicos na bicamada lipídica 1-Palmitoil-2-oleoilfosfatidilcolina (POPC) foi guiada pelos resíduos da região intramembranar da proteína de carrapato. O servidor web PPM Server para recursos de posicionamento de proteínas em membranas foi utilizado na validação dos dados (LOMIZE et al., 2012).

O tamanho da bicamada foi determinado pela proteína receptora, com um espaço de 30 Å entre a proteína e o limite membranar na direção do eixo X e Y. O tamanho da membrana foi de 130 Å no plano XY, contendo 205 moléculas de lipídio por monocamada. A bicamada foi criada utilizando o programa VMD (*Visual Molecular Dynamics*) (Humphrey et al. 1996) com a ferramenta Membrane Builder. A solvatação do sistema foi feita usando o modelo de água TIP3P, seguindo a consideração que a distância entre o extremo da proteína e o limite da caixa de simulação fosse >15 Å em cada direção. Foram adicionadas um total de 31863 moléculas de água e 187 moléculas de cloreto de sódio utilizadas como íons para manter a neutralidade da carga total do sistema. A inserção dos íons foi realizada empregando a ferramenta Autoionize do VMD.

A simulação da proteína na bicamada lipídica foi dividida em cinco etapas conforme descrito abaixo:

- Após 1.000 passos de minimização, o sistema foi simulado a 300 K e pressão de 1 atm por 0.5 ns no ensemble de pressão constante (NPT) mantendo as moléculas de água, íons, proteínas e cabeças lipídicas fixas, exceto pelas caudas da bicamada lipídica. A temperatura foi controlada usando dinâmica de Langevin (*temperature damping coefficient*: 1.0) e a pressão usando o método de piston Langevin, com um período de oscilação de 200 fs. Para as interações não ligadas, o *cutoff* foi ajustado

para 12 Å e para as de longo alcance o *cutoff* foi modelado usando o método de Partícula de Mesh Ewald (PME). A simulação foi executada usando 250.000 passos;

- Seguido de 1000 passos de minimização, o sistema foi simulado a 300 K e pressão de 1 atm por 0.5 ns no ensemble de pressão constante (NPT) com a proteína restrita para permitir que os lipídios, águas e íons se adaptem à proteína em sua forma cristalina. A simulação foi executada usando 250000 passos;
- Foi realizada a terceira minimização de 5.000 passos mantendo a proteína relaxada para melhor equilíbrio e minimização do sistema. A simulação foi executada usando 5.000.000 passos em um tempo de 10 ns;
- Após os 12.000 passos de minimização, o sistema foi livre para ser simulado durante 10 ns com tempo de integração de 2 fs no NPT;
- Através de mais 5.000 passos de minimização, o sistema foi simulado durante 100 ns.

A energia livre da interação foi calculada utilizando a equação:

$$\Delta G_{\text{Bind}} = \beta(\langle V_{l-s}^{\text{el}} \rangle_{\text{bound}} - \langle V_{l-s}^{\text{el}} \rangle_{\text{free}}) + \alpha(\langle V_{l-s}^{\text{vdw}} \rangle_{\text{bound}} - \langle V_{l-s}^{\text{vdw}} \rangle_{\text{free}}) + \gamma$$

onde, $\langle V_{l-s}^{\text{el}} \rangle_Y$ e $\langle V_{l-s}^{\text{vdw}} \rangle_Y$ são a media dos valores de interações eletrostáticas e de van der Waals entre o ligante (*l*) e sua vizinhança (*s*) no estado *Y*, o qual representa a proteína livre em solução ou ligada a outra proteína. Os parâmetros α e β são respectivamente, o fator de escala de energia ajustado para dispersão e eletrostático e eles são constantes.

3.3. RESULTADOS

3.3.1. Interatoma

A rede de interação proteína-proteína foi construída combinando os resultados gerados pelos bancos: Peimpap, String, Psimap e iPfam. A análise de similaridade baseada em multialinhamentos recuperou um número de proteínas a partir do genoma de *R. amblyommatis*: 238 (Peimpap), 217.668 (String), 1.397 (Psimap) e

1.690 (iPfam). Para *I. scapularis*, usando-se os mesmos aplicativos, foram recuperadas 314.606 (Peimap), 273.682 (String), 314.606 (Psimap) e 2.288 (iPfam) proteínas. A partir rede categorizaram-se interações intraespecíficas e interespecíficas, sendo as interespecíficas de maior interesse na compreensão dos aspectos moleculares envolvidos na relação patógeno-vetor. A quantificação das interações interespecíficas (*Rickettsia*-carrapato) revelou 8.630 (Peimap), 2.194 (String), 145.951 (Psimap) e 495.013 (iPfam). A partir da quantificação das interações foi realizado cálculo do score final por banco e, em seguida o score combinado envolvendo todos os bancos. Para cada par de interação (*R. amblyommatis* - *I. scapularis*) foi atribuído um valor, este normalizado entre 0 a 1. Diante do score final combinado para cada interação uma rede consenso final foi reconstruída, considerando o valor de *cutoff* > 0,7. A rede de interação proteína-proteína de *R. amblyommatis* e *I. scapularis* apresentou um total 13.151 interações sendo representadas por 387 proteínas de *R. amblyommatis* e 1.252 proteínas do *I. scapularis* incorporando 21,2% (387) e 5,8% (1.252) das proteínas presentes nos genomas de *R. amblyommatis* e *I. scapularis*, respectivamente (Figura 1).

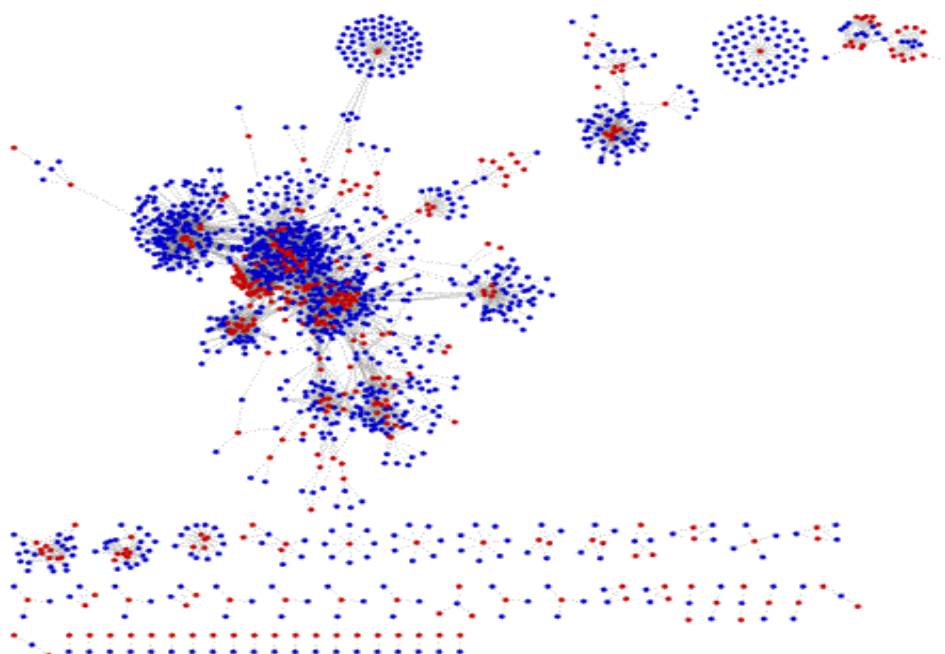


Figura 1. Rede de interação proteína-proteína *R. amblyommatis* e *I. scapularis* visualizada em nós (proteínas) e arestas (interações) pelo Cytoscape. Interatoma total *R. amblyommatis* (vermelho) e *I. scapularis* (azul) contendo 13.151 interações.

Por considerar a natureza intracelular da *Rickettsia*, juntamente com a capacidade de migrar por diferentes órgãos internos do carrapato, e se estabelecer de acordo com o estágio da infecção foram identificadas proteínas secretadas, excretadas e, ou, de superfície de *R. amblyommatis* através da reconstrução do interatoma de proteínas de membrana.

Os resultados dos preditores de localização subcelular: SignalP 4.1, TMHMM 2.0, PsortB 3.0.4, LipoB 1.0a e Phobius 1.0.1 recuperaram respectivamente, 44, 362, 462, 334 e 363 proteínas perfazendo um total de 632 proteínas. Estes dados foram comparados ao interatoma total, de modo que foram selecionadas 3.546 interações. Observou-se 91 proteínas da bactéria interagindo com 980 proteínas do carrapato. O interatoma de membrana abrangeu 5,0% e 4,6% das proteínas presentes nos genomas de *R. amblyommatis* e *I. scapularis*, respectivamente (Figura 2).

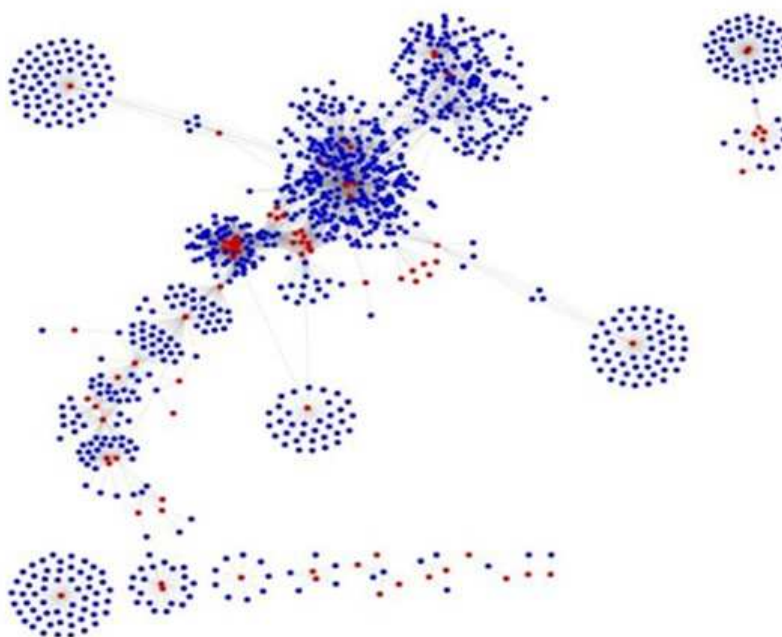


Figura 2. Rede de interação proteína-proteína *R. amblyommatis* e *I. scapularis* visualizada em nós (proteínas) e arestas (interações) pelo Cytoscape. Interatoma de membrana: *R. amblyommatis* (vermelho) e *I. scapularis* (azul) contendo 3.546 interações.

Para a anotação funcional utilizou-se o vocabulário de classificação funcional da Ontologia Gênica (GO), caracterizando 10 *clusters* significativos (valor *p* ajustado <0,05) envolvidos na modulação da fisiologia do carrapato, sendo estes com seus

respectivos p -valor corrigidos, a saber: proteólise ($3.8268E-7$), secreção ($6.1512E-6$), alongamento translacional ($1.7130E-4$), fosforilação ($2.3625E-15$), fosforilação de aminoácidos e proteínas ($3.5471E-15$), transporte mediado por vesículas ($7.8875E-10$), resposta contra vírus (2 módulos: $2.3960E-4$ e $3.4355E-5$), oxidação redutiva ($1.6924E-4$), oxidação fosforilativa ($2.5217E-4$) e regulação do DNA por metilação ($1.3812E-3$) (Figura 3).

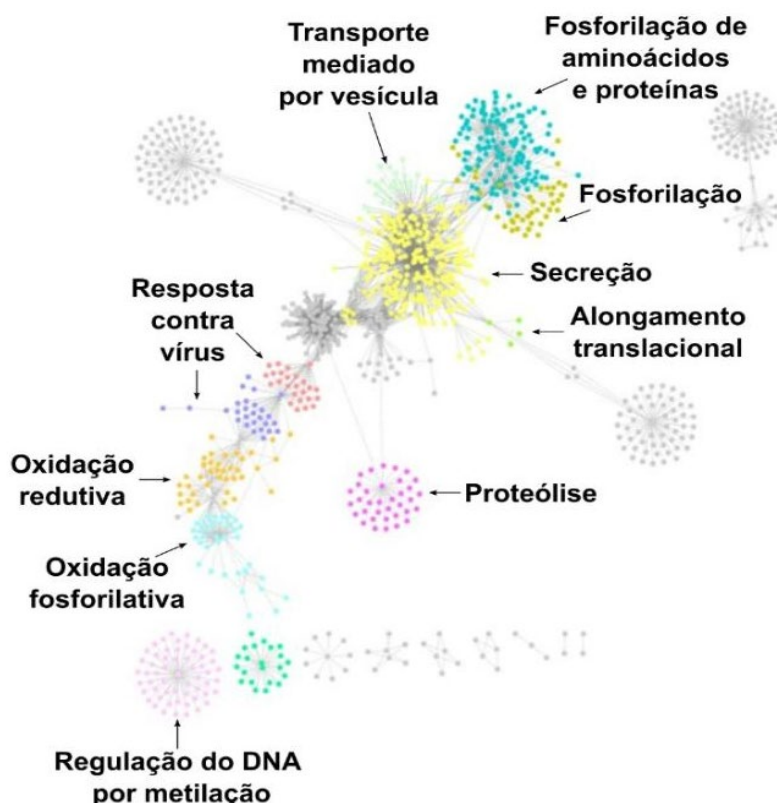


Figura 3. Rede de interação proteína-proteína *R. amblyommatis* e *I. scapularis* visualizada em nós (proteínas) e arestas (interações) pelo Cytoscape. Interatoma de membrana com a anotação funcional (GO) baseada na fisiologia do *I. scapularis*.

3.3.2. Análise do RNA-seq

Inicialmente o estudo baseou-se na identificação de interações putativas a partir de proteínas dos organismos depositadas em banco de dados, reconstruindo assim, uma rede de interação proteína-proteína. Como uma estratégia de validação da rede, comparou-se os transcriptomas do intestino e ovário de fêmeas de *A. sculptum* antes e após a infecção por *R. amblyommatis*. Assim, foram identificadas

proteínas e respectivas interações que melhor delineiam a interface *Rickettsia*-carrapato.

A análise da expressão diferencial foi realizada em órgãos específicos com infecção e sem infecção por *R. amblyommatis*. O nível de expressão do gene quantificado em transcritos por milhão (TPM). Um total de 144.836 genes foi identificado nestes órgãos, com 41.081, 1.399, 3.511 e 1.908 genes únicos transcritos no intestino, no intestino infectado, no ovário e no ovário infectado, respectivamente (Figura 4). A anotação dos transcritos revelou uma mudança no perfil gênico do intestino e do ovário de *A. sculptum* sugerindo que a presença da *R. amblyommatis* diminuiu a expressão de genes únicos em cada órgão.

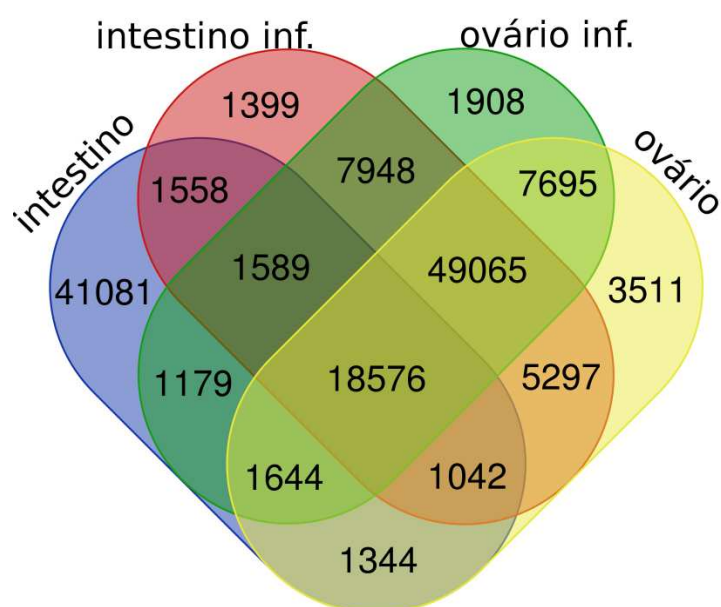


Figura 4. Diagrama de Venn – Quantificação de genes únicos ou comumente expressos em células do intestino e ovário de *A. sculptum* antes e após infecção por *R. amblyommatis*. O número sobreposto representa os genes mútuos diferencialmente expressos entre as diferentes comparações e os números não sobrepostos especificam os genes únicos para cada condição: não infectado e infectado.

O mapeamento dos genes revelou 1.558 genes mútuos no intestino não infectado e infectado e 7.695 para o ovário não infectado e infectado. Verificou-se uma maior co-expressão no ovário em relação ao intestino, reforçando a preferência da bactéria por este órgão. Assim, ao modular a maquinaria do carrapato a *R. amblyommatis* favorece sua manutenção, replicação e transmissão transovariana.

Além disso, observou-se 18.576 genes comumente expressos no intestino e ovário antes e após a infecção por *Rickettsia*, demonstrando se tratar de genes essenciais para a interface *Rickettsia*-carrapato.

A análise dos genes diferencialmente expressos categorizou 745 e 3.298 genes *down* regulados no intestino e no ovário, respectivamente. Para os genes *down* regulados 352 e 1.140 genes no intestino e no ovário, respectivamente (Tabela 01). O *fold change* foi definido pela divisão do TPM órgão infectado/TPM órgão não infectado. Transcritos regulados de forma positiva para *fold change* $\log_2 \geq 2$, e regulados de forma negativa para *fold change* $\log_2 \leq 0,5$.

Tabela 1. Genes *down* e *up* regulados no intestino e ovário de *A. sculptum*.

| Órgãos | Transcritos | Genes <i>down</i> regulados | Genes <i>up</i> regulados |
|-------------------------------------|-------------|-----------------------------|---------------------------|
| Intestino não infectado e infectado | 1.558 | 745 | 352 |
| Ovário não infectado e infectado | 7.695 | 3.298 | 1.140 |

Fold change: TPM órgão infectado/TPM órgão não infectado; genes *down* regulados $\leq 0,5$ e *up* regulados $\leq 2,0$.

Após a categorização dos genes diferencialmente expressos, identificaram-se 12 genes comuns as análises do interatoma de membrana e dos transcriptomas de intestino e ovário. Os genes, a saber: ISCW001070-PA.1 (Calmodulina putativa), ISCW004102-PA.1 (Calmodulina putativa), ISCW001611-PA.1 (Proteína putativa de ligação ao GTP (I) subunidade alfa (GNAI), ISCW004280-PA.1 (Proteína hidrolase hipotética), ISCW007387-PA.1 (Proteína putativa de choque térmico), ISCW000613-PA.1 (Proteína putativa contendo repetição de nhl), ISCW001382-PA.1 (Proteína putativa transmembrana rica em leucina), ISCW002948-PA.1 (Sinalização Gap1 putativa), ISCW003706-PA.1 (Proteína hipotética conservada), ISCW004245-PA.1 (Desidrogenase putativa), ISCW008317-PA.1 (Proteína putativa de ativação da fosfolipase A2), ISCW019641-PA.1 (Complexo putativo exosomal 3' – 5') (Figura 4).

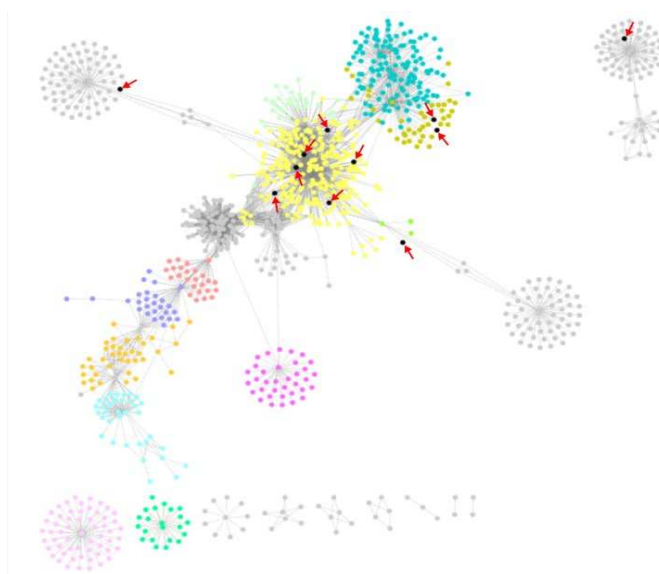


Figura 4. Rede de interação proteína-proteína *R. amblyommatis* e *I. scapularis* visualizada em nós (proteínas) e arestas (interações) pelo Cytoscape. Proteínas recuperadas após a integração do interatoma de membrana aos transcriptomas de ovário e intestino de *A. sculptum* em resposta a infecção por *R. amblyommatis*. Destaque em vermelho as proteínas de *I. scapularis*: ISCW001070-PA.1, ISCW004102-PA.1, ISCW001611-PA.1, ISCW004280-PA.1, ISCW007387-PA.1, ISCW000613-PA.1, ISCW001382-PA.1, ISCW002948-PA.1, ISCW003706-PA.1, ISCW004245-PA.1, ISCW008317-PA.1, ISCW019641-PA.1.

Para este grupo gênico, destaca-se os seguintes módulos funcionais, segundo a anotação do GO: secreção, fosforilação, fosforilação de aminoácidos e proteínas e alongamento translacional (Figura 4). Em uma análise comparativa a anotação funcional da rede de membrana, apenas 4 módulos do total de 10 se mantiveram após a integração da análise dos transcriptomas.

Ao analisar o TPM dos 12 genes e o local de expressão dos mesmos, verificou-se que o ovário se destacou, uma vez que, 10 genes foram expressos neste sítio. Tal achado reforça a hipótese de que o ovário é o órgão mais favorável ao estabelecimento, manutenção e transmissão para riquetsia (Figura 5).

Ao observar os valores de TPM dos 10 genes expressos no ovário e correlacioná-los, observou-se que em 40% dos genes houve um aumento do TPM em resposta à infecção por riquetsia. Este achado sugere que esta bactéria constitui uma carga metabólica e fisiológica para o órgão. A infecção riquetsial sugeriu uma regulação positiva da ISCW007387-PA.1/Proteína de choque térmico (estresse celular), ISCW002948-PA.1/Gap1 putativa (regulação da atividade da GTPase), ISCW001611-PA.1/Proteína putativa de ligação ao GTP (I) subunidade alfa (GNAI) (transdução de sinal) e ISCW004280-PA.1/Proteína hidrolase hipotética (ação catalítica, hidrólise de ligações peptídicas em L-aminoácidos). Os maiores valores de

fold change para o ovário foram 13,52 (ISCW007387-PA.1) e 1,66-2,62 (ISCW001611-PA.1, ISCW004280-PA.1 e ISCW002948-PA.1) (Figura 5).

Ademais, destaca-se a redução da expressão dos genes ISCW001070-PA.1 e ISCW004102-PA.1 no intestino em resposta a infecção por riquetsia, ambos correlacionam a isoformas da proteína calmodulina envolvida em mecanismos de transdução, sinalização do citoesqueleto e divisão celular (Figura 5). Tal, achado reforça que a bactéria pode estar regulando de forma negativa a expressão de genes do intestino e sobrepondo a regulação positiva de genes no ovário. Em resumo, a bactéria manipula diversos processos biológicos importantes na célula do carrapato de modo a garantir sua sobrevivência.

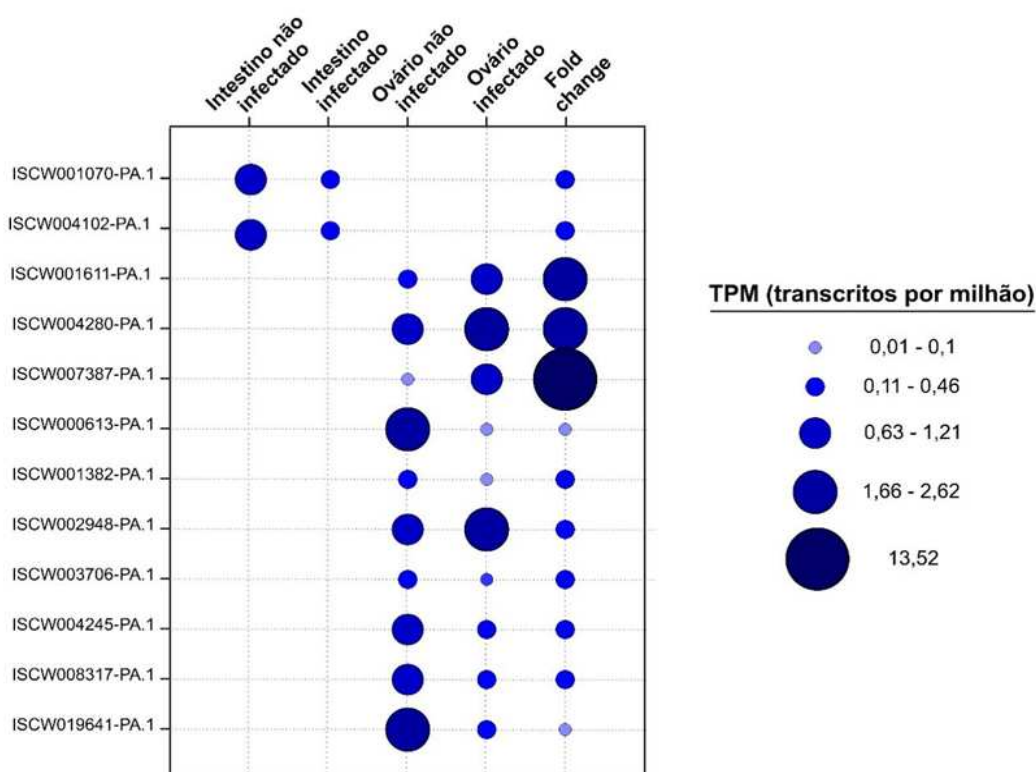


Figura 5. Transcritos por milhão (TPM) dos 12 genes recuperados após a integração do interatoma de membrana aos transcriptomas de ovário e intestino de *A. sculptum* sob condições de infecção e não infecção por *R. amblyommatis*.

Uma maneira de melhor compreender em nível molecular a interface patógeno-vetor é o estudo das interações proteicas estabelecidas pelos organismos. Desta forma, a partir das 12 proteínas do carrapato previamente selecionadas delinearam-se os pares de interação com a bactéria. Totalizando 42 pares de interação

compostos por 12 proteínas do carrapato e 29 proteínas da bactéria, onde o maior número de pares proteicos foi identificado no ovário (39 pares) (Tabela 2).

Tabela 2. Interações proteicas recuperadas após a integração do interatoma de membrana aos transcriptomas de intestino e ovário de *A. sculputum* sob condições de infecção por *R. amblyommatis*.

| ID Carrapato | Anotação | ID <i>Rickettsia amblyommatis</i> | Anotação | Intestino / Ovário | | |
|-----------------|---|-----------------------------------|--|-------------------------|---------------------------------|----------------------|
| ISCW001070-PA.1 | Calmodulina putativa | KJV62290.1 | Proteína do tipo PPIC com domínio PPIASE | Intestino - <i>down</i> | | |
| ISCW004102-PA.1 | Calmodulina putativa | KJV62290.1 | Proteína do tipo PPIC com domínio PPIASE | Intestino - <i>down</i> | | |
| ISCW001611-PA.1 | Proteína putativa de ligação ao GTP (I) subunidade alfa (GNAI) | KJV61259.1 | Fator de ativação | Intestino - <i>down</i> | | |
| | | KJV62362.1 | Fator IF-2 de iniciação da tradução | Ovário - <i>up</i> | | |
| | | KJV62767.1 | EngA GTPase associada ao ribossomo | Ovário - <i>up</i> | | |
| | | KJV61792.1 | Proteína LepA de ligação ao GTP | Ovário - <i>up</i> | | |
| | | KJV62635.1 | Proteína da família GTPase de ligação a subunidade 50S do ribossomo | Ovário - <i>up</i> | | |
| | | KJV62737.1 | Fator de alongamento da tradução Tu | Ovário - <i>up</i> | | |
| | | KJV61557.1 | Fator G de elongação da tradução | Ovário - <i>up</i> | | |
| | | KJV61260.1 | CgtA da família GTPase | Ovário - <i>up</i> | | |
| | | KJV61767.1 | Proteína TypA/BipA de ligação ao GTP | Ovário - <i>up</i> | | |
| | | KJV61115.1 | Guanilato cinase | Ovário - <i>up</i> | | |
| | | KJV61490.1 | Proteína de ligação ao GTP da biogênese do ribossomo YsxC | Ovário - <i>up</i> | | |
| | | KJV61098.1 | TrmE GTPase de modificação de tRNA | Ovário - <i>up</i> | | |
| | | KJV61538.1 | Proteína da família grpE | Ovário - <i>up</i> | | |
| ISCW004280-PA.1 | Proteína hipotética | KJV62746.1 | Proteína da família ecotina | Ovário - <i>up</i> | | |
| ISCW007387-PA.1 | Proteína putativa de choque térmico | KJV62696.1 | Proteína da família grpE | Ovário - <i>up</i> | | |
| | | KJV61666.1 | Proteína da família hsp70 | Ovário - <i>up</i> | | |
| | | KJV61119.1 | Proteína da família MreB/Mrl, determinante da forma celular | Ovário - <i>up</i> | | |
| | | KJV61763.1 | Exodesoxiribonuclease III | Ovário - <i>up</i> | | |
| | | KJV61634.1 | Proteína chaperona DnaK | Ovário - <i>up</i> | | |
| | | KJV62764.1 | Exodesoxiribonuclease III | Ovário - <i>up</i> | | |
| | | KJV61752.1 | Proteína FtsA de divisão celular | Ovário - <i>up</i> | | |
| | | KJV61854.1 | Proteína tol-Pal sistema beta propeller repeat proteina TolB | Ovário - <i>down</i> | | |
| | | ISCW001382-PA.1 | Proteína putativa transmembrana rica em leucina | KJV61538.1 | Proteína Era de ligação ao GTP | Ovário - <i>down</i> |
| | | | | KJV61792.1 | Proteína LepA de ligação ao GTP | Ovário - <i>down</i> |
| KJV62635.1 | Proteína da família GTPase de ligação a subunidade 50S do ribossomo | | | Ovário - <i>down</i> | | |
| ISCW002948-PA.1 | gap1 putativa | KJV61538.1 | Proteína Era de ligação ao GTP | Ovário - <i>down</i> | | |
| | | KJV61792.1 | Proteína LepA de ligação ao GTP | Ovário - <i>down</i> | | |
| | | KJV62635.1 | Proteína da família GTPase de ligação a subunidade 50S do ribossomo | Ovário - <i>down</i> | | |
| ISCW003706-PA.1 | Proteína hipotética conservada | KJV61792.1 | Proteína LepA de ligação ao GTP | Ovário - <i>down</i> | | |
| | | KJV62290.1 | Proteína do tipo PPIC com domínio PPIASE | Ovário - <i>down</i> | | |
| | | KJV62536.1 | Proteína da família com repetição de TRR | Ovário - <i>down</i> | | |
| | | KJV62562.1 | Chaperona LcrH/SycD em resposta a secreção do tipo III de pouco cálcio | Ovário - <i>down</i> | | |
| ISCW003706-PA.1 | Proteína hipotética conservada | KJV62591.1 | Proteína da família com repetição de tetratricopeptídeo | Ovário - <i>down</i> | | |
| | | KJV62635.1 | Proteína da família GTPase de ligação a subunidade 50S do ribossomo | Ovário - <i>down</i> | | |
| ISCW004245-PA.1 | Desidrogenase | KJV61968.1 | Redutase enoil-proteína acil carreadora | Ovário - <i>down</i> | | |

| | | | | |
|-----------------|--|------------|---|----------------------|
| | putativa | KJV62512.1 | Proteína da família redutase enoil-acil carreadora | Ovário - <i>down</i> |
| ISCW008317-PA.1 | Proteína putativa de ativação da fosfolipase A2 | KJV61538.1 | Proteína Era de ligação ao GTP | Ovário - <i>down</i> |
| | | KJV61792.1 | Proteína LepA de ligação ao GTP | Ovário - <i>down</i> |
| | | KJV6263 | Proteína da família GTPase de ligação a subunidade 50S do ribossomo | Ovário - <i>down</i> |
| ISCW019641-PA.1 | Complexo putativo exosomal 3'-5' exoribonuclease, subunidade Rrp43 | KJV62173 | polyribonucleotide nucleotidyltransferase | Ovário - <i>down</i> |

A partir das 42 interações obtidas neste estudo, selecionaram-se sete pares de interação, considerando proteínas envolvidas importante para a dinâmica patógeno-vetor segundo a literatura. Os pares escolhidos para a modelagem e *docking molecular*, a saber: ISCW001070-PA.1-KJV62290.1, ISCW002948-PA.1-KJV61792.1, ISCW003706-PA.1-KJV62290.1, ISCW003706-PA.1-KJV61792.1, ISCW004280-PA.1-KJV62746.1, ISCW008317-PA.1-KJV61538.1 e ISCW008317-PA.1-KJV61792.1. Vale ressaltar que a maior parte das interações ocorreu no ovário, demonstrando a importância deste sítio durante a infecção (Figura 6).

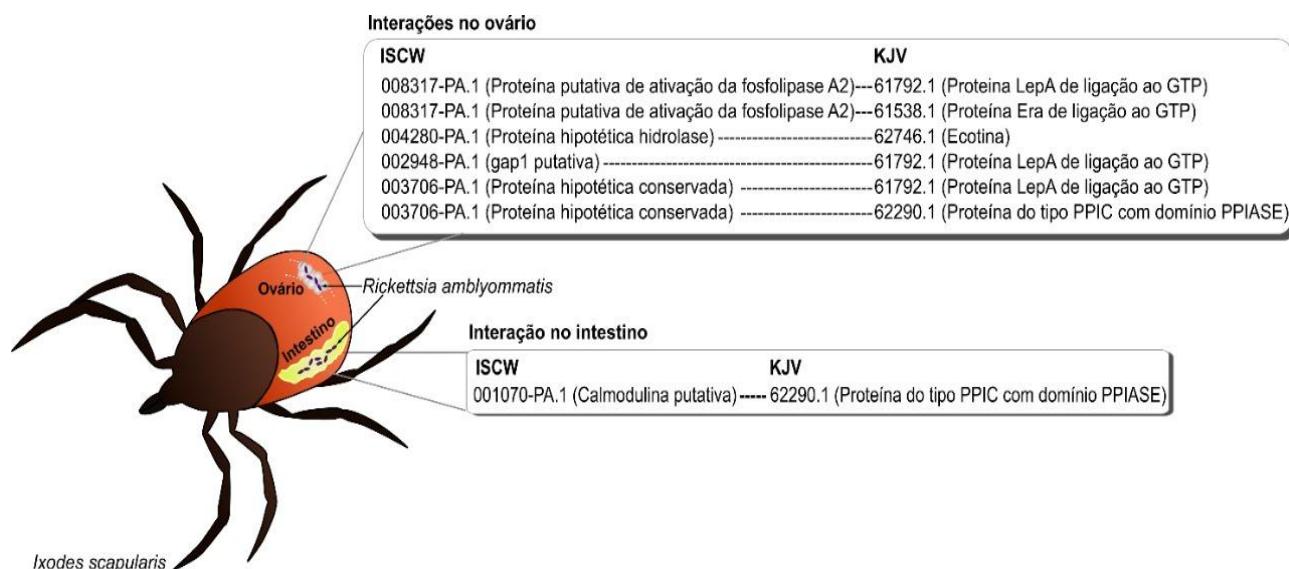


Figura 6. Proteínas do carrapato (ID-ISCW) e da bactéria (ID-KJV) e respectivos pares de interação órgão-específicos.

3.3.3. Modelagem e *Docking Molecular*

A modelagem das proteínas foi realizada com o auxílio do aplicativo do Phyre 2. A análise revelou modelos de alta confiança, uma vez que com todos os modelos gerados para as proteínas do carrapato (ISCW001070-PA.1, ISCW002948-PA.1,

ISCW003706-PA.1, ISCW004280-PA.1, ISCW008317-PA.1) e da bactéria (KJV61.538.1, KJV61.792.1, KJV62.290.1 e KJV62.746.1) apresentavam scores $\geq 99,5\%$ de confiança e $\geq 38\%$ cobertura do alinhamento. Para este estudo, definiu-se o ponto de corte mínimo para os escores de confiança iguais ou superiores a 90% e cobertura do alinhamento em 30% ou mais (Figura Suplementar 1).

O refinamento e minimização de energia dos modelos 3D foi realizado com o auxílio do aplicativo YASARA e, sendo os resultados obtidos em seguida avaliados pelos aplicativos PROCHECK, plot Ramachandran, ERRAT, ProSA e Verify 3D, antes e após o refinamento. O *plot* de Ramachandran realizado pelo PROCHECK revelou que as proteínas do carrapato ISCW001070-PA.1 e ISCW008317-PA.1 e da bactéria KJV61538.1 e KJV62290.1 apresentavam, respectivamente, 96,7%, 91,3%, 90,6% e 91,3% dos resíduos nas regiões mais favorecidas (região A, B, L representadas em vermelho no plot de Ramachandran), sugerindo modelos de boa qualidade estereoquímica (Figuras Suplementares 2, 3, 4, 5 e 6). Para as demais proteínas analisadas observou-se um aumento de resíduos na região mais favorecida após a minimização de energia, porém, não foi satisfatório para caracterizá-las como modelos de boa qualidade. As proteínas do carrapato ISCW001070-PA.1, ISCW004280-PA.1 e da bactéria KJV62746.1 não apresentaram ângulos distorcidos na cadeia principal, sendo considerado um parâmetro de boa qualidade dos modelos (Figuras Suplementares 3, 7 e 8). Os parâmetros do PROCHECK utilizados como definidores de um modelo com qualidade estereoquímica foram: $\geq 90\%$ de seus resíduos nas regiões mais favorecidas: A, B e L; ausência de resíduos na região desfavorável: XX; e ausência de ângulos distorcidos na cadeia principal.

Além disso, a confiabilidade dos modelos 3D foi confirmada pelo aplicativo ERRAT. O servidor utiliza como indicador o fator de qualidade geral, valores >50 caracterizam modelos de alta qualidade (MESSAOUDI et al., 2013). Neste estudo, para todos os modelos 3D analisados após a minimização de energia, o fator de qualidade geral apresentou-se $\geq 90,9091$, sugerindo boa qualidade (Figura Suplementar 2).

Para análise da adequação da sequência à estrutura foram utilizados os aplicativos Verify3D e ProSA. Um dos modelos 3D do carrapato, a proteína hipotética conservada (ISCW003706-PA.1) foi aprovada segundo o Verify3D quando analisada antes da minimização de energia. Observamos que 91,7% dos resíduos

pontuaram $\geq 0,2$ no perfil 3D/1D (Figura Suplementar 9). Os *scores* dos resíduos de todos os modelos analisados após a minimização de energia pelo Verify 3D foram inferiores a 80%, o que significa não atingiram a pontuação $\geq 0,2$ no perfil 3D/1D, sendo assim, categorizados como modelos de baixa qualidade segundo o servidor (Figura Suplementar 2).

Os *scores* Z referentes ao ProSA de todos os modelos analisados antes e após a minimização de energia apresentaram pontuação negativa, o que sugere similaridade com a estrutura de ressonância magnética nuclear (RMN) sendo classificados como modelos de alta qualidade (Figuras Suplementares 3 a 11).

Com base nos parâmetros geométricos e estereoquímicos todos os modelos foram encaminhados para o *docking* molecular. O *docking* molecular realizou a predição de interações funcionais com base no número máximo de resíduos de contato e na inspeção visual da região da interface (SIM & Shruti, 2019). O servidor O servidor ClusPro foi utilizado para encaixe proteína-proteína (Figura 7).

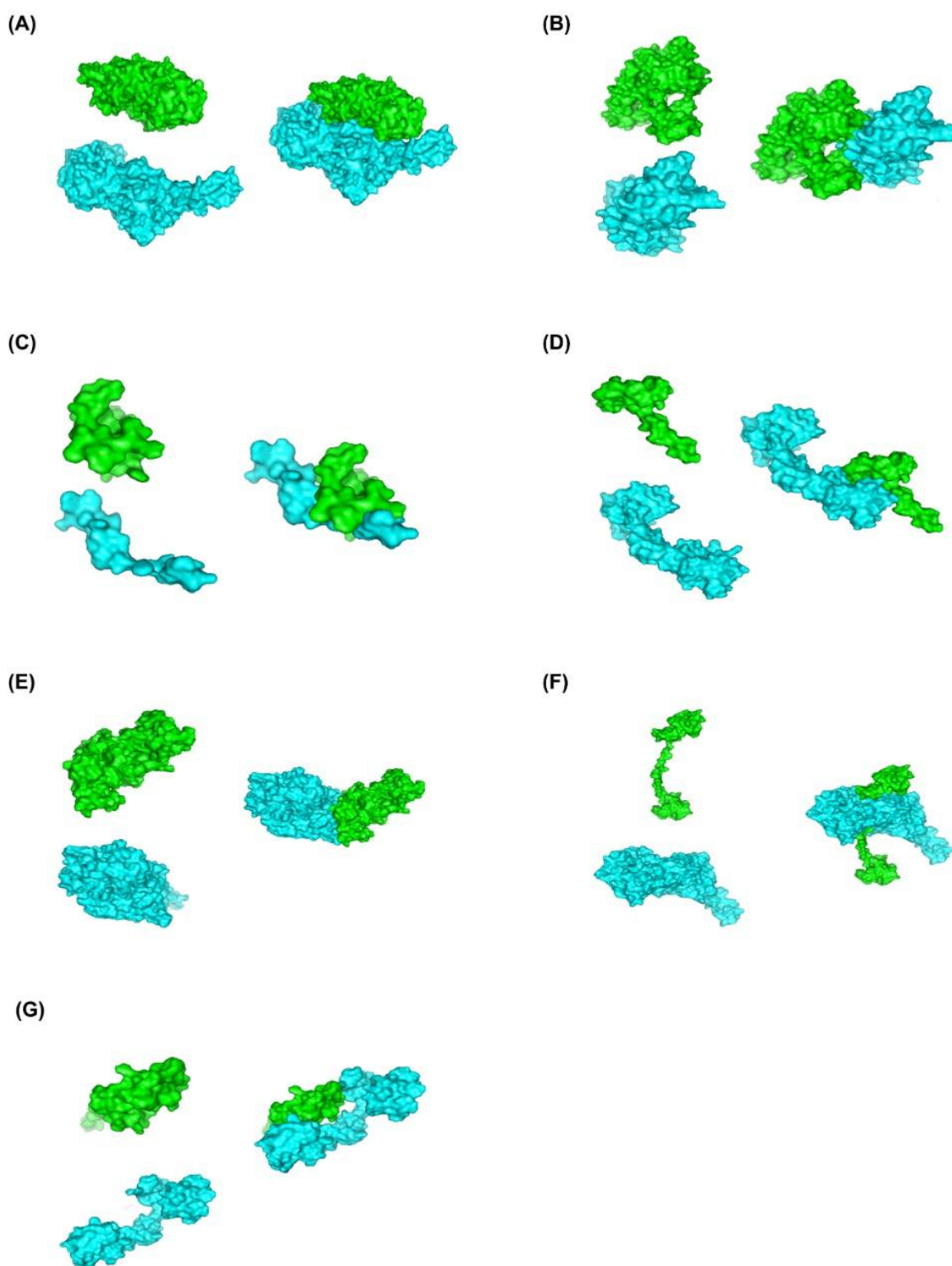


Figura 7. Interação proteína-proteína utilizando o *docking* molecular pelo ClusPro. Em verde, as proteínas do carrapato (receptor) e em azul da bactéria (ligante) (A) Interação A - Proteína putativa de ativação da fosfolipase A2 e Proteína LepA de ligação ao GTP. (B) Interação B- Proteína putativa de ativação da fosfolipase A2 e Proteína Era de ligação ao GTP. (C) Interação C- Proteína hipotética-hidrolase e Ecotina. (D) Interação D- Calmodulina e Proteína do tipo PPIC com domínio PPIASE. (E) Interação E -Gap 1 putativa e Proteína LepA de ligação ao GTP. (F) Interação F- Proteína hipotética conservada e Proteína LepA de ligação ao GTP. (G) Interação G- Proteína hipotética conservada e Proteína do tipo PPIC com domínio PPIASE.

Os resultados do encaixe foram armazenados na forma de diferentes *clusters*; o *cluster* '0' possui a menor energia entre todos os *clusters*. Portanto, para pontuar os resultados de encaixe, o *cluster* '0' foi selecionado para simulações por dinâmica molecular dos complexos de proteína bactéria-carrapato em bicamada lipídica POPC. É importante notar que a energia mais baixa na forma de pontuação de energia negativa reflete maior afinidade. Os perfis de energia obtidos indicaram alta probabilidade para as interações intA, intC, intD e intG, em contrapartida, as intB e intE não foram suportadas termodinamicamente (Tabela 3).

Tabela 3. Simulações por dinâmica molecular dos complexos de proteína bactéria-carrapato em bicamada lipídica POPC (1-Palmitoil-2-oleoilfosfatidilcolina).

| ClusPro | Proteína do carrapato | Proteína da bactéria | Energia livre predita por dinâmica molecular (kcal/mol) |
|---------|-----------------------|----------------------|---|
| IntA | ISCW008317-PA.1 | KJV 61792.1 | -12,81 |
| IntB | ISCW008317-PA.1 | KJV 61538.1 | -1,73 |
| IntC | ISCW004280-PA.1 | KJV 62746.1 | -16,42 |
| IntD | ISCW001070-PA.1 | KJV 62290.1 | -8,59 |
| IntE | ISCW002948-PA.1 | KJV 61792.1 | -2,98 |
| IntF | ISCW003706-PA.1 | KJV 61792.1 | -12,92 |
| IntG | ISCW003706-PA.1 | KJV 62290.1 | -17,46 |

Int: interação

A análise de *docking* molecular revelou cinco complexos potenciais, sendo eles: calmodulina putativa-proteína do tipo PPIC com domínio PPIASE, proteína hipotética conservada-proteína do tipo PPIC com domínio PPIASE, proteína hipotética conservada e proteína LepA de ligação ao GTP, proteína hipotética hidrolase-ecotina, proteína putativa de ativação da fosfolipase A2 e proteína Era de ligação ao GTP. Essas interações servirão como *insights* para elucidar os mecanismos moleculares da interface riquetsia-carrapato.

3.4. DISCUSSÃO

As interações proteína-proteína são a interface funcional de muitos processos biológicos, incluindo interações entre patógenos e vetores (SCHWEPPE et al., 2015). Este estudo recuperou proteínas a partir da abordagem computacional para melhor compreender a relação *Rickettsia*-carrapato através da reconstrução da rede de interação proteína-proteína. O número de proteínas recuperadas (91 *Rickettsia* e 980 carrapato) pode estar associado: (i) uma maior quantidade de proteínas codificadas nos genomas dos organismos, (ii) isoformas para uma mesma proteína que são representadas como vértices diferentes na rede, (iii) e, ou, enriquecimento de interações entre genes de ambos os organismos presente nos bancos de dados utilizados (BRITO & PINNEY, 2017).

Estendeu-se esta análise centrada na rede de membrana por considerar a natureza intracelular obrigatória da *Rickettsia*; o que requer que as bactérias invadam as células hospedeiras (PETCHAMPAI et al., 2015). Desta forma, as proteínas de membrana apresentam maior probabilidade de participar do processo de infecção e serem possíveis alvos moleculares. Proteínas de membrana representam cerca de 60% dos alvos terapêuticos e de medidas de controle (YILDIRIM et al. 2007; BAKHEET & DOIG 2009). Estudos anteriores relataram o funcionamento do secretoma de *Rickettsia* e como ocorre a possível interação entre este repertório de proteínas e vias de sinalização com os carrapatos ((de la FUENTE et al., 2010).

As 42 interações putativas identificadas neste estudo foram associadas aos processos sinalização, secreção de proteínas, regulação da localização celular, regulação do transporte intracelular e fosforilação, tendo como significado fisiológico o estabelecimento e sobrevivência da bactéria nas células do carrapato, principalmente no ovário.

Diante do contexto, nossos achados catalogaram interações que ocorreram envolvendo proteínas de membrana de riquétsia. As proteínas de membrana HSP70 (KJV61666.1) e DnaK (KJV61634.1) de riquétsia parecem atuar em vias de sinalização (DUMLER, et al., 2001), ambas tem como ligante a proteína de choque térmico putativa do carrapato. Estudos demonstraram que proteína de superfície principal 4 de *Anaplasma phagocytophilum* (MSP4) e a proteína de choque térmico

70 (HSP70) mostraram-se localizadas na membrana bacteriana, com um possível papel durante a infecção por patógenos em carrapatos e na multiplicação em neutrófilos dos hospedeiros (CONTRERAS et al., 2017).

A proteína de riquetsia KJV61792.1/LepA de ligação ao GTP que interagiu com cinco diferentes proteínas do carrapato (ISCW001611-PA.1/proteína putativa de ligação ao GTP, ISCW008317-PA.1/proteína putativa de ativação da fosfolipase A2, ISCW002948-PA.1/gap 1 putativa, ISCW004280-PA.1/proteína hipotética conservada, ISCW001382-PA.1/proteína putativa transmembranarica em leucina e ISCW003706-PA.1). Numa abordagem molecular, verificou-se que as interações proteína LepA de ligação ao GTP e seus respectivos interativos, gap1 putativa e proteína hipotética conservada foram termodinamicamente viáveis. A proteína LepA é um fator secretado pelo T4SS (sistema de secreção tipo IV) regulando positivamente a síntese de nutrientes que permitem que as riquetsias sobrevivam em ambientes estressantes (UCHIYAMA, 2012). Nossos resultados reafirmaram a importância desta proteína, uma vez que participou de um número maior de interações.

As interações carrapato-riquetsia, a saber, calmodulina-proteína tipo PPIC com domínio PPIASE e proteína hipotética conservada-proteína tipo PPIC com domínio PPIASE foram descritas como termodinamicamente estáveis. As proteínas PPIase bacteriana estão envolvidas no tráfico de proteínas, arranjo ou maturação de proteínas da membrana externa e virulência bacteriana (BEHRENS et al., 2001; EMELYANOV & DEMYANOVA, 1999). Em relação à infecção bacteriana, foi proposto que as proteínas com domínio PPIase podem alterar a conformação dos componentes bacterianos ou da superfície da célula hospedeira, facilitando a adesão entre bactérias e células hospedeiras (THEPPARIT et al., 2010).

As proteínas do tipo GTPases de riquetsia também se destacaram principalmente promovendo interações proteicas com proteínas do carrapatos durante a infecção no ovário, tais proteínas são essenciais para a dinâmica celular. As GTPases Rac1 e Rac2, o complexo fator promotor da nucleação WAVE e o complexo Arp2/3 de *Rickettsia parkeri* contribuem para a invasão em hospedeiros (REED et al., 2012).

Também verificamos que o complexo proteína hidrolase conservada-ecotina se mostrou termodinamicamente estável. A proteína ecotina de *E.coli* protege da

elastase de neutrófilos (GILLESPIE et al., 2008), inibe várias serinas proteases pancreáticas, incluindo quimotripsina, tripsina, elastases, fator X, calicreína, bem como uma variedade de outras proteases (EGGERS et al., 2004).

Ao correlacionar as interações ao local da infecção, observamos que exposição à riquetsia modulou a transcrição gênica do carrapato, numa resposta tecido-específica para intestino e ovário. Este estudo sugere que a *Rickettsia amblyommatis* encontrou no ovário um nicho mais favorável, uma vez que foram recuperadas neste órgão um maior número de interações. O ovário se destaca como a principal estrutura biológica do carrapato para a biologia, uma vez que desempenha um papel na transmissão e amplificação de riquetsias no ambiente (SOARES et al., 2012).

Finalizamos destacando cinco potenciais complexos proteicos: calmodulina putativa-proteína do tipo PPIC com domínio PPIASE, proteína hipotética conservada-proteína do tipo PPIC com domínio PPIASE, proteína hipotética conservada e proteína LepA de ligação ao GTP, proteína hipotética hidrolase-ecotina, proteína putativa de ativação da fosfolipase A2 e proteína Era de ligação ao GTP. Provavelmente esses complexos tenham significado funcional na manutenção da riquetsia em células do carrapato, principalmente no ovário. Sugerimos que a espécie *R. amblyommatis* modulou a expressão de genes preferencialmente no ovário, destacando este órgão como um importante sítio para sua manutenção, multiplicação e transmissão.

Em conclusão o estudo catalogou 42 interações proteicas da interface *Rickettsia*-carrapato, expressos no carrapato em resposta à infecção por *Rickettsia*, 5 complexos proteicos candidatos a validação experimental derivados de evidências *in silico* que servirão como *insights* para decifrar os mecanismos de infecção, além da seleção de alvos moleculares.

3.5. REFERÊNCIAS BIBLIOGRÁFICAS

ALTSCHUL, S. F. et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389-3402, 1997.

BARABASI, A. L.; OLTVAI, Z. N. Network biology: understanding the cell's functional organization. **Nature Reviews Genetics**, v. 5, n. 2, p. 101-113, 2004.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114-2120, 2014.

BRAY, N. et al. Near-optimal RNA-Seq quantification. **Nature Biotechnology**, v. 34, p. 525–527, 2016.

BRITO, A. F.; PINNEY, J. W. Protein–protein interactions in virus–host systems. **Frontiers in Microbiology**, v. 8, p. 1557, 2017.

CAMON, E. et al. The gene ontology annotation (goa) database: sharing knowledge in uniprot with gene ontology. **Nucleic Acids Research**, v. 32, n. suppl_1, p. D262-D266, 2004.

DE LA FUENTE, J. et al. Strategies for new and improved vaccines against ticks and tick-borne diseases. **Parasite Immunology**, v. 38, n. 12, p. 754-769, 2016.

DURMUŞ TEKİR, S. D.; ÜLGEN, K. Ö. Systems biology of pathogen-host interaction: Networks of protein-protein interaction within pathogens and pathogen-human interactions in the post-genomic era. **Biotechnology Journal**, v. 8, n. 1, p. 85-96, 2013.

EREMEEVA, M. E.; DASCH, G. A. Challenges posed by tick-borne rickettsiae: eco-epidemiology and public health implications. **Frontiers in Public Health**, v. 3, p. 55, 2015.

FINN, R. D. et al. Pfam: clans, web tools and services. **Nucleic Acids Research**, v. 34, n. suppl_1, p. D247-D251, 2006.

FLÓREZ, A. F. et al. Protein network prediction and topological analysis in *Leishmania major* as a tool for drug target selection. **BMC Bioinformatics**, v. 11, n. 1, p. 484, 2010.

FREY, B. J.; DUECK, D. Clustering by passing messages between data points. **Science**, v. 315, n. 5814, p. 972-976, 2007.

GARDY, J.L. et al. PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. **Bioinformatics**, v. 21, n. 5, p. 617-623, 2005.

GRABHERR, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. **Nature Biotechnology**, v. 29, n. 7, p. 644, 2011.

GULIA-NUSS, M. et al. Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease. **Nature Communications**, v. 7, n. 1, p. 1-13, 2016.

HASE, Takeshi et al. Structure of protein interaction networks and their implications on drug design. **PLoS Computational Biology**, v. 5, n. 10, 2009.

HESS, B. et al. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. **Journal of Chemical Theory and Computation**, v. 4, n. 3, p. 435-447, 2008.

HUBBARD, T. J. P. et al. SCOP: a structural classification of proteins database. **Nucleic Acids Research**, v. 25, p. 236-239, 1997.

HUMPHREY, W. et al. VMD: visual molecular dynamics. **Journal of Molecular Graphics**, v. 14, n. 1, p. 33-38, 1996.

HWANG, W-C.; ZHANG, A.; RAMANATHAN, M. Identification of information flow-modulating drug targets: a novel bridging paradigm for drug discovery. **Clinical Pharmacology & Therapeutics**, v. 84, n. 5, p. 563-572, 2008.

JONGEJAN, F.; UILENBERG, G. The global importance of ticks. **Parasitology**, v. 129, n. S1, p. S3-S14, 2004.

KALL, L. et al. An HMM posterior decoder for sequence feature prediction that includes homology information. **Bioinformatics**, p.251-257, 2005.

KAR, G.; GURSOY, A.; KESKIN, O. Human cancer protein-protein interaction network: a structural perspective. **PLoS Computational Biology**, v. 5, n. 12, 2009.

KELLEY, L. A, STERNBERG, M. J. Protein structure prediction on the Web: a case study using the Phyre server. **Nature Protocols**, v. 4, p. 363-371, 2009.

KERSEY, P. J. et al. Ensembl Genomes: an integrative resource for genome-scale data from non-vertebrate species. **Nucleic Acids Research**, v. 40, n. D1, p. D91-D97, 2012.

KIM, J. G. et al. Predicting the interactome of *Xanthomonas oryzae* pathovar *oryzae* for target selection and DB service. **BMC Bioinformatics**, v. 9, n. 1, p. 41, 2008.

KOZAKOV, D. et al. How good is automated protein docking? **Proteins: Structure, Function, and Bioinformatics**, v. 81, n. 12, p. 2159-2166, 2013.

KOZAKOV, Dima et al. The ClusPro web server for protein–protein docking. **Nature Protocols**, v. 12, n. 2, p. 255, 2017.

KRIEGER, E. et al. Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. **Proteins: Structure, Function, and Bioinformatics**, v. 77, n. S9, p. 114-122, 2009.

KROGH, A. et al. Hidden Markov models in computational biology. Applications to protein modeling. **Journal of Molecular Biology**, v. 235, n. 5, p. 1501-1531, 1994.

LACOUNT, D. J. et al. A protein interaction network of the malaria parasite *Plasmodium falciparum*. **Nature**, v. 438, n. 7064, p. 103-107, 2005.

LOMIZE, M. A. et al. OPM database and PPM web server: resources for positioning of proteins in membranes. **Nucleic Acids Research**, v. 40, n. D1, p. D370-D376, 2012.

MAERE, S.; HEYMANS, K.; KUIPER, M. BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. **Bioinformatics**, v. 21, n. 16, p. 3448-3449, 2005.

MCCARTHY, F. M. et al. AgBase: a functional genomics resource for agriculture. **BMC Genomics**, v. 7, n. 1, p. 229, 2006.

MENDES, T. A. O. et al. Repeat-enriched proteins are related to host cell invasion and immune evasion in parasitic protozoa. **Molecular Biology and Evolution**, v. 30, n. 4, p. 951-963, 2013.

MOREIRA, H. N. S. et al. A deep insight into the whole transcriptome of midguts, ovaries and salivary glands of the *Amblyomma sculptum* tick. **Parasitology International**, v. 66, n. 2, p. 64-73, 2017.

MORRIS, G. M. et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. **Journal of Computational Chemistry**, v. 30, n. 16, p. 2785-2791, 2009.

MULLARD, A. Protein-protein interaction inhibitors get into the groove. **Nature Reviews Drug Discovery**, v. 11, p. 1730175, 2012.

NUNES, P. H.; BECHARA, G. H.; CAMARGO-MATHIAS, M. I. Morphological changes in the salivary glands of *Amblyomma cajennense* females (*Acari: Ixodidae*) in different feeding stages on rabbits at first infestation. **Experimental and Applied Acarology**, v. 45, n. 3-4, p. 199, 2008.

OTTMAN, C. Protein-protein interactions: an overview. In MANNHOLD R et al. (eds): **Protein-Protein Interactions in Drug Discovery**. Singapore, Wiley-VCH Verlag GmbH & Co, p. 1-19, 2013.

PARK, D. et al. Comparative interactomics analysis of protein family interaction networks using PSIMAP (protein structural interactome map). **Bioinformatics**, v. 21, n. 15, p. 3234-3240, 2005.

PAROLA, P. et al. Update on tick-borne rickettsioses around the world: a geographic approach. **Clinical Microbiology Reviews**, v. 26, n. 4, p. 657-702, 2013.

PAROLA, P.; LABRUNA, M. B.; RAOULT, D. Tick-borne rickettsioses in America: unanswered questions and emerging diseases. **Current Infectious Disease Reports**, v. 11, n. 1, p. 40-50, 2009.

PETCHAMPAL, N. et al. Identification of host proteins involved in rickettsial invasion of tick cells. **Infection and Immunity**, v. 83, n. 3, p. 1048-1055, 2015.

PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nature Methods**, v. 8, n. 10, p. 785, 2011.

- RAMAN, K.; CHANDRA, N. Mycobacterium tuberculosis interactome analysis unravels potential pathways to drug resistance. **BMC Microbiology**, v. 8, n. 1, p. 234, 2008.
- RAOULT, D. Emerging rickettsioses reach the United States. **Clinical Infectious Diseases**, v. 51, n. 1, p. 121-122, 2010.
- RICHARDS, A. L. Worldwide detection and identification of new and old rickettsiae and rickettsial diseases. **FEMS Immunology & Medical Microbiology**, v. 64, n. 1, p. 107-110, 2012.
- SAITO, R. et al. A travel guide to Cytoscape plugins. **Nature Methods**, v. 9, n. 11, p. 1069, 2012.
- SCHWEPPE, D.K.; HARDING, C.; Chavez, J.D.; WU X.; RAMAGE, E.; SINGH, P.K.; MANOIL, C.; BRUCE, J.E. Host-microbe protein interactions during bacterial infection. *Chem. Biol.*, v. 22, p.1521–1530, 2015.
- SNEL, B.; BORK, P.; HUYNEN, M. Genome evolution: gene fusion versus gene fission. **Trends in Genetics**, v. 16, n. 1, p. 9-11, 2000.
- SONNHAMMER, E. L. et al. A hidden Markov model for predicting transmembrane helices in protein sequences. In: **ISMB**, p. 175-182, 1998.
- STAPELS, D. A. C; GEISBRECHT, B. V.; ROOIJAKKERS, S. H. M. Neutrophil serine proteases in antibacterial defense. **Current Opinion in Microbiology**, v. 23, p. 42-48, 2015.
- SZKLARCZYK, D. et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. **Nucleic Acids Research**, v. 39, n. suppl_1, p. D561-D568, 2010.
- TAYLOR, C. M. et al. Targeting protein-protein interactions for parasite control. **PLoS One**, v. 6, n. 4, 2011.
- UCHIYAMA, T. Tropism and pathogenicity of rickettsiae. **Frontiers in Microbiology**, v. 3, p. 230, 2012.
- VAN DER AUWERA, G. A. et al. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. **Current Protocols in Bioinformatics**, v. 43, n. 1, p. 11.10. 1-11.10. 33, 2013.
- VON MERING, C. et al. STRING: a database of predicted functional associations between proteins. **Nucleic Acids Research**, v. 31, n. 1, p. 258-261, 2003.
- WALLNER, B.; ELOFSSON, A. Can correct protein models be identified?. **Protein Science**, v. 12, n. 5, p. 1073-1086, 2003.
- WIEDERSTEIN, M.; SIPPL, M. J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, v. 35, n. suppl_2, p. W407-W410, 2007.
- YANAI, I.; DERTI, A.; DELISI, C. Genes linked by fusion events are generally of the same functional category: a systematic analysis of 30 microbial genomes.

Proceedings of the National Academy of Sciences, v. 98, n. 14, p. 7940-7945, 2001.

YU, C. S.; LIN, C. J.; HWANG, J. K. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. **Protein Science**, v. 13, n. 5, p. 1402-1406, 2004.

CONSIDERAÇÕES FINAIS

Este estudo é inovador para área ao utilizar a abordagem computacional para reconstruir o interatoma e elucidar os mecanismos moleculares e funcionais que caracterizam a relação *Rickettsia*-carrapato. Observou-se uma modulação gênica preferencial de *R. amblyommatis* pelo ovário do carrapato. Uma vez, que tal ação contribui para a sua replicação eficiente e sua manutenção em longo prazo no ambiente celular do carrapato, assim como a transmissão.

Em resumo, destacamos 42 interações dinâmicas presentes nesta interface, 12 genes diferencialmente expressos no carrapato em resposta à infecção por *Rickettsia* e 5 complexos proteicos candidatos a validação experimental derivados de evidências *in silico* que servirão como *insights* para decifrar os mecanismos de infecção, além da seleção de alvos moleculares.

MATERIAL SUPLEMENTAR

Figura Suplementar 1. Avaliação estrutural e anotação das proteínas de *I. scapularis* e *R. amblyommatis* pelo aplicativo Phyre 2.

| ID <i>I. scapularis</i> | Número de resíduos | Score Confiança | Score Cobertura | Anotação |
|--|----------------------------|------------------------|------------------------|---|
| ISCW001070-PA.1 | 70 (99% da sua sequência) | 99,5% | 99% | Calmodulina putativa |
| ISCW002948-PA.1 | 300 (40% da sua sequência) | 100% | 40% | Gap1 putativa |
| ISCW003706-PA.1 | 156 (45% da sua sequência) | 100% | 45% | Proteína hipotética conservada |
| ISCW004280-PA.1 | 56 (98% da sua sequência) | 99,5% | 98% | Proteína hidrolase hipotética |
| ISCW008317-PA.1 | 309 (38% da sua sequência) | 100% | 38% | Proteína putativa de ativação da fosfolipase A2 |
| ID <i>Rickettsia amblyommatis</i> | Número de resíduos | Score Confiança | Score Cobertura | Anotação |
| KJV61538.1 | 290 (85% da sua sequência) | 100% | 85% | Proteína Era de ligação ao GTP |
| KJV61792.1 | 569 (95% da sua sequência) | 100% | 95% | Proteína LepA de ligação ao GTP |
| KJV62290.1 | 247 (88% da sua sequência) | 100% | 88% | Proteína do tipo PPIC com domínio PPIASE |
| KJV62746.1 | 31 (84% da sua sequência) | 99,6% | 84% | Ecotina, inibidor da tripsina |

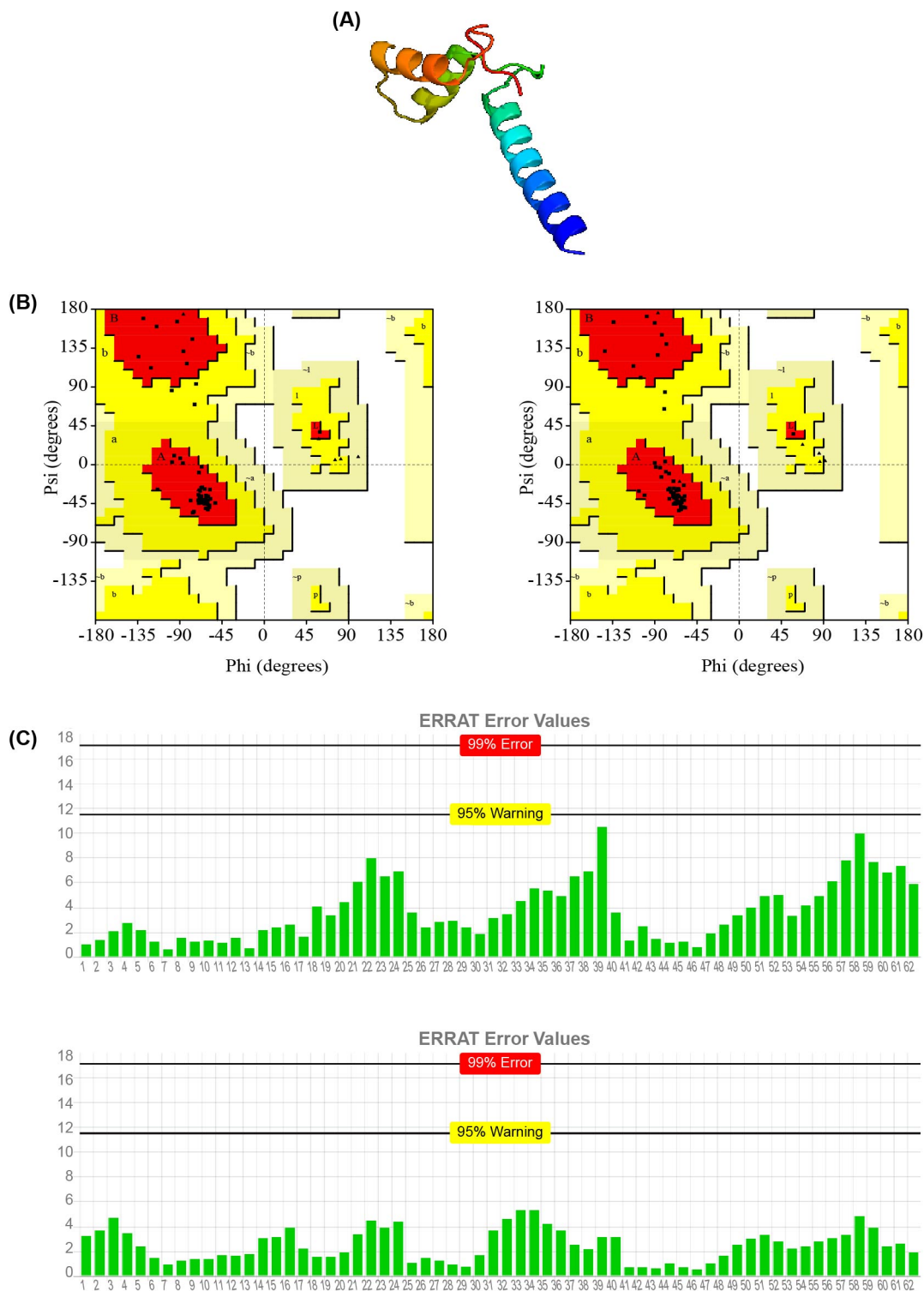
Nota: Parâmetros de avaliação da estrutura 3D: Score de confiança >90% e score de cobertura ≥30%.

Figura 2 Suplementar. Validação dos modelos 3D de proteínas de *I. scapularis* e *R. amblyommatis* gerados pelo aplicativo Phyre 2 antes e após minimização de energia pelo aplicativo YASARA.

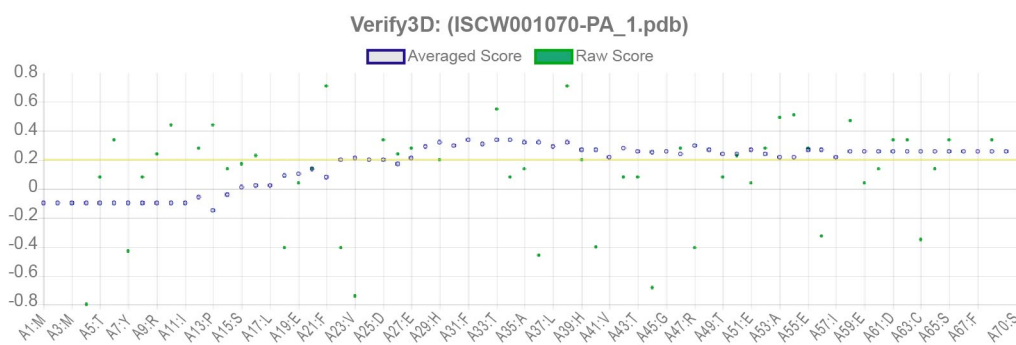
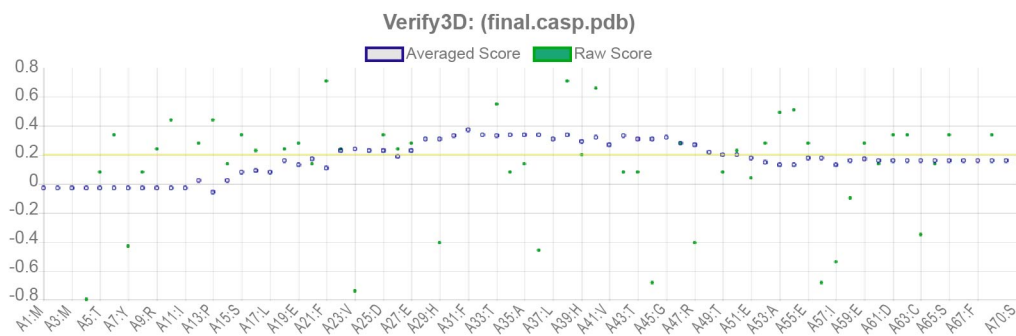
| Proteína | PROCHECK | | ERRAT | | ProSA | | Verify 3D | |
|----------------------------------|-------------------------------|------|----------------------|---------|-----------|-------|------------------------|-------|
| | (Distorted main-chain angles) | | (Fator de qualidade) | | (Score Z) | | (Score dos resíduos %) | |
| | Antes | Após | Antes | Após | Antes | Após | Antes | Após |
| ID <i>I. scapularis</i> | | | | | | | | |
| ISCW001070-PA.1 | 0 | 0 | 100 | 100 | -2,26 | -3,35 | 40,00 | 68,57 |
| ISCW002948-PA.1 | 78 | 9 | 50,1684 | 94,3662 | -4,40 | -5,05 | 74,03 | 68,51 |
| ISCW003706-PA.1 | 16 | 2 | 91,6667 | 92,3077 | -5,35 | -4,47 | 91,72 | 65,61 |
| ISCW004280-PA.1 | 6 | 0 | 40,505 | 90,9091 | -3,12 | -3,78 | 48,21 | 64,29 |
| ISCW008317-PA.1 | 56 | 13 | 58,8235 | 97,9592 | -6,95 | -7,05 | 52,65 | 64,17 |
| ID <i>R. amblyommatis</i> | | | | | | | | |
| KJV61538.1 | 17 | 1 | 58,3039 | 94,6996 | -8,64 | -8,64 | 66,67 | 75,6 |
| KJV61792.1 | 35 | 21 | 39,1304 | 100 | -8,87 | -9,84 | 42,71 | 77,68 |
| KJV62290.1 | 44 | 1 | 55,9671 | 97,1193 | -4,77 | -5,88 | 26,98 | 31,35 |
| KJV62746.1 | 0 | 0 | 100 | 100 | -1,77 | -1,92 | 0 | 0 |

Nota: Parâmetros de qualidade dos modelos 3D: PROCHECK: ângulos distorcidos da cadeia principal = 0; ERRAT: fator de qualidade geral >50; ProSA: Score Z <0; e Verify 3D: Score dos resíduos ≥80% com pontuação ≥0, 2 no perfil 3D/1D.

Figura Suplementar 3. (A) Predição do modelo 3D da proteína ISCW001070-1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score* Z obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

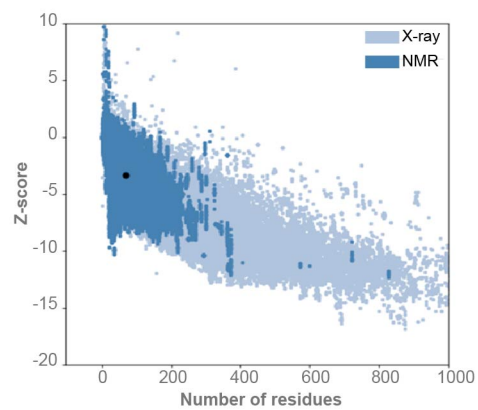
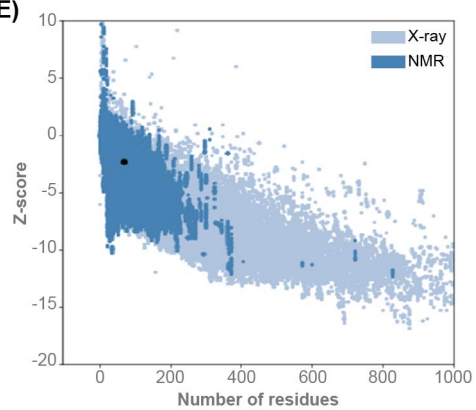
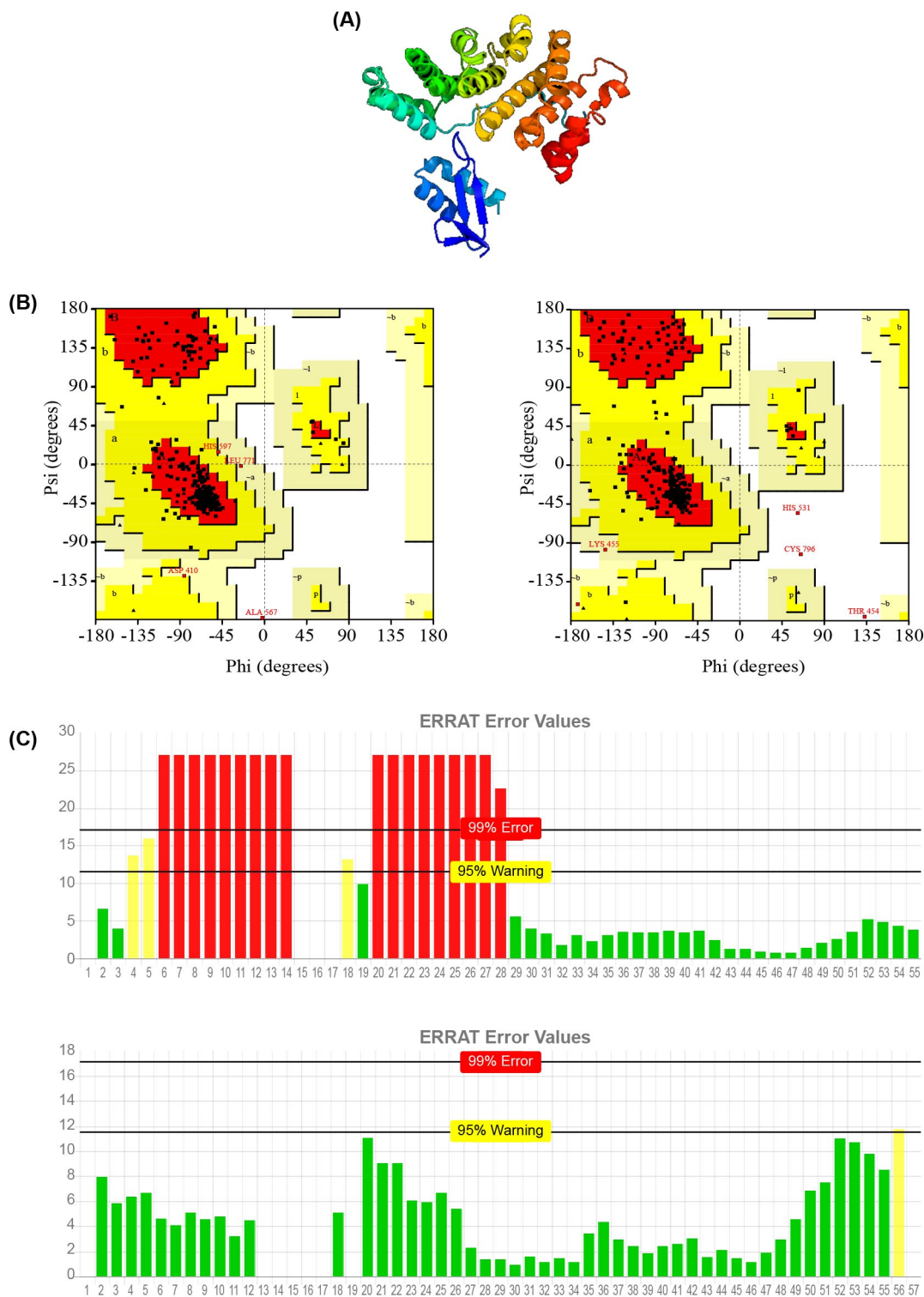
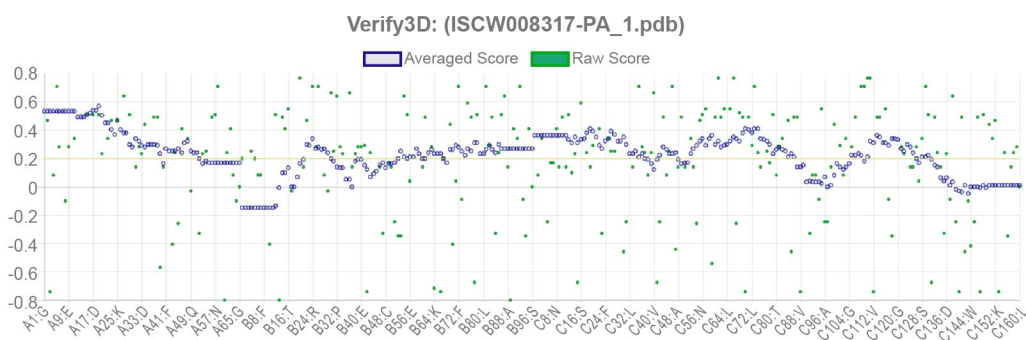
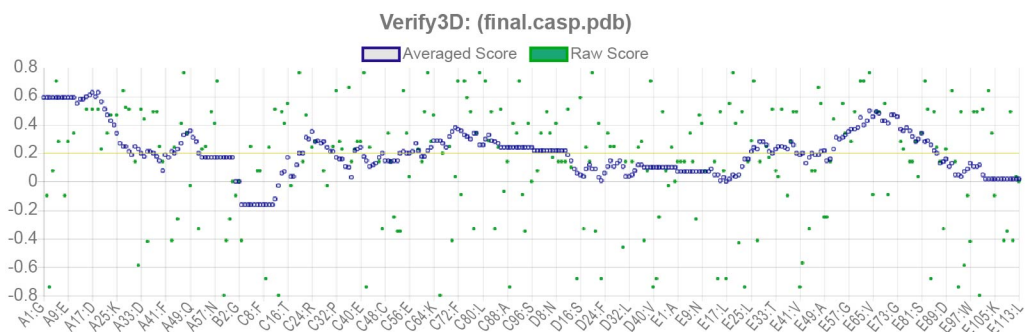


Figura Suplementar 4. (A) Predição do modelo 3D da proteína ISCW008317-1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score* Z obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

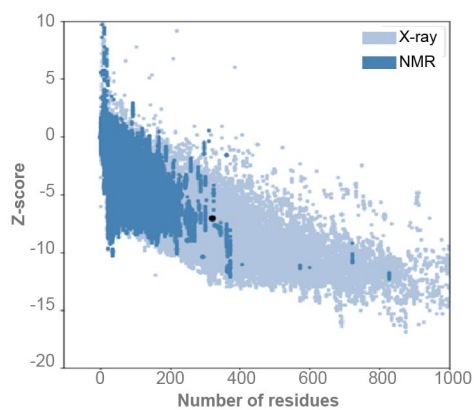
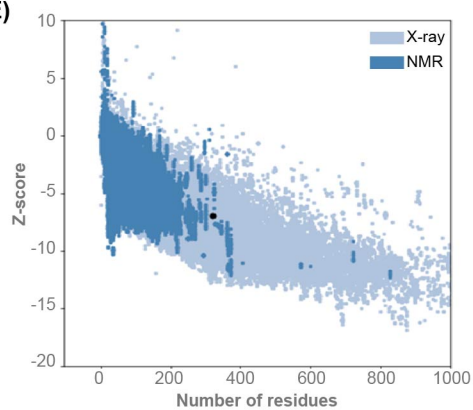
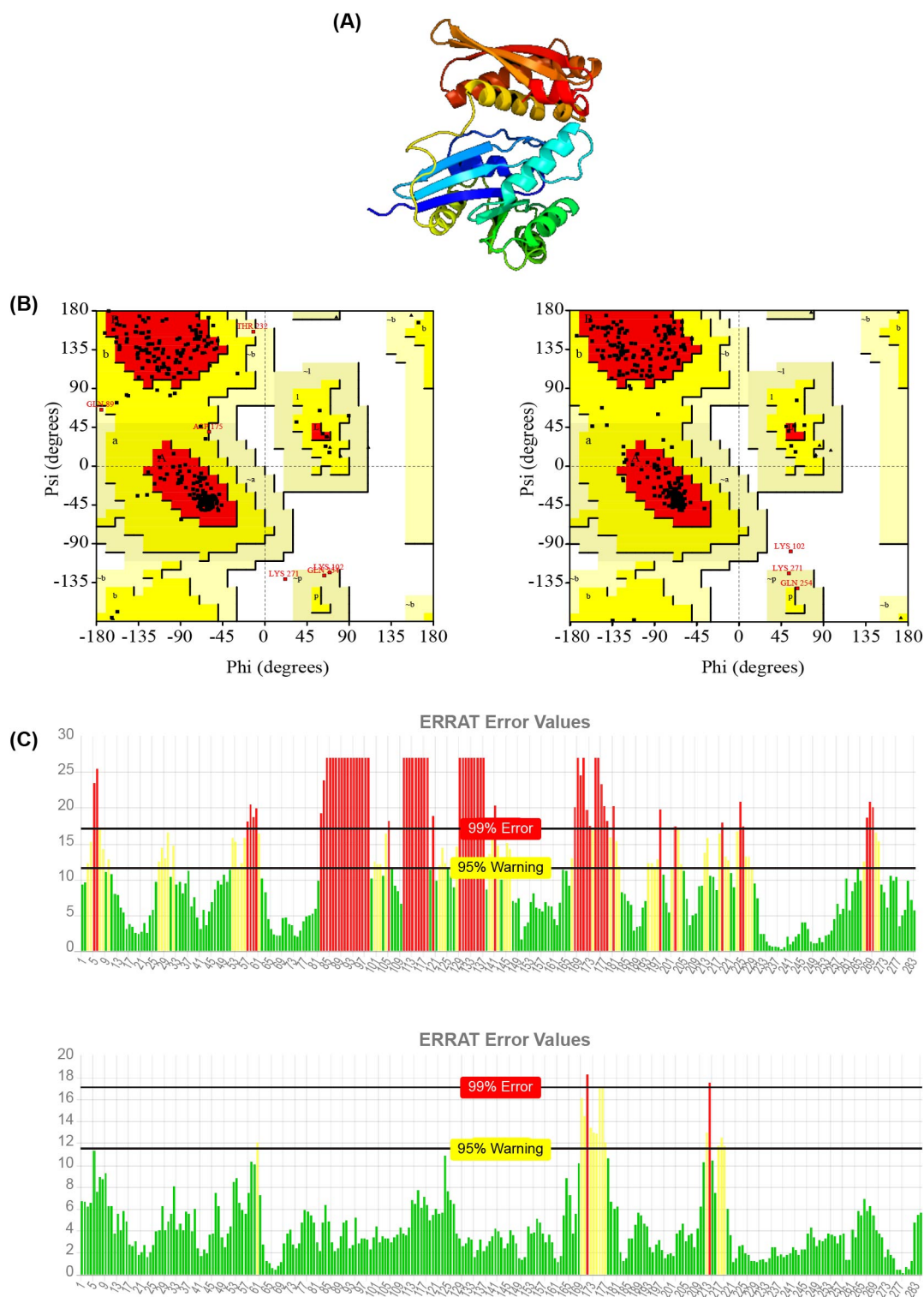
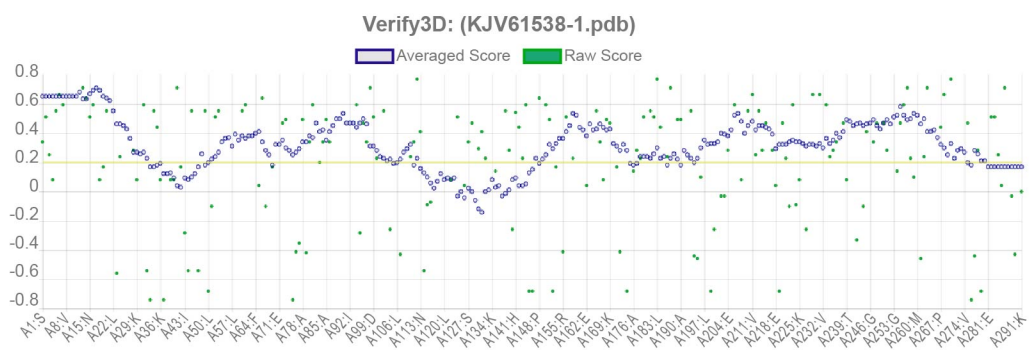
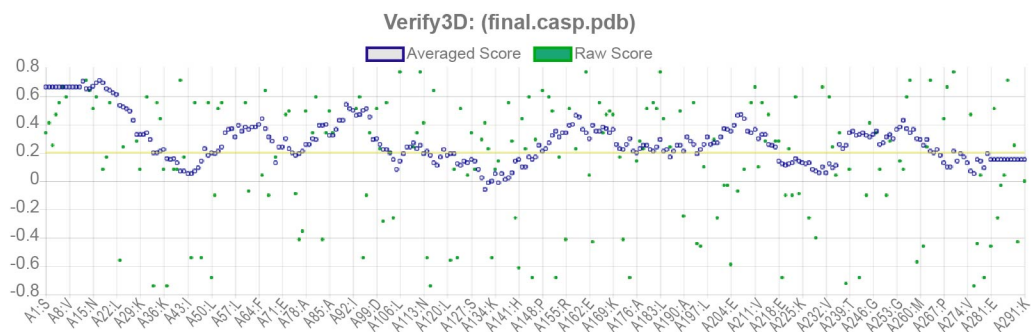


Figura Suplementar 5. (A) Predição do modelo 3D da proteína KJV61538.1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score Z* obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

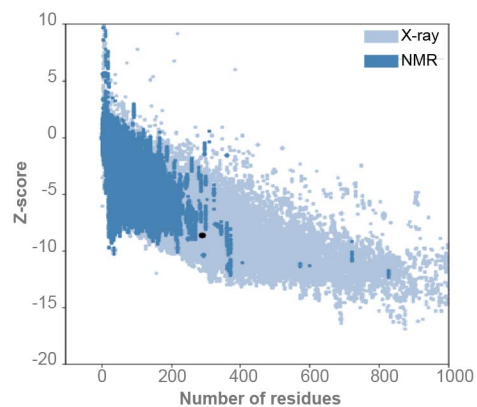
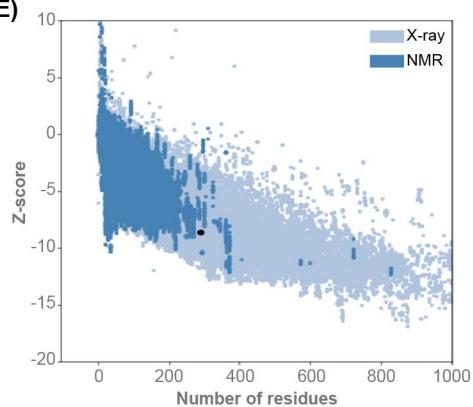
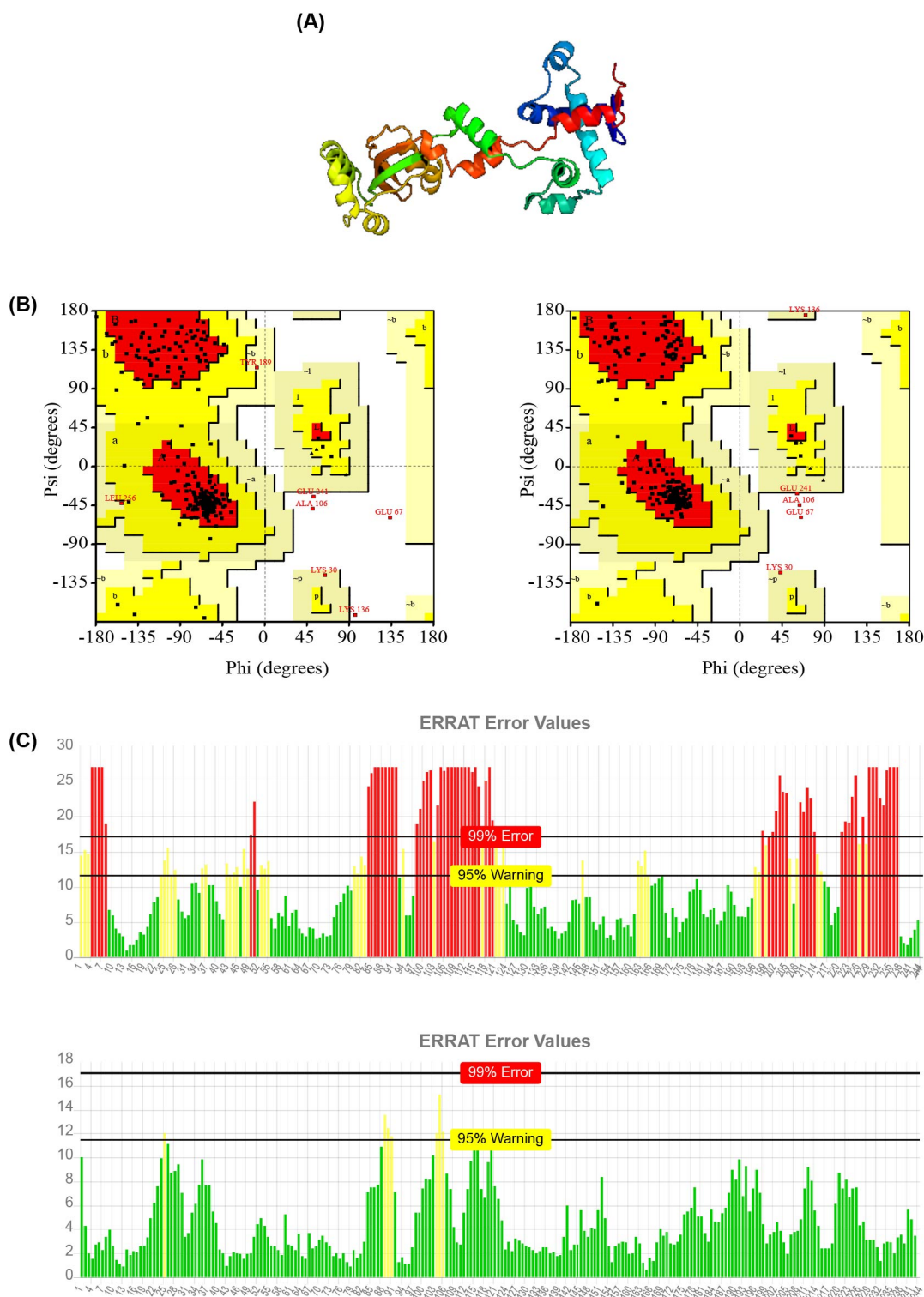
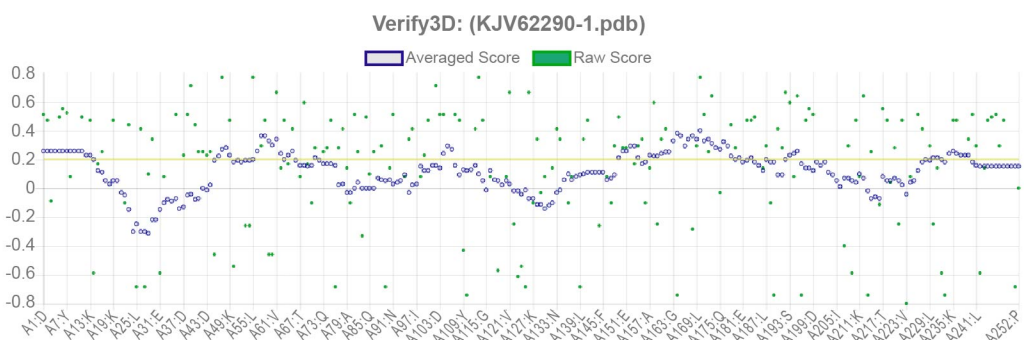
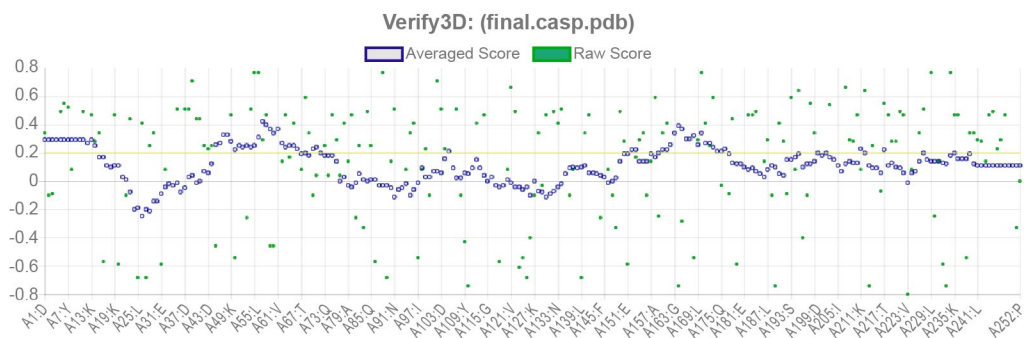


Figura Suplementar 6. (A) Predição do modelo 3D da proteína KJV62290.1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score Z* obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

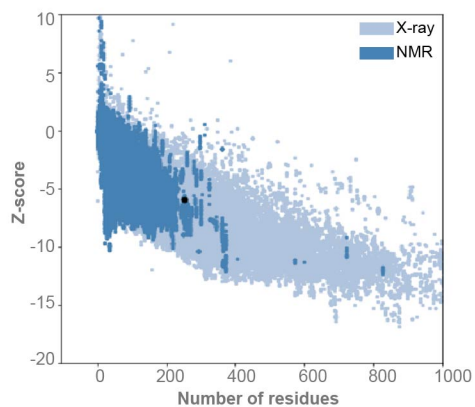
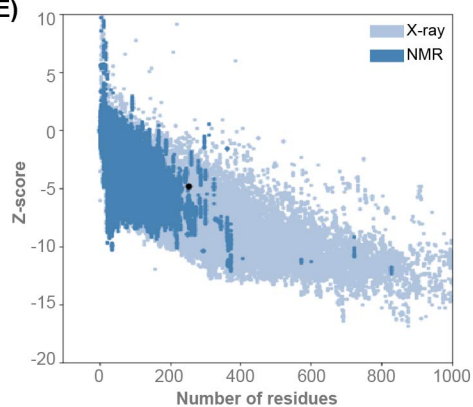
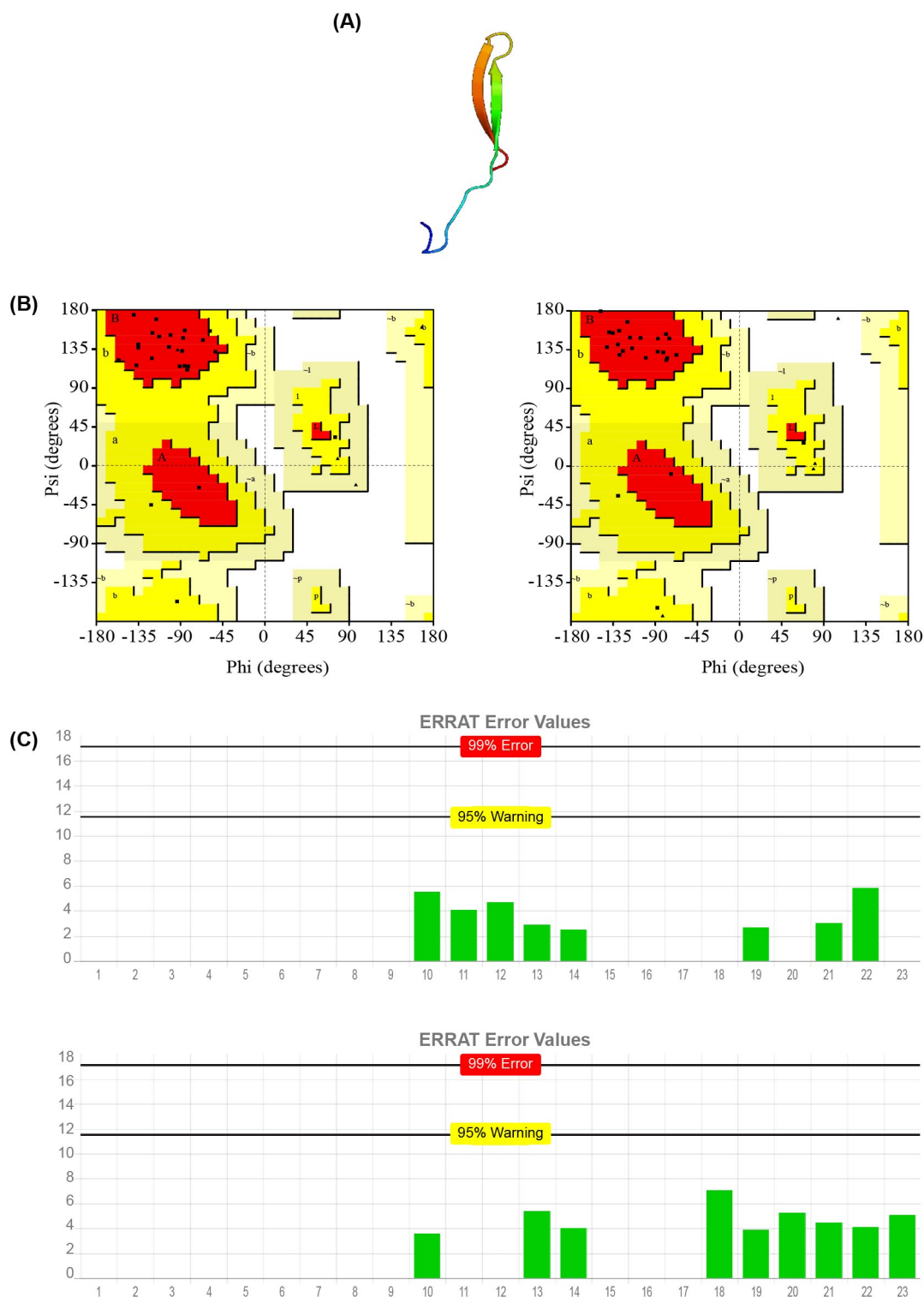
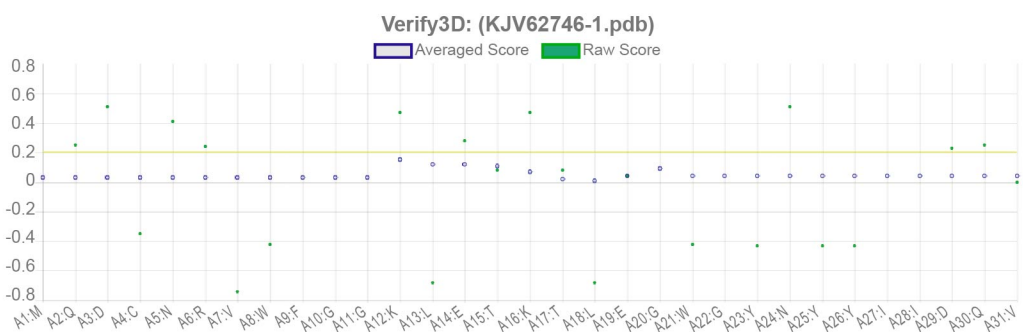
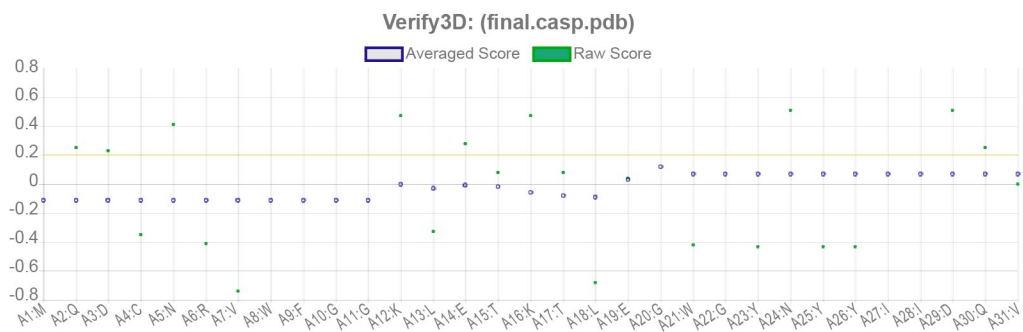


Figura Suplementar 7. (A) Predição do modelo 3D da proteína KJV62746.1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score Z* obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

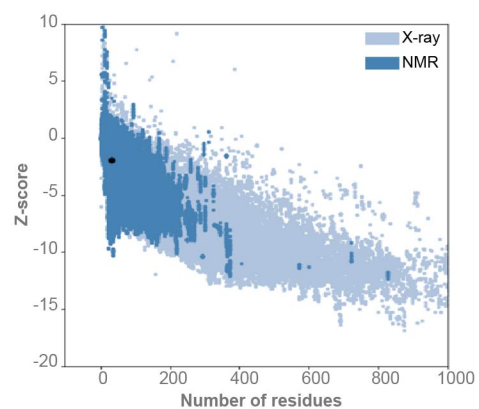
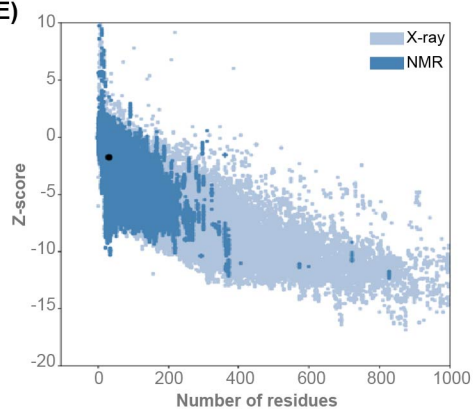
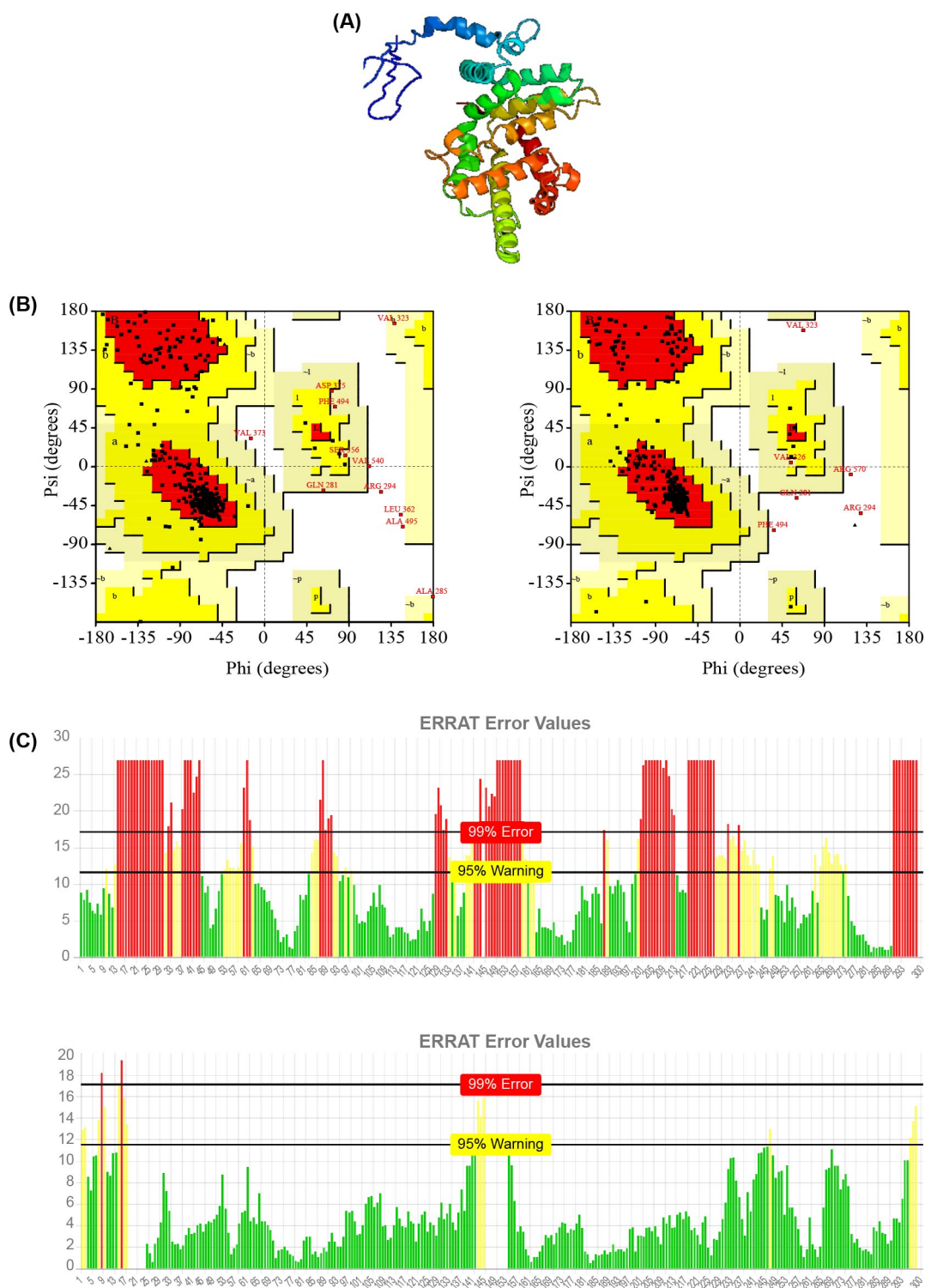
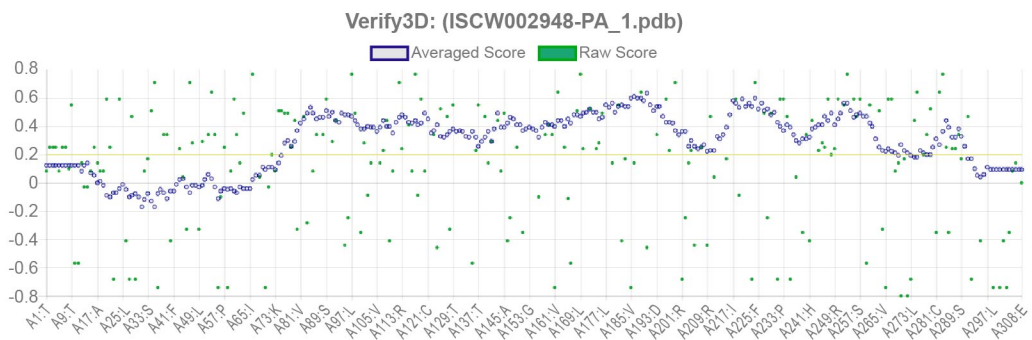
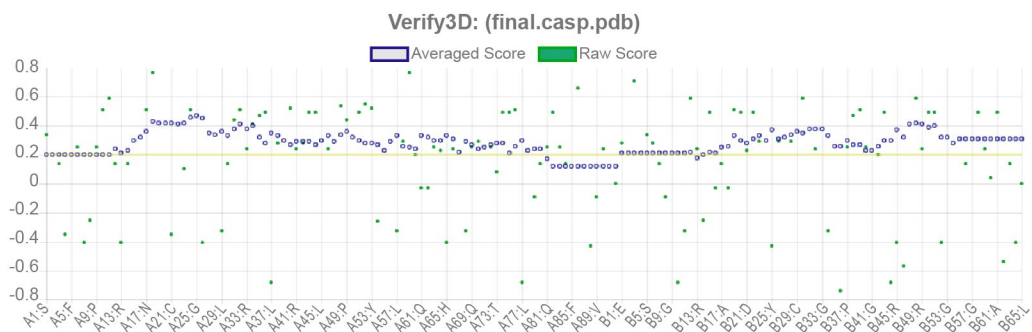


Figura Suplementar 8. (A) Predição do modelo 3D da proteína ISCW002948-PA.1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score* Z obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

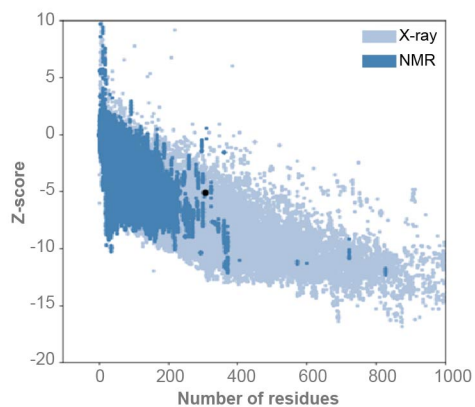
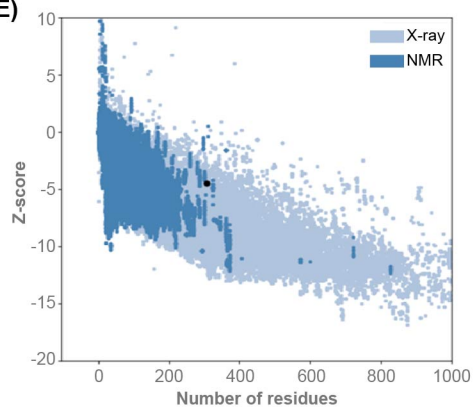
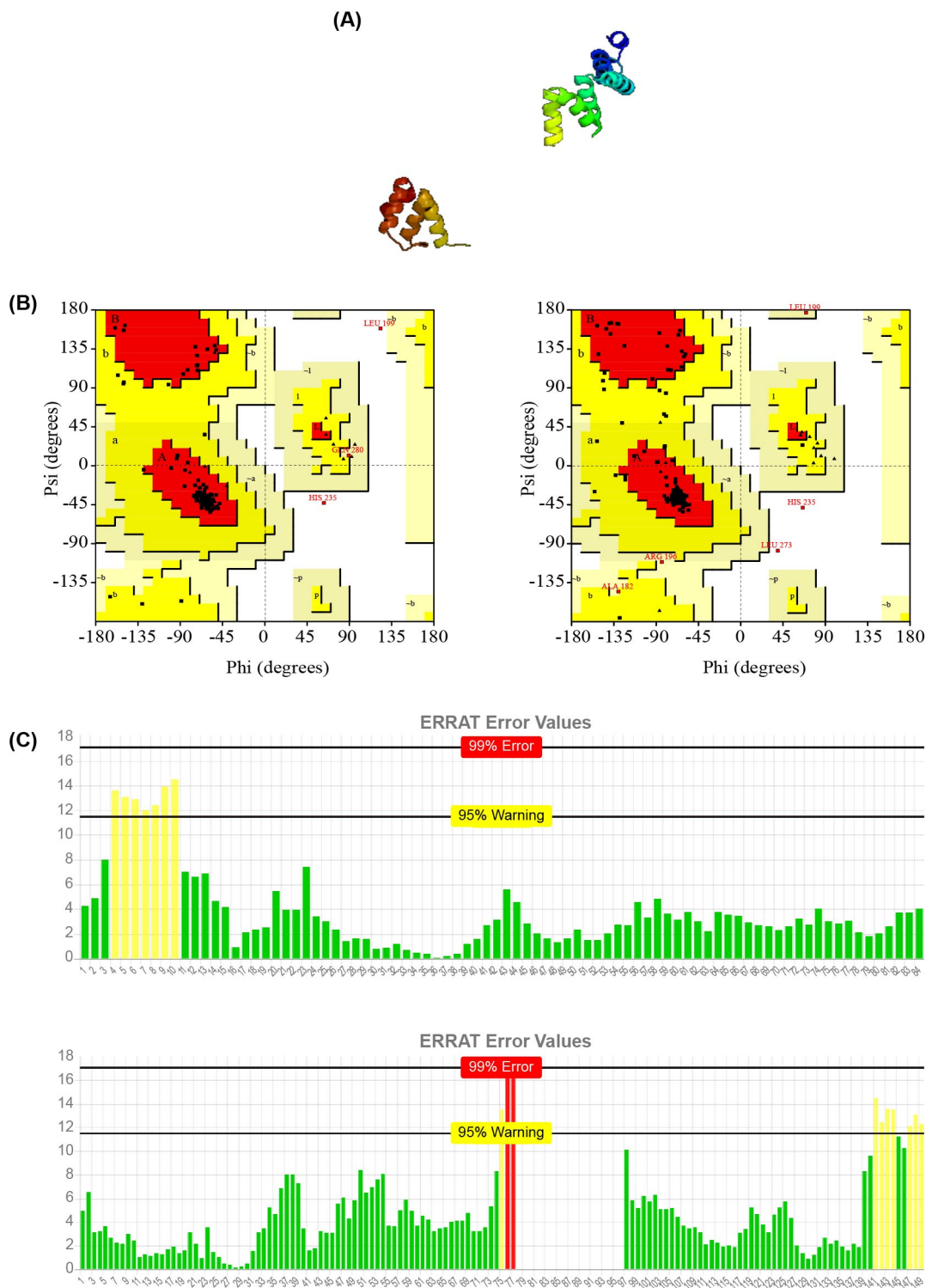
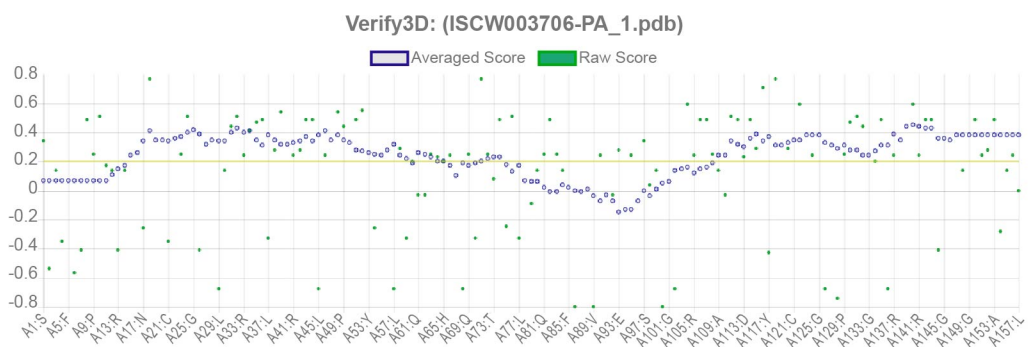
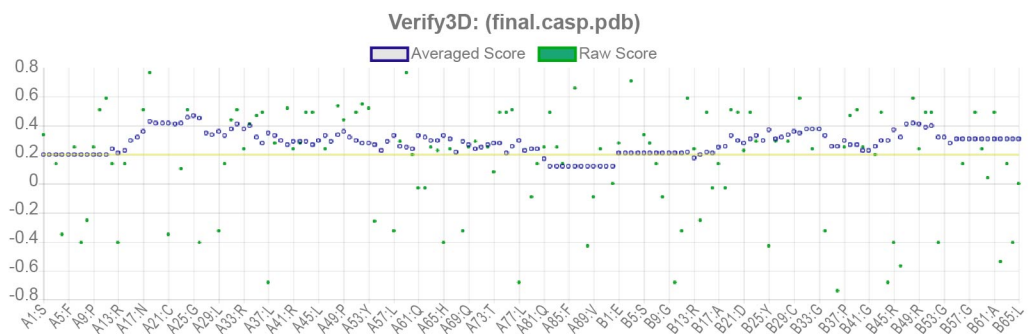


Figura Suplementar 9. (A) Predição do modelo 3D da proteína ISCW003706-PA.1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score* Z obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

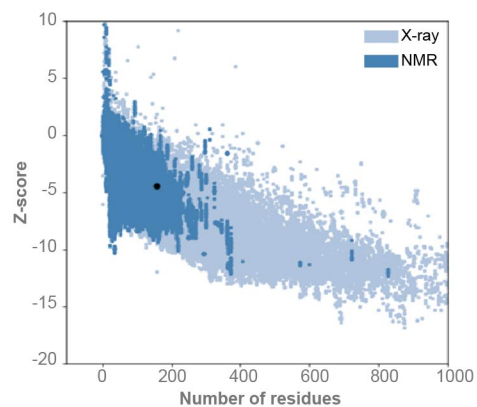
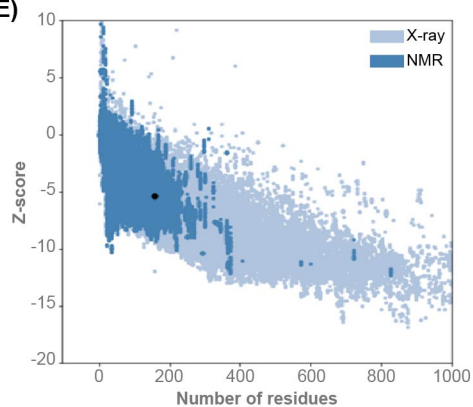
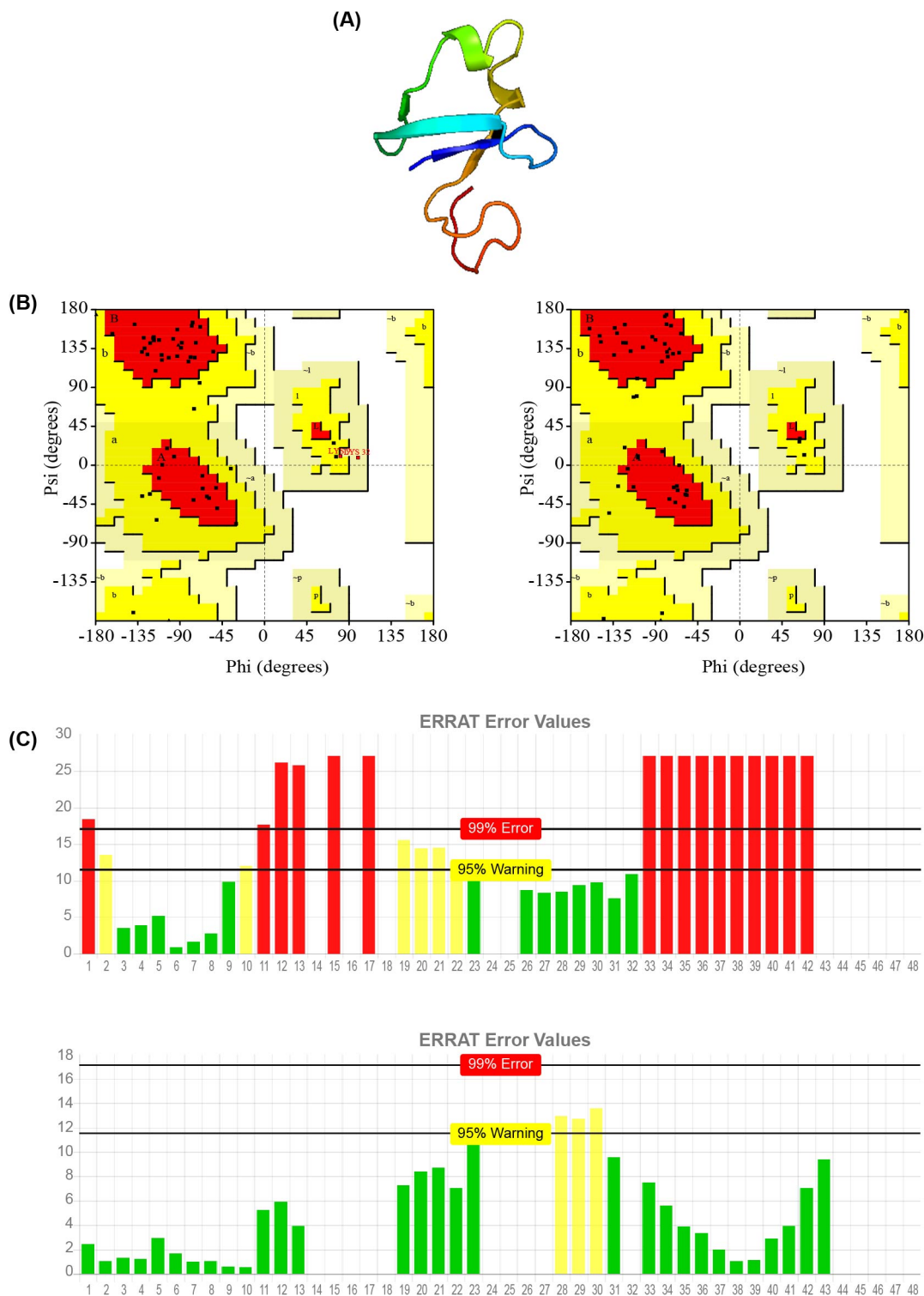
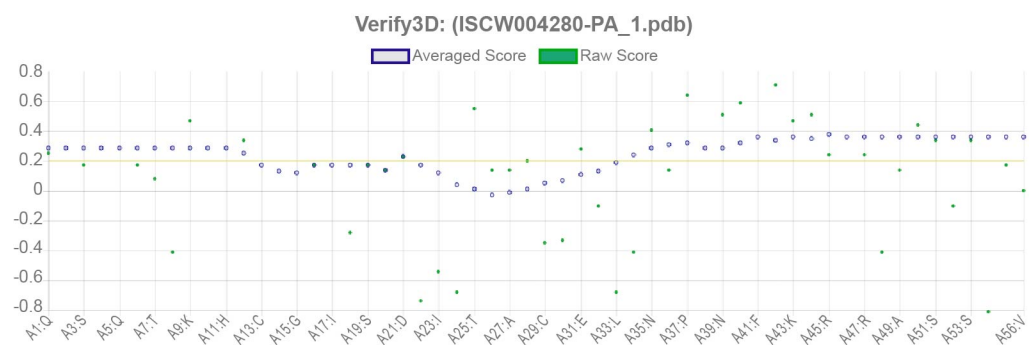
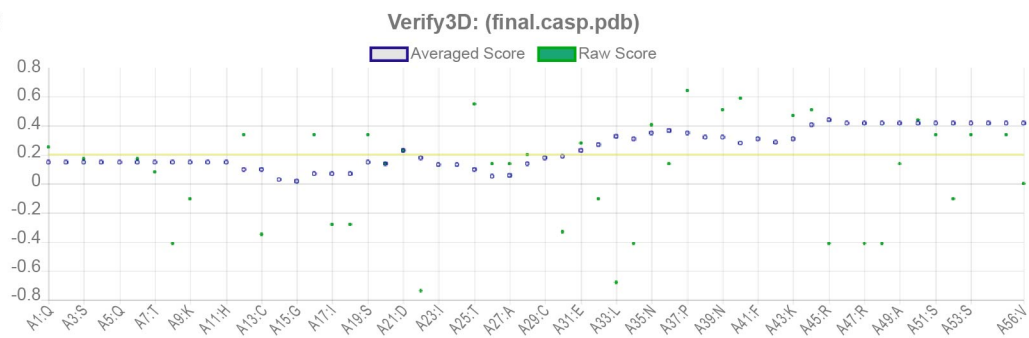


Figura Suplementar 10. (A) Predição do modelo 3D da proteína ISCW004280-PA.1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score* Z obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

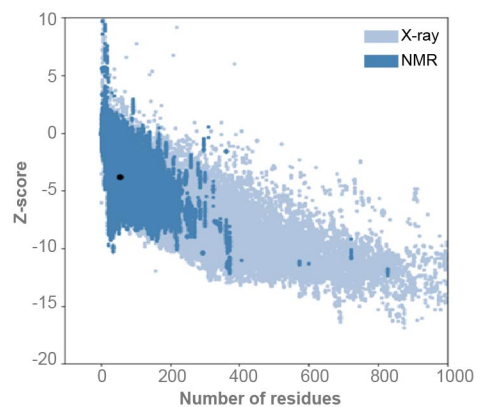
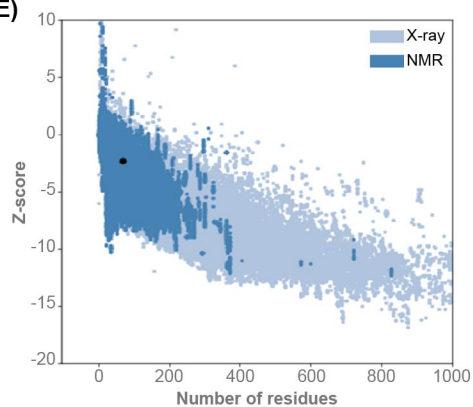
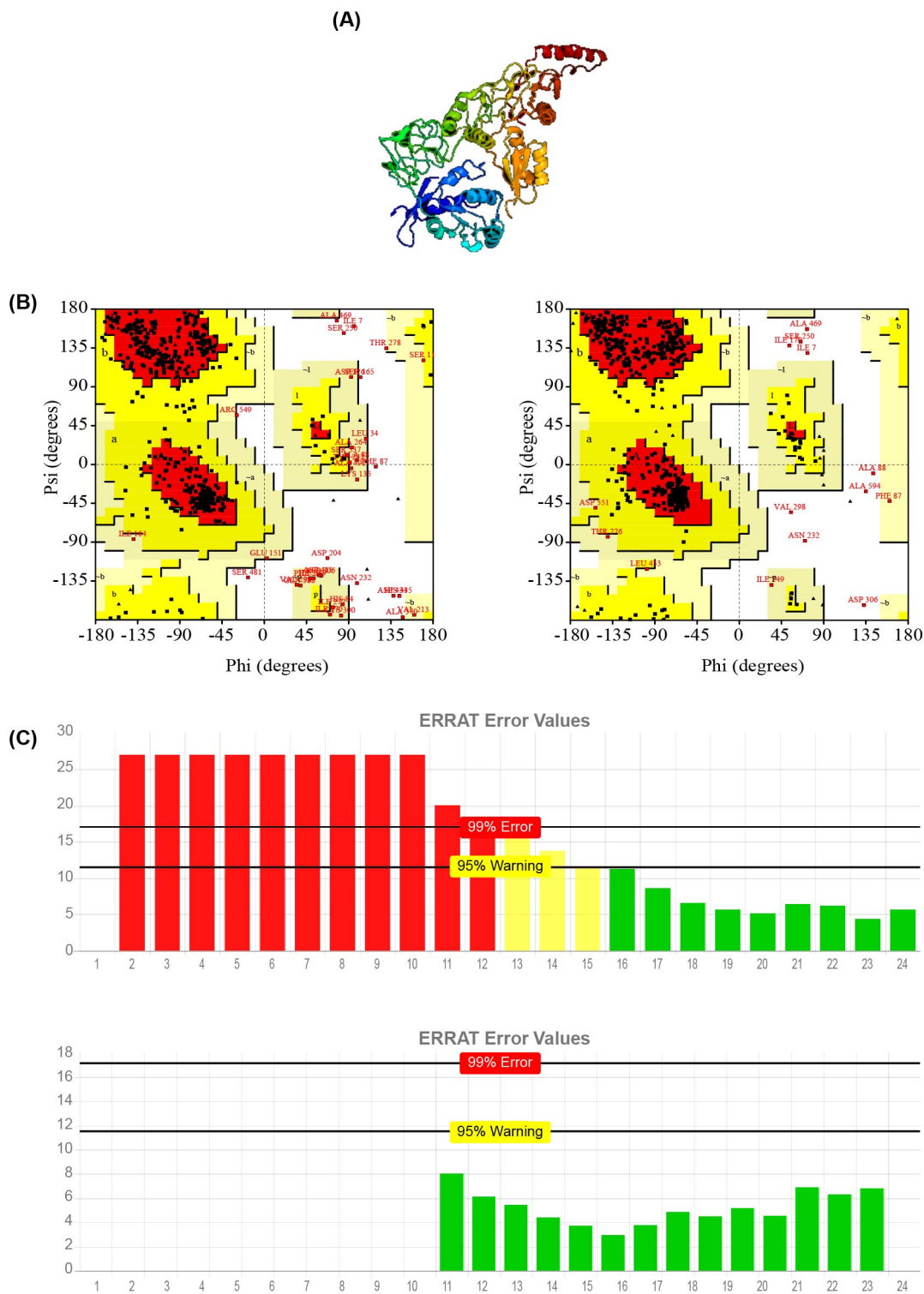
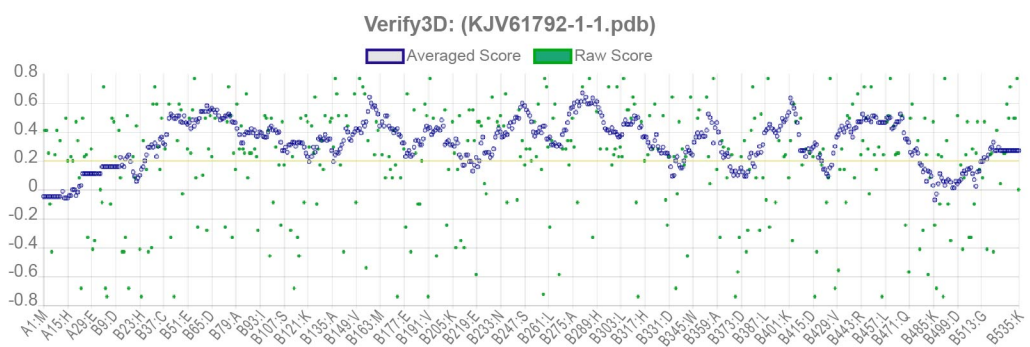
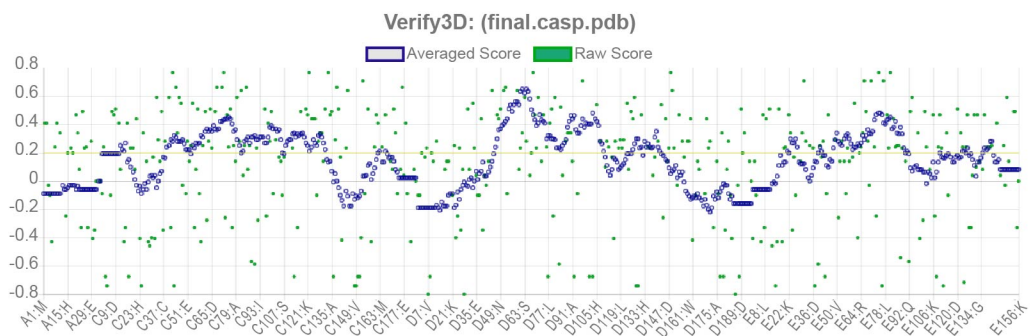


Figura Suplementar 11. (A) Predição do modelo 3D da proteína KJV61792-1 pelo aplicativo Phyre2; (B) Ramachandran plot antes e após a minimização de energia, sendo regiões mais favorecidas, permitidas, generosamente permitidas e não permitidas (região vermelha, amarela escura, amarela clara e branca, respectivamente); (C) Fator de qualidade obtidos com a utilização do aplicativo ERRAT antes e após a minimização de energia; (D) *Score* dos resíduos obtidos com a utilização do aplicativo Verify-3D, destacando limite do *Score* $\geq 0, 2$, representado pela linha amarela; (E) *Score Z* obtidos com a utilização do aplicativo ProSA, com o *score* representado pelo ponto preto.



(D)



(E)

